Surgical management of burn wounds

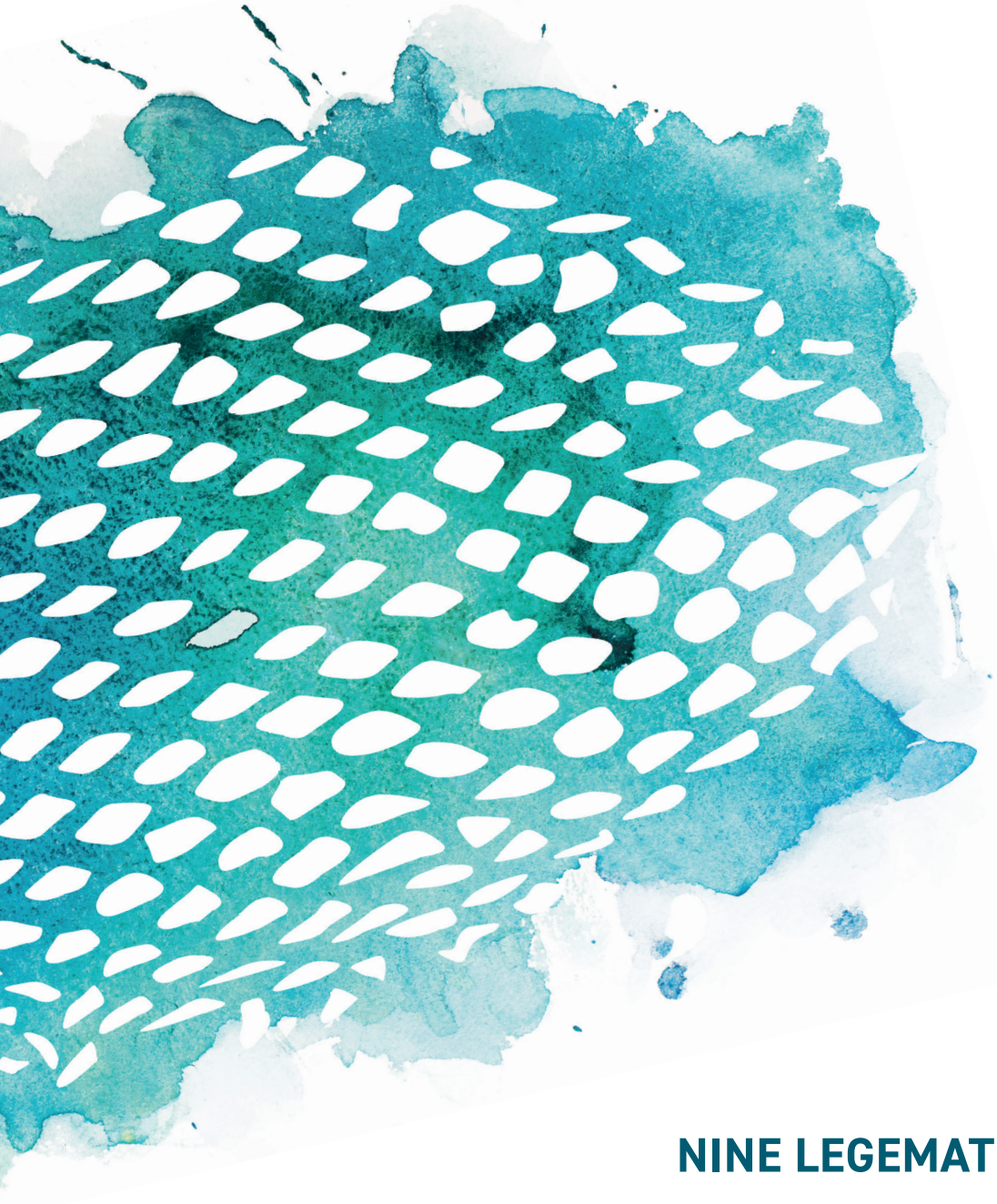
Legemate, Catherine Maria

2022

document version
Publisher's PDF, also known as Version of record

SURGICAL MANAGEMENT OF BURN WOUNDS

CLINICAL AND PATIENT-REPORTED OUTCOMES



NINE LEGEMATE

SURGICAL MANAGEMENT OF BURN WOUNDS

CLINICAL AND PATIENT-REPORTED OUTCOMES

Nine Legemate

Design & layout: Bregje Jaspers | ProefschriftOntwerp.nl
Printed by: Ridderprint BV
ISBN: 978-94-6458-098-3

The research in this thesis was embedded in Amsterdam Movement Sciences Research Institute, at the Department of Plastic, Reconstructive and Hand Surgery, Amsterdam UMC, Vrije Universiteit Amsterdam, the Netherlands.
Financial support for printing this thesis was provided by: Stichting Brandwonden Research Instituut/HUMECA, Nederlandse Vereniging voor Plastische Chirurgie, Nederlandse Brandwonden Stichting, Junior Vereniging Plastische Chirurgie, Amsterdam Movement Sciences Research Institute, Maasstad academie, Smit + Nephew, WCS Kenniscentrum wondzorg.

VRIJE UNIVERSITEIT

SURGICAL MANAGEMENT OF BURN WOUNDS

CLINICAL AND PATIENT-REPORTED OUTCOMES

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op dinsdag 22 maart 2022 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Catherine Maria Legemate

geboren te Utrecht

promotor:	prof.dr. E. Middelkoop
copromotoren:	dr. C.H. van der Vlies dr. M.E. van Baar
promotiecommissie:	prof.dr. M.B. Bouman prof.dr. C.C. Breugem prof.dr. M.H.J. Verhofstad prof.dr. A.E.A.M. Weel - Koenders dr. E. Bosma

CONTENTS

1

Introduction & Outline

Burns are among the most devastating injuries encountered in medical practice. Each year more than 11 million people suffer from burn injuries worldwide. It is estimated that burns cause 180.000 deaths annually throughout the world of which the vast majority occur in low and middle-income countries.[1] In the Netherlands, about 92.000 people visit the general practitioner with burn injuries every year. Approximately 3.900 people are treated at an emergency department because of burns and 1.100 people are admitted to a hospital.[2, 3] Of these, 800 people are treated in one of the three Dutch burn centers.[4]

TREATMENT OF BURNS

Burn wound treatment can be roughly divided in non-operative and operative care. Superficial burns can be adequately treated non-operatively with various dressings and will heal with few or no functional and aesthetical problems.[5] Deep partial and full thickness burns, however, can result in functional and aesthetical problematic scars if re-epithelialization (wound healing) does not occur within two to three weeks.[6, 7] Therefore, these burns are often treated with debridement of the burned tissue followed by autologous split skin grafting. In the majority of hospitals sharp tangential (i.e. layer after layer) excision of the burned tissue with a guarded knife is the standard of care for wound debridement. To cover the wound, donor skin is grafted form a healthy part of the patients' body with a dermatome that leaves sufficient residual epidermal cells to allow spontaneous re-epithelization.[8] This area is referred to as the donor-site and is commonly situated on the patients' thigh, back or upper arm.[9] In addition to the burn scar there will be a scar on the donor-site.

With the evolution of burn wound debridement techniques such as hydrosurgical and enzymatic debridement, sharp tangential excision is no longer the only option.[10, 11] In contrast to conventional surgical debridement, one of the main advantages of these techniques is that they accurately debride burned tissue with maximal preservation of viable dermis and prevent unnecessary tissue loss. A recent review showed that next to increased dermal preservation, these techniques seem to reduce complete healing time and the necessity of skin grafting and therefore have the potential to improve scar outcomes.[11] However, the evidence was of limited quality.

OUTCOMES OF BURN INJURIES

Over the past decades, the number of people surviving burns has increased substantially due to advances in critical and surgical care management.[12-14] Therefore, more patients are confronted with lifelong disabilities, disfigurements and often subsequent psychological problems.[15] This has led to a shift in attention from clinically-led short-term outcomes (e.g. improvement of survival)

to longer-term patient-reported outcome measures (PROMs), which focus on psychological sequelae and physical appearance.[16] PROMs are questionnaires that are completed by patients themselves to measure their perspectives on disease and healthcare outcomes. They provide important insights into what truly matters to (specific) patients since they are directly assessed by the patient without the interpretation by another persona. This information can be used to inform patients, manage expectations of certain treatment options and support shared decision making. Also, outcomes can raise awareness of potential complications or problems during recovery and help clinicians to tailor treatment. In modern day burn research, particularly health related quality of life (HRQL) and scar quality outcomes are progressively used to improve patient-centred care.[17, 18]

HRQL is a self-reported outcome that reflects the patients' perception of his/her health status, based on physical, psychological and social wellbeing after an injury or disease.[19] Although many different PROMs have been developed, no consensus exists which PROM is optimal in the assessment of HRQL after burn injury.

Scar quality measurement can be performed by different types of instruments, including scar assessment scales and devices. Devices often measure just one scar characteristic (like colour or pliability), while scar assessment scales allow for the evaluation of multiple scar characteristics. Clinician-reported outcomes capture the clinicians' opinion on visual and tactile characteristics, whereas patient-reported outcomes enable the assessment of sensory characteristics in addition to the patients' opinion on the visual and tactile characteristics of their scar(s).

Measurements are crucial in clinical practice and medical research. They form the basis of diagnosis, prognosis and evaluation of the results of medical interventions. A decision-maker should know that the measurement instrument used is adequate for its purpose, how it compares with similar instruments and how to interpret the results it produces. When multiple measurement instruments are available, choosing the most appropriate instrument can be challenging. Clinimetrics is a methodological discipline which focusses on the quality of measurements in medical research and clinical practice. Recently, COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) have been published to guide the selection of an appropriate measurement instrument. These guidelines advice that selection should be based on 1) feasibility of use (e.g. availability, patient compliance), 2) the clinimetric properties (e.g. reliability, validity), and 3) interpretability of results (e.g. normative values, minimal important change).[20-22]

AIMS AND OUTLINE OF THIS THESIS

Research in this thesis focusses on short- and long term patient-reported outcomes of burn care and clinical outcomes of burn surgery. An overview of available evidence on HRQL after burn injury was provided (Part I), clinimetric properties of PROMs used in burn care were assessed (Part II) and clinical studies into outcomes of burn surgery were performed (Part III). The aim was to critically appraise available evidence of often used measurement instruments and to improve surgical burn care through the evaluation of available surgical techniques.

PART I Health-Related Quality of Life after burn injury

PROMs to assess HRQL after burn injury are either generic (assessing general aspects of health) or disease specific (covering aspects that are specifically relevant for burn patients). Generic instruments allow comparison with the general population and other diseases, whilst burn-specific instruments include disease specific items that are relevant for burn patients e.g. body appearance.

We conducted an extensive systematic literature review in order to reveal and summarize the current knowledge on HRQL after burn injuries. Chapter 2 provides a comprehensive overview of all generic and burn-specific instruments that have been used to measure HRQL in adult burn patients. This study also examines recovery patterns of HRQL after burn injury. The identification of specific predictors of HRQL after burn injury is essential to optimize the treatment and rehabilitation of these patients. Information regarding predictors may help caregivers to select patients that need specific care or to identify those who require special attention during their treatment. All the available evidence on predictors of HRQL in adult burn patients is presented and summarized in Chapter 3.

Although HQRL assessment in adult burn patient has gained interest, children and adolescents form a special subgroup. With 43% of patients admitted to the Dutch burn centres being younger than 19 years of age and infants <5 years being the most frequently admitted age category, these patients need special attention.[4] Especially, since more children have to deal with life-long consequences of burns due to improved survival.[23-25] Results of the first systematic review on study design, instruments, methodological quality, outcomes and predictors of HRQL in children after burn injury are presented in Chapter 4.

PART II Clinimetric studies on outcomes

Results of part 1 of this thesis show that there are many measures of HRQL and that there is no standard to measure HRQL after burn injury. To make a recommendation on the most suitable HRQL measure for burn patients, it is valuable to assess which HRQL instrument has the best measurement properties. Measurement properties are quality aspects of a measurement

instrument and provide information on whether the results obtained by an instrument can be trusted. Hence, HRQL instruments with robust measurement properties in burn patients are required to draw *reliable* and *valid* conclusions about outcomes and, ultimately, provide high quality evidence to improve patient care.

The extended definition of *reliability* is 'the extent to which scores for patients that have not changed are the same for repeated measurement under several conditions'.[26, 27] The more outcomes vary between repeated measures without true changes of the outcome, the lower the reliability is. This variation can be due to the measurement instrument itself, the person performing the measurement, the patient who is undergoing the measurement, or circumstances under which the measurement is taken.[27]

The general definition of *validity* is 'the degree to which an instrument measures the construct (in our case HRQL) it purports to measure'.[26] Three types of validity van be distinguished: content validity, criterion validity and construct validity. The *content validity* of a questionnaire arises from assessment of relevance, comprehensiveness, and comprehensibility of items. Evidence of these parameters can be derived from PROM development studies, content validity studies and expert opinion. *Criterion validity* refers to how well the scores of the measurement instrument agrees with the scores of the gold standard. It is important to notice that there is no gold standard for the measurement of HRQL. *Construct validity* is applicable in situations in which there is no gold standard and refers to whether the instrument provides scores based on what you expect (i.e. a lower HRQL in patients after severe burns, compared to patients after minor burns).

Another important measurement property is *responsiveness*, which is 'the ability of an instrument to detect change over time in the construct to be measured'.[27] If a patient changes on the construct of interest, scores on the measurement instrument assessing the construct has to change accordingly. In Chapter 5 the COSMIN methodology and guidelines were used to critically appraise the measurement properties of HRQL instruments in burn patients.[20-22]

Next to robust measurement properties, results of PROMs need to be *interpretable*. Interpretability is 'the degree to which it is clear what scores or change in scores mean' and is not a measurement property because it does not refer to the quality of a measurement instrument. One PROM that is often used as an outcome measure in this thesis is the Patient and Observer Scar Assessment Scale (POSAS). This scale is used to measure scar quality from the perspectives of clinician/researcher ('observer') and the patient and has evidenced adequate measurement properties in burn scars.[28-31]

Key concepts of interpretability are the concept of the minimal important change (MIC) and Minimal Clinically Important Difference (MCID). The MIC represents the smallest change in score

that would be perceived by the patient as important and refers to changes in scores within one patient over time.[32] The MCID represents the smallest difference in score that would be perceived by the patient as important and is used to compare scores between individuals or groups at the same time (i.e. the difference between two trial arms).[33, 34] So, the MIC tends to be used to refer to longitudinal within-person changes in scores, and MCID (which is the equivalent of the minimal important difference (MID)) for cross-sectional between-person differences.[27] If the outcome of a study is less than the MIC or MCID, although statistically significant, it may not represent a clinically relevant alteration to a patient. Therefore, the MIC and MCID represent the smallest change or difference in score in the domain of interest which patients perceive as beneficial and which would therefore mandate, in the absence of troublesome side effects and excessive cost, a change in the management of patients. The aim of Chapter 6 was to determine the MIC and MCID of the POSAS. In addition to an increased interpretability of results of future studies, this helps to serve as a basis for estimating the necessary sample size in designing future studies.

PART III Outcomes of burn surgery

Healthcare professionals are increasingly encouraged to bring scientific evidence into their care planning and decision making. Nevertheless, the choice of the debridement technique that is used to treat deeper burns still depends on the decision of the burn specialists and no algorithm is available to guide them in this decision. Conventional surgical debridement of acute burn wounds consists of sharp tangential excision of non-viable tissue with hand-held knives (Watson knife, the Humby knife, the Goulian or Weck knife, or a simple scalpel).[35] Adequate debridement with these knives is determined by the presence of punctuate bleeding and viable dermis. This procedure is not only associated with substantial blood loss, but also with the unnecessary removal of viable dermis.[36, 37] During the last decade, hydrosurgery has become popular as a new option for excision of non-viable tissue prior to skin grafting.[38] The hydrosurgery system console pressurizes saline solution and generates a jet, emitted from the tip of a hand-held instrument (Figure 1). The high-pressure jet evaporates non-viable tissue from the wound, which is then sucked in the device through a vacuum generated by the Venturi effect.[39, 40] The disposable instrument attaches to a console and is activated by the surgeon using a foot-pedal. The surgeon can adjust the pressure of the fluid jet, allowing for a precise depth of debridement. In this way, an accurate wound debridement with complete removal of non-viable tissue and maximal preservation of healthy tissue can be achieved.[36, 38, 41] There is little or no evidence on beneficial effects of hydrosurgery over conventional knife debridement whilst the instrument costs are higher.[11, 42]

Hydrosurgery for burn wound debridement was introduced in Dutch burn care in 2006 and was widely adopted from 2008 onwards. We analysed data from the Dutch Burn Repository R3 to gain insight into which patients received hydrosurgery in specialized burn care centres

in the Netherlands and to evaluate surgical outcomes compared to conventional debridement techniques. Results of this study are presented in Chapter 7.

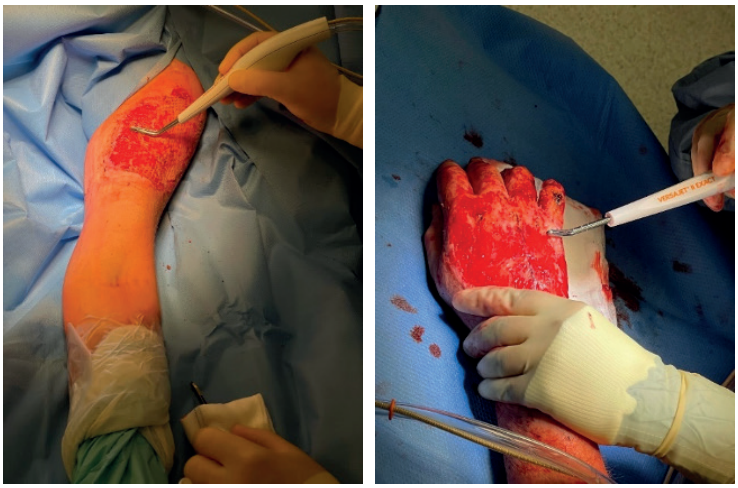
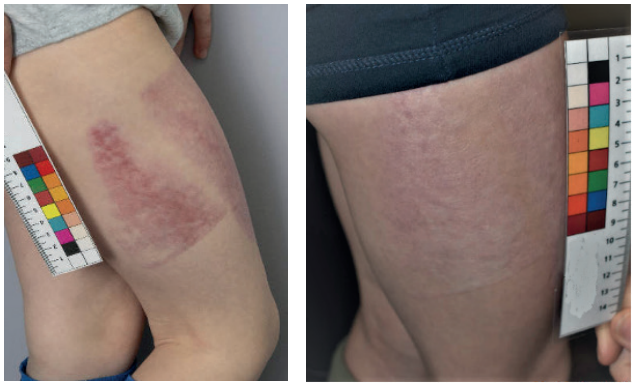


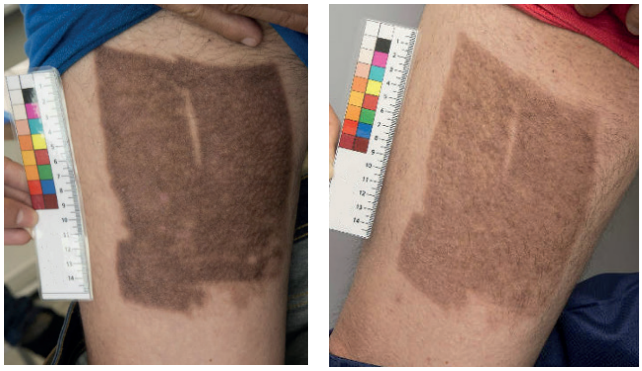
Figure 1. Versajet hydrosurgery system (Smith+Nephew, London UK) for burn wound debridement.

One of the main factors determining the quality of the scar and the degree of contraction of the healing wound is the amount of remaining viable dermis.[43] Thus, debridement methods which maximally preserve dermis are essential. To assess if hydrosurgical debridement leads to the long-standing assumption of maximal dermal preservation with a better scar outcome, we conducted a double-blind within-patient randomized controlled trial (RCT) in the three Dutch Burn Centres. The study protocol and results of this study are described in Chapter 8 and Chapter 9. Outcomes of this RCT included subjective assessments of scar quality by patients and caregivers next to objective scar quality measures.

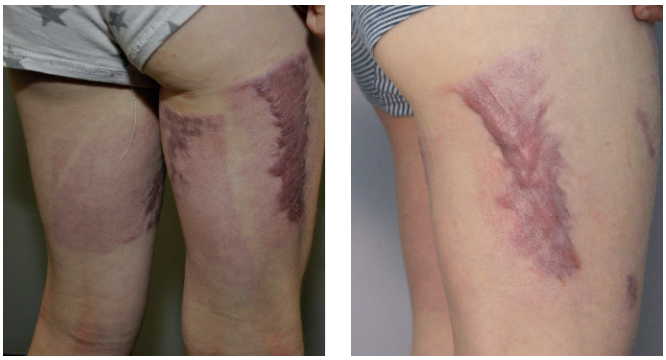
In most of the cases surgical treatment is followed by autologous split skin grafting, leaving a donor-site wound. These donor-sites are known to cause pain, infection, itching and even scarring itself (Figure 2).[44] Although scar quality is currently one of the most important outcomes of burn surgery and split skin grafting is still the cornerstone in the treatment of deeper burns, a recent systematic review showed that studies that have reported donor-site scar morbidity are scarce and only a few studies have assessed the opinion of the patient by the use of PROMs.[45] If poor patient satisfaction regarding a donor-site is expected, this may justify the use of more costly and/or time-consuming novel methods like tissue engineered skin constructs, or more selective debridement techniques like hydrosurgery or proteolytic enzymes. It might even be an argument to refrain from skin grafting with prolonged wound healing and risk of infection or pathological scarring at the burn-site.



Donor-site scar at the upper leg of a young boy at 3 months post-surgery (left) and 12 months post-surgery (right).



Donor-site scar at the upper leg of an adult male at 3 months post-surgery (left) and 12 months post-surgery (right).



Donor-site scar at the upper leg of a young girl at 6 months post-surgery (left) and 12 months post-surgery (right).

Figure 2. Several examples of (pathological) donor-site scarring

On the contrary, if no problems regarding donors-site scarring are expected, early excision and grafting leads to rapid wound closure and a shorter in hospital stay which may result in better quality of life and lower costs. In recent years, a growing trend has emerged towards more patient involvement in treatment decision making (shared decision making (SDM)). One of the most important steps in the process of SDM is that the physician carefully explains the benefits and harms of all reasonable treatment options. Donor-site scarring is an important harm that patients should be informed about if (early) excision and split skin grafting is considered.

In a prospective cohort study, we followed burn patients up to one year post-surgery to assess scar quality of donor-sites after skin graft harvesting. Other aims were to explore changes in scar quality over time, assess the agreement between patients' and caregivers' judgment of donor-site scars, determine predictors of donor-site scar quality and compare scar quality of donor-sites with scar quality of the recipient site (burn scar). The results are described in Chapter 10 and Chapter 11 and can be used to adequately inform patients of their treatment and manage their expectations.

At the end of this thesis, in Chapter 12, the findings of the presented studies are discussed and future perspectives on burn treatment and research are delineated.

REFERENCES

1. Fact sheet Burns. https://www.who.int/news-room/fact-sheets/detail/burns 2020.
2. Brandwondenrapportage Cijfers 2016: VeiligheidNL; 2018.
3. Zoonen E, Koppes L, van Baar ME, van Schie C, Verheij R. Brandwonden bij de huisarts vóórkomen, risicofactoren en behandeling. Utrecht: Nivel.
4. Dokter J, Vloemans AF, Beerthuizen GI, van der Vlies CH, Boxma H, Breederveld R, Tuinebreijer WE, Middelkoop E, van Baar ME, Dutch Burn Repository G. Epidemiology and trends in severe burns in the Netherlands. *Burns* 2014;40(7): 1406-1414.
5. Wasiak J, Cleland H, Campbell F, Spinks A. Dressings for superficial and partial thickness burns. *Cochrane Database of Systematic Reviews* 2013(3).
6. Cubison TCS, Pape SA, Parkhouse N. Evidence for the link between healing time and the development of hypertrophic scars (HTS) in paediatric burns due to scald. *Burns* 2006;32(8): 992-999.
7. Deitch EA. A POLICY OF EARLY EXCISION AND GRAFTING IN ELDERLY BURN PATIENTS SHORTENS THE HOSPITAL STAY AND IMPROVES SURVIVAL. *Burns* 1985;12(2): 109-114.
8. Ratner D. Skin grafting. *Seminars in Cutaneous Medicine and Surgery* 2003;22(4): 295-305.
9. Coull A. Making sense of ... split skin graft donor sites. *Nursing times* 1991;87(40): 52-54.
10. Kwa KAA, Goei H, Breederveld RS, Middelkoop E, van der Vlies CH, van Baar ME. A systematic review on surgical and nonsurgical debridement techniques of burn wounds. *J Plast Reconstr Aesthet Surg* 2019;72(11): 1752-1762.
11. Edmondson SJ, Murray A. Is the time right to put down the knife? A letter in response with a call for change in primary research outcomes to tackle the global burden of long-term scarring. *Burns* 2018;44(7): 1860-1861.
12. Brusselaers N, Hoste EA, Monstrey S, Colpaert KE, De Waele JJ, Vandewoude KH, Blot SI. Outcome and changes over time in survival following severe burns from 1985 to 2004. *Intensive Care Med* 2005;31(12): 1648-1653.
13. Bloemsma GC, Dokter J, Boxma H, Oen IM. Mortality and causes of death in a burn centre. *Burns* 2008;34(8): 1103-1107.
14. Tompkins RG. Survival from burns in the new millennium: 70 years' experience from a single institution. *Ann Surg* 2015;261(2): 263-268.
15. van Baar ME, Essink-Bot ML, Oen IM, Dokter J, Boxma H, van Beeck EF. Functional outcome after burns: a review. *Burns* 2006;32(1): 1-9.
16. Guyatt G. Measurement of health-related quality of life in chronic airflow limitation. *Monaldi Arch Chest Dis* 1993;48(5): 554-557.
17. Griffiths C, Guest E, White P, Gaskin E, Rumsey N, Pleat J, Harcourt D. A Systematic Review of Patient-Reported Outcome Measures Used in Adult Burn Research. *J Burn Care Res* 2017;38(2): e521-e545.
18. Carriere ME, Kwa KAA, de Haas LEM, Pijpe A, Tyack Z, Ket JCF, van Zuijlen PPM, de

Vet HCW, Mokkink LB. Systematic Review on the Content of Outcome Measurement Instruments on Scar Quality. *Plast Reconstr Surg Glob Open* 2019;7(9): e2424.

19. Guyatt GH JR, Feeny DH, Parick DL. *Measurements in clinical trials: choosing the right approach*. Lippincott—Raven.
20. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018;27(5): 1171-1179.
21. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27(5): 1147-1157.
22. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, Bouter LM, de Vet HCW, Mokkink LB. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27(5): 1159-1170.
23. Duke JM, Rea S, Boyd JH, Randall SM, Wood FM. Mortality after burn injury in children: a 33-year population-based study. *Pediatrics* 2015;135(4): e903-910.
24. Bakker A, Maertens KJ, Van Son MJ, Van Loey NE. Psychological consequences of pediatric burns from a child and family perspective: a review of the empirical literature. *Clin Psychol Rev* 2013;33(3): 361-371.
25. Sheridan RL, Hinson MI, Liang MH, Nackel AF, Schoenfeld DA, Ryan CM, Mulligan JL, Tompkins RG. Long-term outcome of children surviving massive burns. *JAMA* 2000;283(1): 69-73.
26. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7): 737-745.
27. de Vet HCW TC, Mokknk LB, Knol DL. *Measurement in medicine. A practical guide to biostatistics and epidemiology* (1st edn). Cambridge Univ Press: Cambridge, **2011.**
28. 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(1): 6-18.
29. Draaijers LJ, Tempelman FRH, Botman YAM, Tuinebreijer WE, Middelkoop E, Kreis RW, van Zuijlen PPM. The patient and observer scar assessment scale: A reliable and feasible tool for scar evaluation. *Plastic and Reconstructive Surgery* 2004;113(7): 1960-1965.
30. van de Kar AL, Corion LU, Smeulders MJ, Draaijers LJ, van der Horst CM, van Zuijlen PP. Reliable and feasible evaluation of linear scars by the Patient and Observer Scar Assessment Scale. *Plast Reconstr Surg* 2005;116(2): 514-522.
31. Truong PT, Lee JC, Soer B, Gaul CA, Olivotto IA. Reliability and validity testing of the Patient and Observer Scar Assessment Scale in evaluating linear scars after breast cancer surgery. *Plast Reconstr Surg* 2007;119(2): 487-494.
32. De Vet HCW TC, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge

University Press: Cambridge, United Kingdom, 2011.

33. de Vet HC, Ostelo RW, Terwee CB, van der Roer N, Knol DL, Beckerman H, Boers M, Bouter LM. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2007;16(1): 131-142.
34. Terwee CB, Peipert JD, Chapman R, Lai JS, Terluin B, Cella D, Griffith P, Mokkink LB. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Qual Life Res* 2021.
35. Orgill DP. Excision and skin grafting of thermal burns. *N Engl J Med* 2009;360(9): 893-901.
36. Jeffery SL. Device related tangential excision in burns. *Injury* 2007;38 Suppl 5: S35-38.
37. Gurfinkel R, Rosenberg L, Cohen S, Cohen A, Barezovsky A, Cagnano E, Singer AJ. Histological assessment of tangentially excised burn eschars. *Can J Plast Surg* 2010;18(3): e33-36.
38. Cubison TC, Pape SA, Jeffery SL. Dermal preservation using the Versajet hydrosurgery system for debridement of paediatric burns. *Burns* 2006;32(6): 714-720.
39. (NICE) NIfHaCE. The Versajet II hydrosurgery system for surgical debridement of acute and chronic wounds and burns. Medtech innovation briefing. 2014.
40. Sainsbury DC. Evaluation of the quality and cost-effectiveness of Versajet hydrosurgery. *Int Wound J* 2009;6(1): 24-29.
41. Matsumura H, Nozaki M, Watanabe K, Sakurai H, Kawakami S, Nakazawa H, Matsumura I, Katahira J, Inokuchi S, Ichioka S, Ikeda H, Mole T, Smith J, Martin R, Aikawa N. The estimation of tissue loss during tangential hydrosurgical debridement. *Ann Plast Surg* 2012;69(5): 521-525.
42. Kakagia DD, Karadimas EJ. The Efficacy of Versajet Hydrosurgery System in Burn Surgery. A Systematic Review. *J Burn Care Res* 2018;39(2): 188-200.
43. Mensik I, Lamme EN, Brychta P. Depth of the graft bed influences split-skin graft contraction. *Acta Chir Plast* 2003;45(3): 105-108.
44. Brolmann FE, Eskes AM, Goslings JC, Niessen FB, de Bree R, Vahl AC, Pierik EG, Vermeulen H, Ubbink DT, group Rs. Randomized clinical trial of donor-site wound dressings after split-skin grafting. *Br J Surg* 2013;100(5): 619-627.
45. Asuku M, Yu TC, Yan Q, Boing E, Hahn H, Hovland S, Donelan MB. Split-thickness skin graft donor-site morbidity: A systematic literature review. *Burns* 2021.

PART I
HEALTH-RELATED QUALITY OF LIFE AFTER BURN INJURY

2

Health-related quality of life in adults after burn injuries: a systematic review

I. Spronk, **C.M. Legemate**, I.M.M.H. Oen, N.E. van Loey, S. Polinder, M.E. van Baar

PLoS One. 2018 May 24;13(5):e0197507

ABSTRACT

Objectives

Measurement of health-related quality of life (HRQL) is essential to qualify the subjective burden of burns in survivors. We performed a systematic review of HRQL studies in adult burn patients to evaluate study design, instruments used, methodological quality, and recovery patterns.

Methods

A systematic review was performed. Relevant databases were searched from the earliest record until October 2016. Studies examining HRQL in adults after burn injuries were included. Risk of bias was scored using the Quality in Prognostic Studies tool.

Results

Twenty different HRQL instruments were used among the 94 included studies. The Burn Specific Health Scale–Brief (BSHS-B) (46%), the Short Form–36 (SF-36) (42%) and the EuroQol questionnaire (EQ-5D) (9%) were most often applied. Most domains, both mentally and physically orientated, were affected shortly after burns but improved over time. The lowest scores were reported for the domains 'work' and 'heat sensitivity' (BSHS-B), 'bodily pain', 'physical role limitations' (SF-36), and 'pain/discomfort' (EQ-5D) in the short-term and for 'work' and 'heat sensitivity', 'emotional functioning' (SF-36), 'physical functioning' and 'pain/discomfort' in the long-term. Risk of bias was generally low in outcome measurement and high in study attrition.

Conclusion

Consensus on preferred validated methodologies of HRQL measurement in burn patients would facilitate comparability across studies, resulting in improved insights in recovery patterns and better estimates of HRQL after burns. We recommend to develop a guideline on the measurement of HRQL in burns. Five domains representing a variety of topics had low scores in the long-term and require special attention in the aftermath of burns.

INTRODUCTION

Surviving a severe burn injury is considered a traumatic experience. Due to substantial improvements in burn treatment, an increasing number of patients survive burns[1,2]. This increases the importance of documenting outcomes of burns on both the short- and long-term as a significant number of patients face physical and/or psychological consequences, such as post-traumatic stress symptoms, depression, and limited physical functioning[3-5]. Moreover, disabilities and disfigurement are frequently accompanied with burn injury.

Health related quality of life (HRQL) is an outcome measure that reflects a patient's perception of his or her health condition on physical, psychological and social wellbeing after an injury or disease[6]. In general, HRQL is assessed by questionnaires filled in by patients. HRQL instruments are either generic (i.e. applicable to any illness) or disease-specific. Generic instruments facilitate comparison between different diseases, whereas burn-specific instruments take the specific effects of burns into account[7]. HRQL measurement is increasingly used in both clinical practice and burn research to qualify the impact of burns[3,8]. It may help to tailor aftercare to the patient's need.

Although, some earlier reviews on the HRQL of burn patients have been performed, there is no recent systematic review on this topic. Yoder et al. conducted a systematic review on the evolution of one burn-specific HRQL instrument; the burn specific health scale (BSHS)[9]. Outcomes were, however, not reported. Stavrou et al. only provided a narrative overview of the domains that could be impaired after burns[10,11].

In conclusion, there is a need for a systematic review to identify which HRQL instruments are used in burns and to examine recovery patterns after burns. Therefore, the aims of this review are 1) to identify which generic and burn specific instruments are used for the measurement of HRQL after burn injuries in adults and 2) to examine recovery patterns of HRQL after burns.

METHODS

The present review was conducted and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Statement[12]. The protocol for this systematic review was registered on PROSPERO (ID=CRD42016048065) and is available online (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016048065).

Search strategy and eligibility criteria

A systematic search using terms covering HRQL and burns (Appendix 1) was conducted from the earliest record until October 2016. The search strategy was developed in collaboration with a librarian with extensive experience in systematic reviews. The databases included Embase, Medline, CINAHL, Cochrane, Web of Science and Google scholar. Original research studies conducted in adult burn patients, written in English and published in a peer-reviewed journal were included. Studies were required to have a generic or disease-specific HRQL as outcome measure and burn patients had to be treated at a health care facility. This includes patients that required inpatient hospitalisation, but also patients treated at an emergency department as well as in outpatient care. Studies that included data on other patient groups, in addition to burn patients, and that not present HRQL outcomes for burn patients separately were excluded.

Selection of studies and data extraction

After removal of duplicates, articles were excluded on the basis of title by one reviewer (IS). Two reviewers (IS and CL) independently evaluated a random sample of ten percent of the abstracts. As there was no disagreement between the reviewers, the remaining abstracts were appraised by one reviewer (IS). In case of any doubt, a title or abstract was screened by a second reviewer. Screening of full texts and extraction of data was done independently by two researchers (IS and CL). The titles, abstracts or full texts were evaluated using the inclusion criteria described above. Extracted information included study characteristics, patient characteristics, details on the instruments used to assess HRQL and HRQL outcomes at each assessment point. Disagreements around article inclusion or extraction of data were resolved by discussion with a third researcher (MvB).

Risk of bias assessment

The risk of bias of all eligible studies was assessed using four of the six domains of the Quality in Prognostic Studies (QUIPS) risk of bias tool[13]. We included the domains: study participation, study attrition, outcome measurement and statistical analysis and presentation. Two domains 'prognostic factor measurement' and 'study confounding' were not included as these domains are specific for prognostic studies and thus fell outside the scope of the review. The domains were rated as 'low' bias (all items 'low risk'), 'moderate' bias (max. 50% items with high or unknown risk of bias) or 'high' risk of bias (>50% items high of unknown risk of bias). First, two researchers (IS and CL) were trained to use the QUIPS and independently assessed the risk of bias of eighteen eligible studies (19%)). Discrepancies were discussed with a third researcher (MvB). Then, the researchers independently assessed a random sample of 25 of the remaining articles (33%). There was only a slight disagreement (7%) and therefore the remaining studies were appraised by one researcher (IS). In case of any doubt, a study was appraised by a second reviewer.

Data analysis

In case of multiple studies using an identical dataset, the study that included the most assessment points, the most patients or the most HRQL domains was chosen. If no decision could be made, the most recent publication was selected. If scores were only presented in figures, authors were asked to provide the scores. If authors did not respond, the scores were read from the graph and were rounded to the nearest 0.5 points. If domain scores were only presented as norm scores, authors were asked to provide the non-normalized domain scores. If no scores were received, the outcomes were not included in the recovery pattern analyses. Outcomes of studies were only included when the study population included at least 10 patients.

RESULTS

Identification and selection of studies

The search resulted in 3,788 unique articles. Screening of titles resulted in 255 potentially relevant articles. Of these, 111 were excluded on the basis of abstract and 144 were retrieved for full-text review (Figure 1). Fifty-one of these articles did not meet all inclusion criteria, resulting in the inclusion of 94 articles.

Study characteristics

Most studies were conducted in Europe (n=37), the USA (n=19) and Australia (n=14). More than half (n=54) of the studies were published after 2010 and most had a cross-sectional design (n=57) (Table 1). Sample sizes of the studies varied between 914 and 1,58715 burn patients, with most studies having a sample size below 200 patients (86%). In most studies (n=83) more males than females were included, although not all studies provided details on the sex distribution (n=6)[15-20]. The mean %TBSA burned ranged from 3.5%[21] to 83.5%[22]. Eight studies did not report the %TBSA burned of the included patients. Mean LOS was between 10 and 30 days in most studies. In total, 35 studies failed to report the mean LOS.

Measurement of HRQL

Twenty different instruments, of which eight are validated in the burn population, were used to assess HRQL. The three most often applied instruments were the Burn Specific Health Scale - Brief (BSHS-B) (n=44), the Medical Outcome Study Short Form - 36 items (SF-36) (n=40), and the EuroQol five dimensions questionnaire (EQ-5D) (n=8) (Figure 2). Eight instruments were only used in one single study. Thirty-one studies used more than one instrument to assess the HRQL (Table 1). Twenty-four of these used a burn-specific and a generic HRQL instrument. Most used both the SF-36 and the BSHS-B (n=18). Thirty-two studies (34%) used a longitudinal design with multiple HRQL assessments over time; twenty-three studies used at least four assessment points. The most frequently used assessment time points were during hospital admission, and at 3 months, 6 months, 12 months and at 24 months after injury (Figure 3).

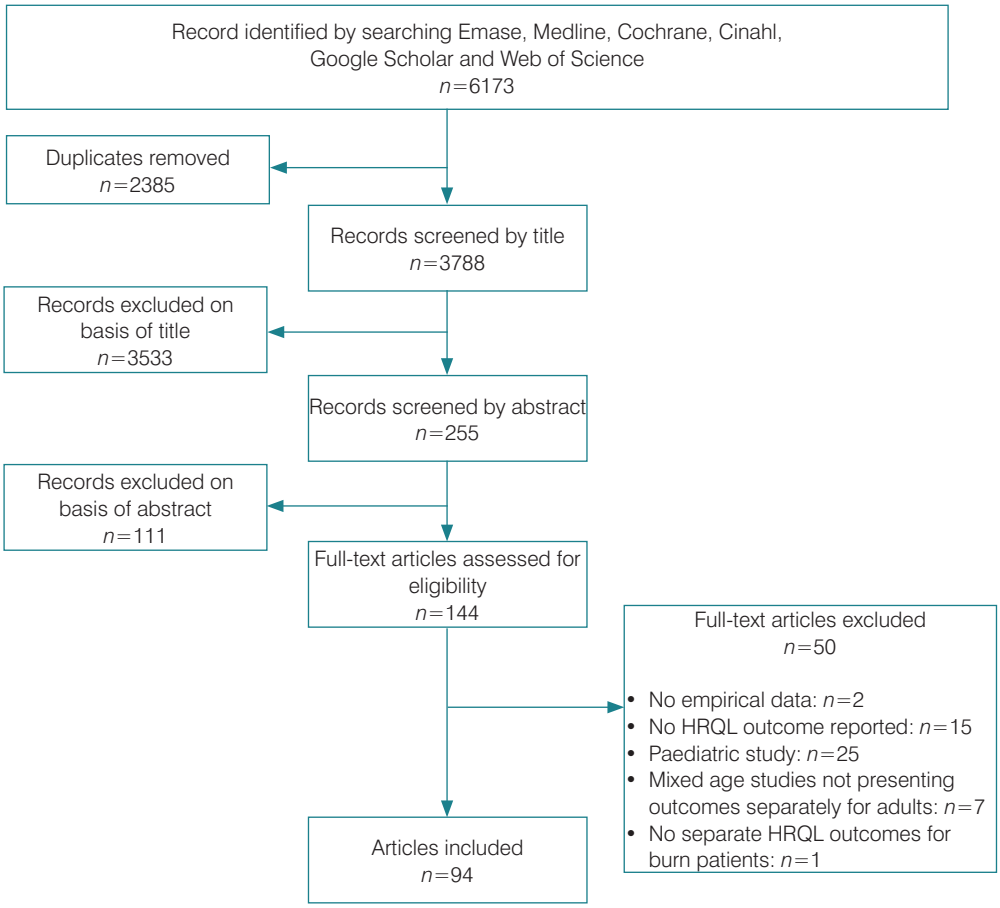


Figure 1. Flowchart outlining selection of studies

Table 1. Study characteristics of 94 studies measuring HRQL in adult burn patients

Study characteristics	Studies (n)	References
Study type		
Case-control	3	14,21,23
Cohort	32	15,20,24-52
Cross-sectional	56	10,16-19,22,53-101
Trial	3	102-104
Patient sample size		
0-20	7	14,22,40,48,58,81,87
>20-50	19	16-18,21,25,33,34,51,54,55,59,60,66,71,72,83,84,96,99
>50-100	30	19,28,30,31,36-39,44,46,50,52,53,56,57,62,64,74,77-80,85,88,90,92,97,104,105
>100-200	25	24,29,35,42,43,47,49,61,63,65,68-70,75,76,82,89,91,93-95,100,102,103,106

Table 1. Continued

>200-500	9	26,41,45,67,73,86,98,101,107
>500	3	15,27,32
NA	1	20
Mean %TBSA burned		
0-10%	9	21,26,27,30,39,51,66,71,86
>10-20%	32	28,29,31,40-48,53,60,63,70,76-80,82,84,88,90,94,95,97,101,102,105,106
>20-30%	26	17-19,24,25,32,35-38,54,58,61,64,65,67-69,72,74,85,91,96,99,103 52
>30-40%	6	16,23,34,73,81,89
>40-50%	5	14,83,93,98,107
>50-60%	4	56,57,87,92
>60-70%	2	55,59
>70-80%	0	-
>80-90%	2	22,100
NA	8	15,20,33,49,50,62,75,104
Mean length of stay (days)		
0-10	2	50,71
>10-20	15	21,26,31,40,47,51,53,54,60,70,86,98,103,106 46
>20-30	29	19,25,29,32,34-38,45,49,52,61,69,74,76-80,84,90,94-97,101,102,105
>30-40	8	18,44,65,67,68,72,87,99
>40-50	1	58
>50-60	0	-
>60-70	1	104
>70-80	1	55
>80	2	22,73
NA	35	14-17,20,23,24,27,28,33,39,41-43,48,56,57,59,62-64,66,75,81,83,85,88,89,92,93,100,107
Number of HRQL instruments		
1 instrument	63	16-21,23-25,28-35,37,38,42-45,50,51,53,54,58,59,62,63,66-72,74,75,79,81,84-87,89-95,97-104,106,107
2 instruments	24	14,22,26,27,40,41,46-49,52,55-57,60,64,65,78,80,82,83,88,96,105
3 instruments	7	15,36,39,61,73,76,77
Number of assessment time points		
1 time point	61	14,16-19,21-23,39,53-101,104,105,107
2 time points	5	28,34,35,46,102
3 time points	4	24,44,52,103
4 time points	12	29,30,33,36,38,42,43,45,47-49,51
5 time points	10	15,25-27,32,37,40,41,50,106
>5 time points	1	31
NA	1	20

Note. NA=not applicable, TBSA=total body surface area

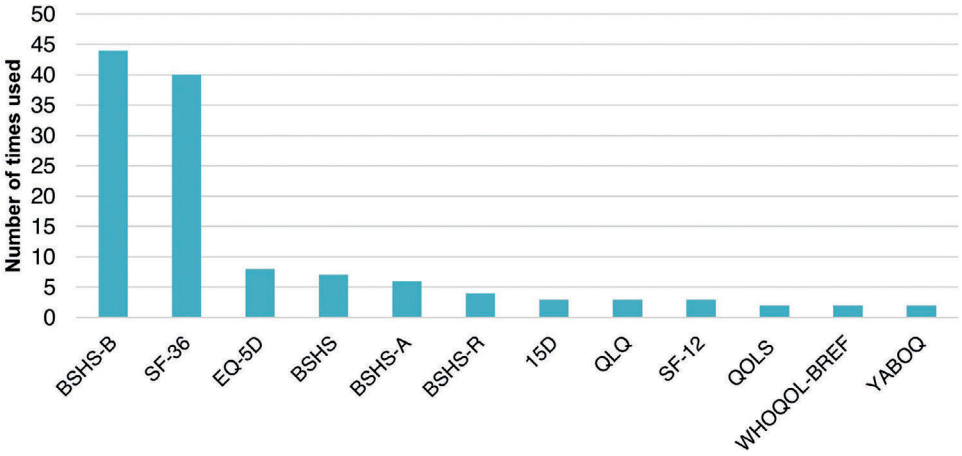


Figure 2. Instruments used to measure health-related quality of life in >1 study.

BSHS-B=Burn Specific Health Scale-Brief, SF-36=Medical Outcome Study Short Form-36 items, EQ-5D=EuroQol five dimensions questionnaire, BSHS=Burn Specific Health Scale, BSHS-A=Burn Specific Health Scale - Abbreviated, BSHS-R=Burn Specific Health Scale Revised, 15D=15-dimensional health-related quality of life instrument, QLQ=Quality of Life Questionnaire, SF-12=Medical Outcome Study Short Form-12 items, QOLS=Quality of Life Scale, WHOQOL-BREF=World Health Organization Quality of Life - BREF, YABOQ=Young Adult Burn Outcome Questionnaire

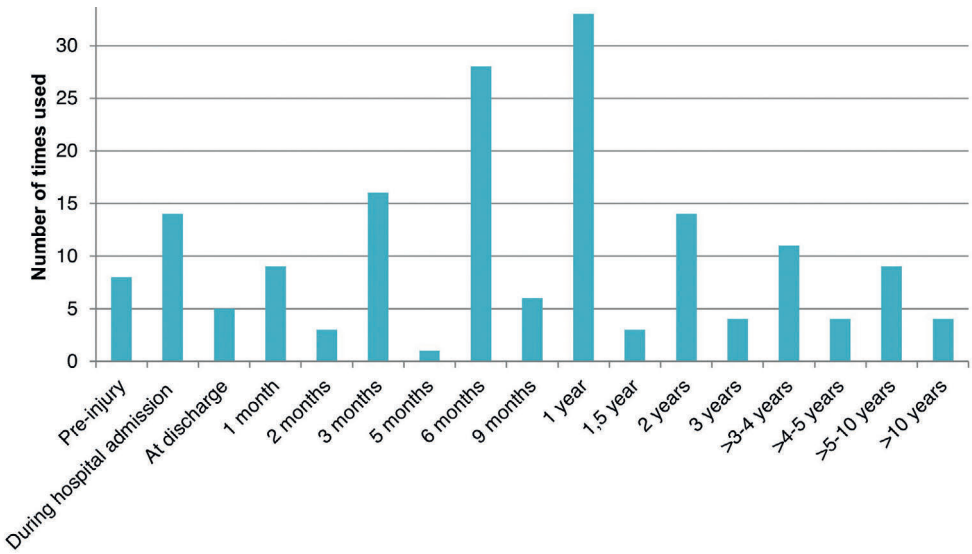


Figure 3. Time points at which health-related quality of life in burn patients was assessed.

Note. Data on pre-burn HRQL is collected retrospectively

Quality assessment

The risk of bias was evaluated using the QUIPS tool. Whilst most studies had low risks of bias on 'outcome measurement' (n=87) and 'statistical analysis and reporting' (n=75), a moderate or high risk was evident in many studies for 'study attrition' (n=88) (Figure 4). This was in particular caused by a lack of reporting of attempts to collect information on drop-outs and of key characteristics of non-responders. Four studies[27,32,33,54] scored a low risk of bias on all four evaluated items of the QUIPS.

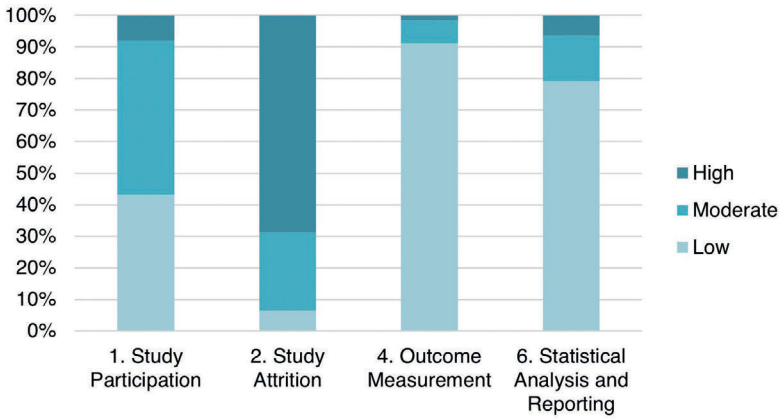


Figure 4. Risk of bias assessed with four domains of the Quality in Prognostic Studies (QUIPS) risk of bias tool

Recovery patterns of HRQL after burns in adults

Recovery patterns of the most applied instruments, the BSHS-B, the SF-36 and the EQ-5D, which are all validated within the burn population, were analysed. All studies that reported a BSHS-B or BSHS-R outcome, a SF-36 outcome or an EQ-5D outcome on at least one time point were included.

BSHS-B recovery patterns

The BSHS-B includes 40 items comprising nine HRQL domains: simple abilities, heat sensitivity, hand function, treatment of regimens, work, body image, affect, interpersonal relationships and sexuality[108]. Responses on individual items are scored on a five-point scale ranging from 0 (extremely) to 4 (not at all). Mean scores per domain were assessed and high scores refer to a good perceived health status. Of the 47 studies that used the BSHS-B or BSHS-R, 17 could be used to analyze HRQL recovery patterns[19,26,38,40,49,60,64,67,76,81-83,88,90,91,102,105]. Overall, shortly after burns, scores on the different domains were low and most increased with time (Figure 5). In the short-term, most problems were reported for the domains 'work' and 'heat sensitivity'.

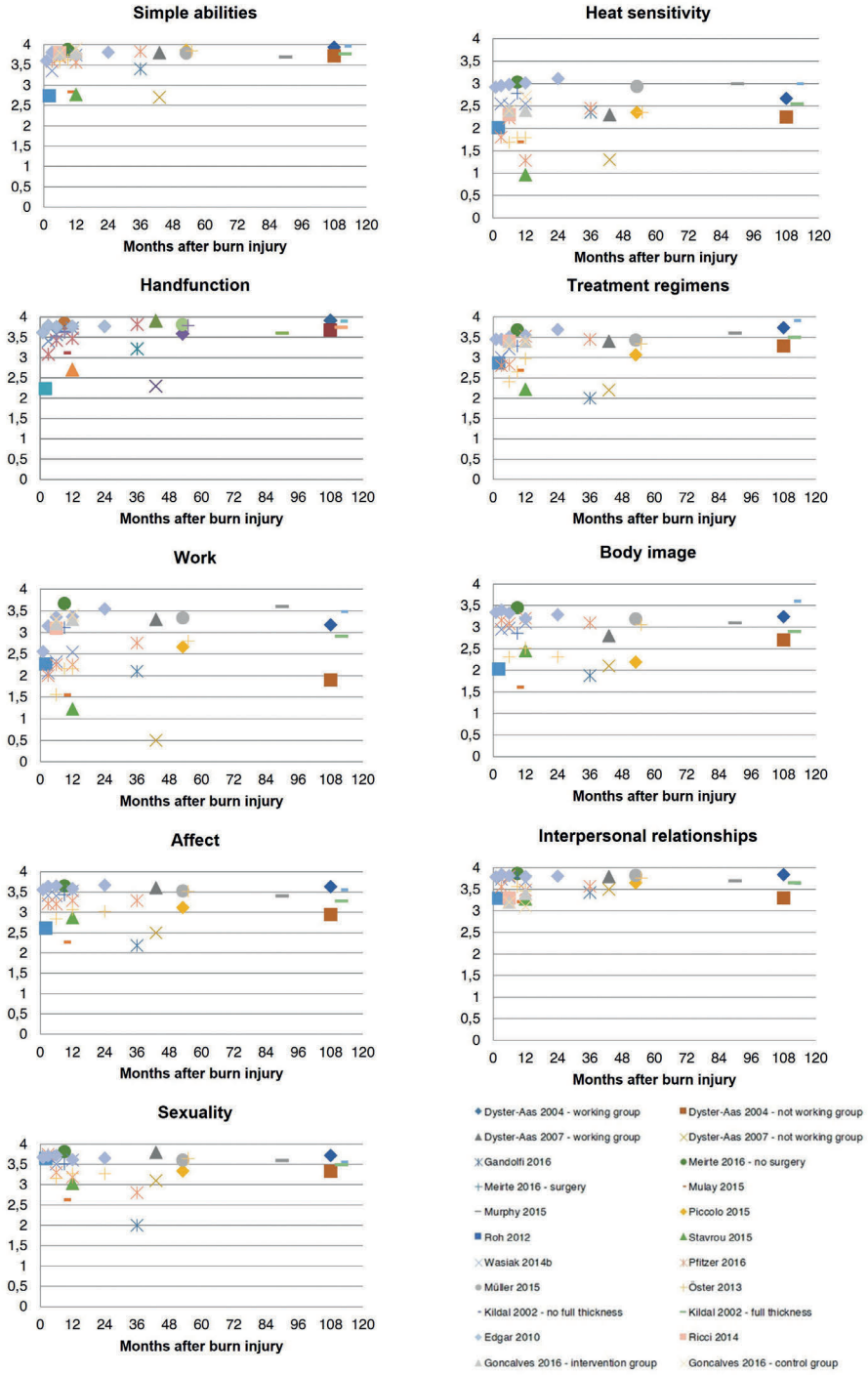


Figure 5. BSHS-B domain scores for seventeen studies.

The self-reported outcomes of the domains 'simple abilities', 'hand function', 'affect', 'heat sensitivity', 'body image' and 'treatment regimens' showed improvement over time. Low scores were especially seen in the first 12 months after burns and improved afterwards. On average, outcomes of the domains 'simple abilities' and 'hand function', improved towards the maximum score, whereas the domains 'affect' and 'treatment regimens' improved to 3.5 out of 4, the domain 'body image' improved towards 3 out of 4 and the domain 'heat sensitivity' towards 2.5 out of 4. The domain 'sexuality' remained relatively stable, only few studies reported somewhat lower scores in the short-term. The outcomes of the 'interpersonal relationships' domain were relatively high during the entire follow-up. The self-reported outcomes of the last domain, the domain 'work', varied widely among studies. In general, subgroups with less severe problems (i.e. no surgery, no full thickness burn) had higher scores on all domains.

SF-36 recovery patterns

The SF-36 consists of 36 items comprising eight domains: physical functioning, role limitations-physical, bodily pain, general health, vitality, social functioning, and role limitations emotional, and mental health. Mean domain scores that were transformed to a 0 (the worst) to 100 (the best) scale were used. Higher scores indicate a greater perceived health. The SF-36 domains can be summarized into the physical component summary (PCS) and the mental component summary (MCS)[109]. These measures are transformed to norm-based scores with a mean of 50 and a standard deviation (SD) of 10. Scores lower than 50 indicate scores below the average. Analyses of recovery patterns of the SF-36 outcome data were based on 17 studies of the 40 studies that assessed HRQL with the SF-36[23,26,28,29,32,35,40,44,49,60,64,66,72,73,76,80,99]. Four out of the 17 studies described all eight domains of the SF-36 as well as the PCS and MCS. Ten studies included the eight domains, one study included seven domains[76], and one study described both summary scores[32]. The MCS scores showed variation in the short-term, with studies reporting scores just above and below the norm score (Figure 6). In the longer-term, scores moved towards the norm score. PCS scores were almost all below the norm score and an improvement towards the norm was seen in the longer-term. The lowest scores were reported for the domains 'bodily pain' and 'physical role limitations' in the short-term and for the domains 'physical role limitations' and 'emotional role limitations' in the longer-term (Figure 7a and 7b). Four domains, including 'physical functioning', 'bodily pain', 'social functioning' and 'mental health', showed a similar pattern with lower scores shortly after burns and these improved towards the norm afterwards. The other four domains showed different patterns. The domain 'vitality' showed a large variety in obtained scores in the short-term, both below and above the US-norm score.

However, afterwards, scores were closer to the norm score. The self-reported outcomes of the 'general health' domain remained constant during the whole follow-up time. Scores of the domain 'emotional role limitations' were relatively high shortly after burns, but lower scores were reported in the longer-term. The outcomes of the remaining domain, 'physical role limitations',

varied widely among studies during the entire follow-up period. Overall, subgroups with less severe injury (i.e. no surgery, no contractures) had higher scores on all domains.

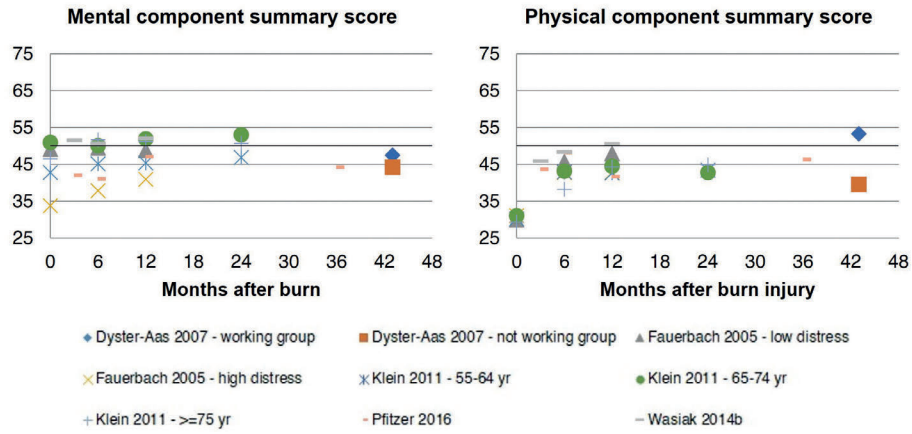


Figure 6. SF-36 physical component summary scores and mental component summary scores for five studies. The black line in the figures represents the US-norm score.

EQ-5D recovery patterns

The EQ-5D consists of five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression and a visual analogue scale (VAS) for general health. Each dimension has three levels of severity: no problems, moderate problems or severe problems[110]. Based on the answers of the five dimensions, a single index value can be derived ranging from 0 (death) to 1 (full health). Eight studies used the EQ-5D; data of 5 studies could be used to examine the recovery patterns based on the EQ-5D. Three studies were based on the same data source as studies already included in the analyses and were therefore not used[36,61,76]. As only two studies included a time point after 12 months (resp. 18 months[45] and on average 55 months[37]), no firm conclusions can be drawn on longer-term recovery. All studies reported the EQ-5D VAS score for general health. Reported scores were lower shortly after burns and increased with time towards the norm score (Figure 8). The study reporting lower scores at 12 months was the only study in more severe burn patients[87]. Lowest scores shortly after burns were seen for the EQ-5D index and the 'pain/discomfort' domain. The EQ-5D VAS score improved towards the norm score in the longer-term, just as the 'mobility' and 'self-care' domain. The self-reported outcomes of two other domains, 'usual activities', and 'anxiety/depression' and the EQ-5D index showed some improvement over time, but did not reach the level of the norm scores. The outcomes of the last domain 'pain/discomfort' did not show much improvement over time.

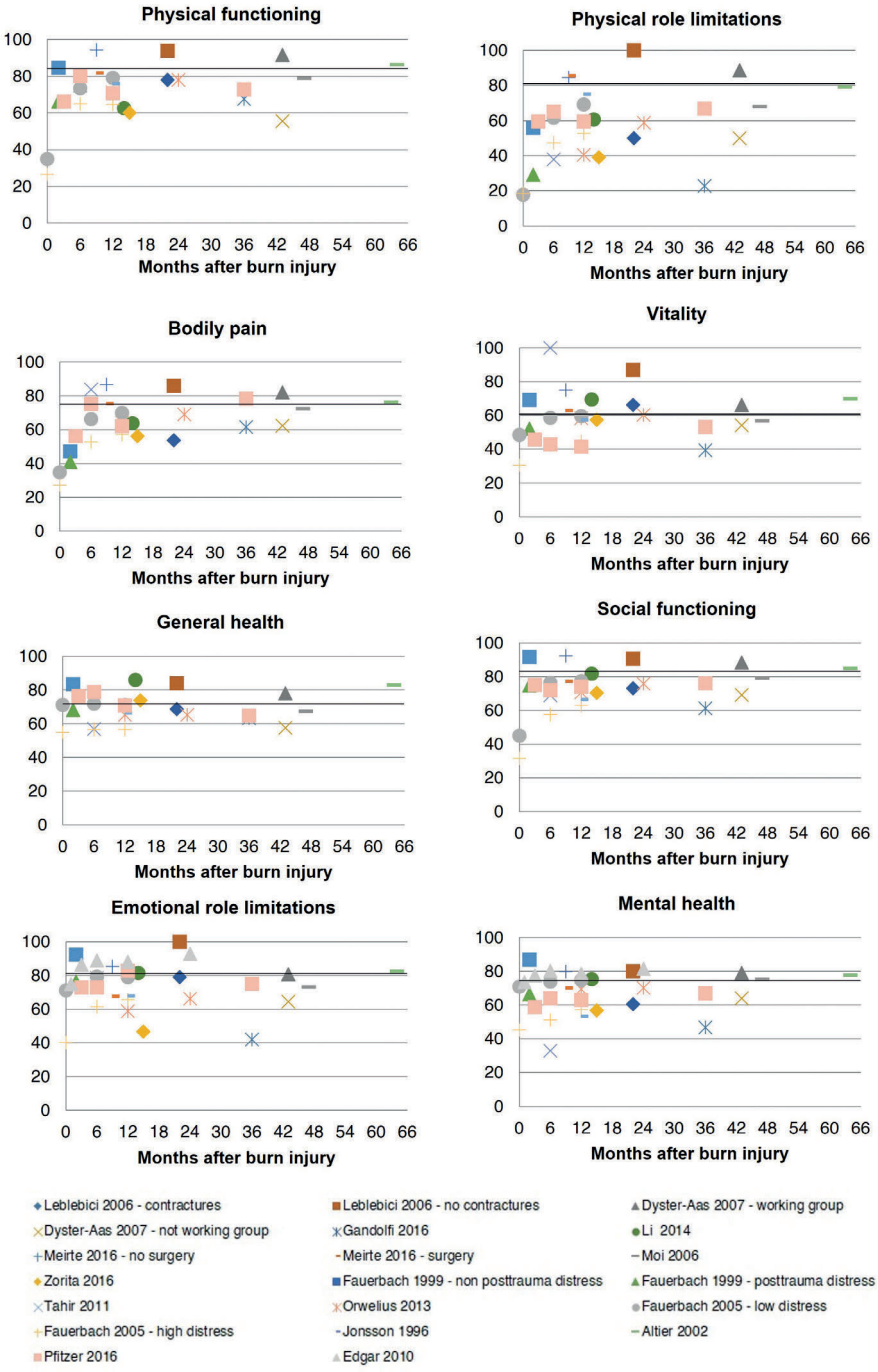


Figure 7a. SF-36 domain scores for fourteen studies.

The black line in the figures represents the US-norm score.

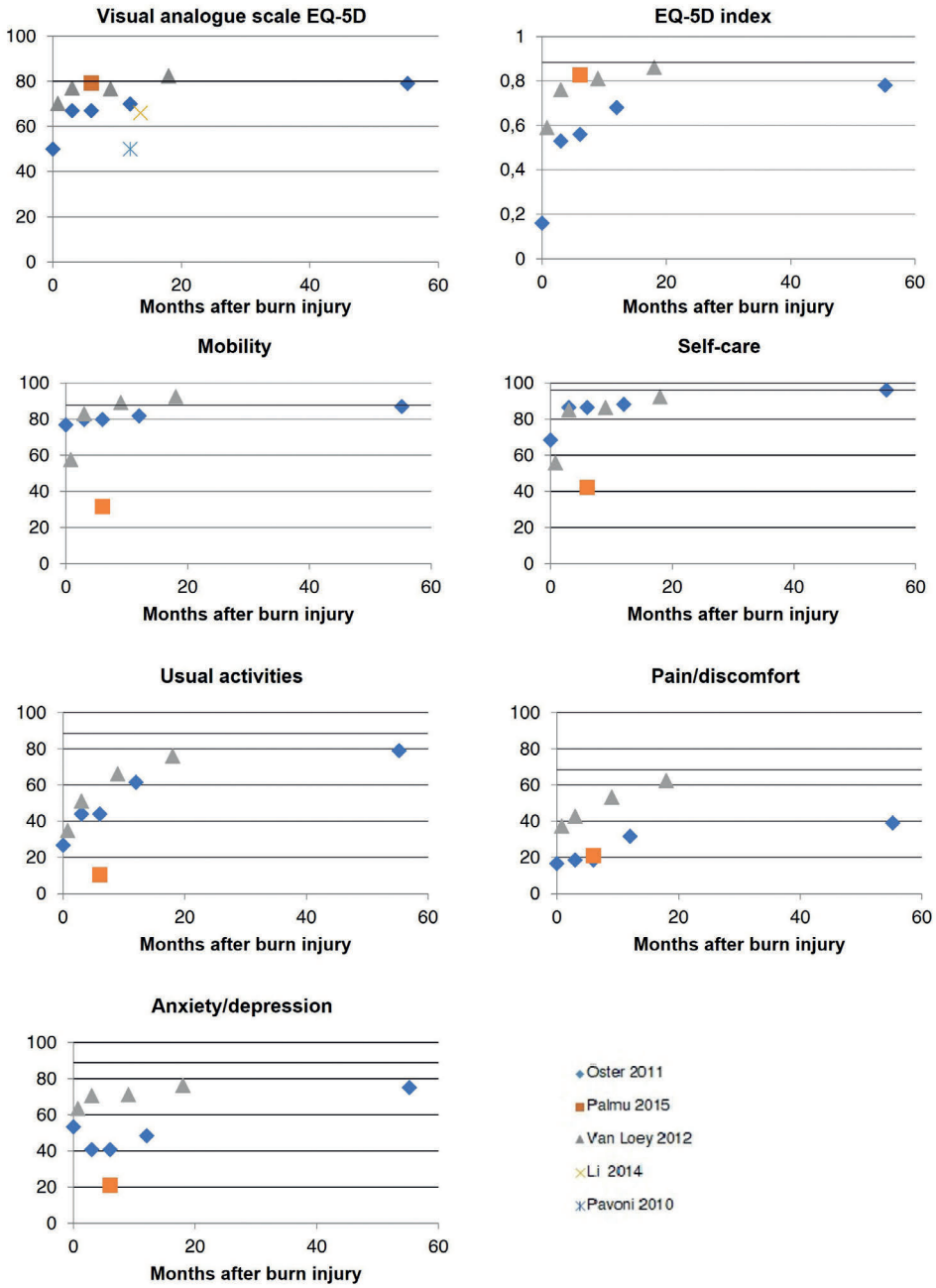


Figure 8. EQ-5D scores the visual analogue scale, the EQ-5D index and five dimensions for three to five studies.

The line in the figures represent the composed norm score based on norm scores of the countries where the studies were conducted [111]. The y-axis represents 0 - 100% patients with no problems on a specific domain.

DISCUSSION

This review provides a comprehensive overview of generic and burn specific instruments used to measure HRQL in adult burn patients and examined recovery patterns of HRQL in burns. Twenty HRQL instruments were used among the 94 studies. The BSHS-B and the SF-36 were most widely applied followed by the EQ-5D. It was seen that scores on most domains, both mental and physically orientated, were lower shortly after burns and improved over time. However, the BSHS-B domains 'work' and 'heat sensitivity', the SF-36 domains 'emotional role limitations' and 'physical role limitations', and the EQ-5D domain 'pain/discomfort' showed considerable variation across studies and low scores were also reported in the longer-term. The methodological quality of the included papers was in general moderate.

This review showed that there is some agreement on instruments used for the measurement of HRQL in adults after burns. Both instruments that are validated and that are not validated in the burn population are used. The majority of studies (70%) used the BSHS-B, the SF-36, or a combination of both instruments and eight studies (9%) used the EQ-5D, which are all validated in the burn population. It is recommended to use both a validated generic and burn specific instrument to assess the HRQL to capture the full impact of a health condition[112]. However, only 24 (26%) of the included studies used a combination of instruments. The (additional) use of a generic instrument, the SF-36 or the EQ-5D has the advantage that norm scores are available. The use of norm scores facilitates the comparison with other populations and interpretation of the outcomes. For the BSHS-B, partial population norm scores are available, including 30 of 40 items of the BSHS-B; the remaining ten items were considered too specific for burns[113]. Unfortunately, the results are not summarized on domain level. This would have provided norm scores for six of the BSHS-B domains. In the absence of population norm scores, domain scores reported by burns survivors in the long-term can be used as norm values.

Despite the widespread use of the BSHS-B, there is discussion about this instrument. A study comparing the SF-36 with the BSHS-B found that the SF-36 domains are more sensitive than the BSHS-B domains from 1 month post burn[26]. Besides, there is no evidence on test-retest ability, validity of hypothesis testing and item-total correlations of the BSHS-B[114]. Currently several new instruments are being developed by different research groups[114-117], resulting in different instruments which may hamper the comparison of outcomes in the future. There is a need to achieve consensus on which HRQL instruments are best to use in burn populations and at which time points. The studies with a longitudinal design (n=32) showed overlap in their assessment points. Most studies assessed HRQL at baseline, 3 months, 6 months, 12 months and 24 months post burn. Given the high attrition rates in burn studies, it may be difficult to obtain longer follow-up. However, a further improvement of HRQL beyond this period may be expected as it is known that HRQL further improves after 24 months[37,38].

The three HRQL questionnaires have overlapping domains[118]. For example, the domains 'simple abilities' (BSHS-B), 'physical functioning' (SF-36), 'mobility' (EQ-5D) and 'self-care' (EQ-5D) all measure activity limitations. Results on the different questionnaires show congruent results; activity is limited shortly after burns and improves with time. This is in line with the course of the recovery of burns as shortly after burns wounds are healing and physical capability is impaired. When wounds are healed activity improves. However, participation restrictions due to physical functioning are seen in both the short- and longer-term. The three domains covering this ('work' (BSHS-B), 'physical role limitations' (SF-36), and 'usual activities' (EQ-5D)) show mixed results, with also reduced scores in the longer-term. Simple activities like walking and dressing improve towards the level of the average population, however, more advanced functioning like working is more affected by burns and varies among the population, which might be explained by the heterogeneous nature of the burn population in combination with reported substantial effects on work situation, also in burns of limited severity[119].

Participation restrictions due to emotional and mental well-being ('interpersonal relationships' (BSHS-B), 'social functioning' (SF-36) and 'emotional role limitations' (SF-36)) are less prevalent after burns. In the short term there are some limitations with social activities, but this improves over time. In the longer-term, limitations of regular daily activities, including work, because of emotional problems seem to develop. Patients accomplish less than they would like and work not as carefully as usual. Mental function improved over time. This was consistent across the questionnaires ('affect' (BSHS-B), 'mental health' (SF-36) and 'anxiety/depression' (EQ-5D)). However, the scores for anxiety and depression did not reached the level of the general population, indicating that burn patients are on average more anxious or depressed.

Results on pain varied between the domains measuring this construct. According to the 'bodily pain' domain of the SF-36, the level of pain decreases with time and is comparable to the level of the general population in the longer-term, whereas the domain 'pain/discomfort' from the EQ-5D shows that the majority of patients experience pain or other discomfort in the longer-term. This is a much higher percentage than the proportion of the general population experiencing pain. Pain might thus be an issue in some patients in the longer-term, but does not seem to interfere with daily activities.

This review has a number of strengths and limitations. Strengths include the comprehensive overview of HRQL instruments used in burn populations, based on six databases, and the identification of HRQL domains that need more attention in the aftermath of burn injuries. However, some limitations also merit note. The scope of the review was limited to English-language studies, which might have resulted in missed studies that were published in foreign language journals. Another limitation is the wide variation in both study designs and instruments used, impeding a meta-analysis. Besides, due to the low number of longitudinal studies, we had

to use cross-sectional studies to examine recovery patterns. Also, the review was hampered by different ways of reporting the results, including mean or median scores, domain scores versus total scores and 0-100 scores or standardized norm scores, which makes it hard to compare results. Besides, the methodological quality of included studies varied widely. The most alarming was the general high risk of bias on study attrition. Only few studies adequately reported attempts to collect information on participants who dropped out and key characteristics on those lost to follow-up. In future articles it is important to include description of these factors in order to reach a low risk of bias on study attrition and improve the overall study quality.

CONCLUSION

This review demonstrates that most domains of HRQL, frequently measured using the BSHS, SF-36 or EQ-5D, are affected shortly after the burn event. Most domains will recover over time excluding physical and emotional role participation, anxiety, depression and pain. This reflects the need for both mental and physical support in the aftermath of burns. To further facilitate the comparability of burn-related HRQL outcomes across the world, use of uniform validated instruments, time points and data presentation is needed. It is therefore recommended to develop a guideline on the measurement of HRQL in burn patients.

Acknowledgements

The authors gratefully acknowledge Wichor Bramer (Biomedical information specialist, Medical Library, Erasmus MC) for performing the database search. We would also like to acknowledge the authors (Dr Müller, Dr Orwelius, Professor Ekholm, Professor Pfitzer, Dr Öster and Dr Edgar) for provision of data from their articles.

REFERENCES

1. Van den Kerckhove E, Stappaerts K, Boeckx W, et al. Silicones in the rehabilitation of burns: a review and overview. Burns 2001; 27:205-14.
2. Bloemsma G, Dokter J, Boxma H, et al. Mortality and causes of death in a burn centre.Burns 2008;34:1103-7.
3. Falder S, Browne A, Edgar D, et al. Core outcomes for adult burn survivors: a clinical overview. Burns 2009; 35:618-41.
4. Jasper S, Rennekampff H, de Zwaan M. Psychiatric co-morbidity, body image problems and psychotherapeutic interventions for burn survivors: a review. Psychother Psychosom Med Psychol 2013; 63:423-8.
5. Van Baar M, Essink-Bot M-L, Oen I, et al. Functional outcome after burns: a review. Burns 2006; 32:1-9.
6. Guyatt GH, Jaeschke R, Feeny DH, et al. Measurements in clinical trials: choosing the right approach. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 21996. p. 41-8.
7. Mustoe TA, Cooter RD, Gold MH, et al. International clinical recommendations on scar management. Plast Reconstr Surg 2002; 110:560-71.
8. Pereira C, Murphy K, Herndon D. Outcome measures in burn care: is mortality dead? Burns 2004; 30:761-71.
9. Yoder LH, Nayback AM, Gaylord K. The evolution and utility of the burn specific health scale: a systematic review. Burns 2010; 36:1143-56.
10. Stavrou D, Weissman O, Tessone A, et al. Health related quality of life in burn patients–a review of the literature. Burns 2014; 40:788-96.
11. Jabir S, Dziewulski P. Response to: Health related quality of life in burn patients–a review of the literature. Burns 2015; 41:645.
12. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 6:e1000097.
13. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med 2013; 158:280-6.
14. Grisbrook TL, Reid SL, Edgar DW, et al. Exercise training to improve health related quality of life in long term survivors of major burn injury: A matched controlled study. Burns 2012; 38:1165-73.
15. Miller T, Bhattacharya S, Zamula W, et al. Quality-of-life loss of people admitted to burn centers, United States. Qual Life Res 2013; 22:2293-305.
16. Blades B, Munster AM. A burn specific health scale. J Trauma 1982; 22:872-5.
17. Blades BC, Jones C, Munster AM. Quality of life after major burns. J Trauma 1979; 19:556-8.
18. Cakir U, Terzi R, Abaci F, et al. The prevalence of post-traumatic stress disorder in patients with burn injuries, and their quality of life. Int J Psychiatry Clin Pract 2015; 19:56-9.
19. Dyster-Aas J, Kildal M, Willebrand M, et al. Work status and burn specific health after work-related burn injury. Burns 2004; 30:839-42.
20. Munster AM, Fauerbach JA, Lawrence J. Development and utilization of a psychometric instrument for measuring

quality of life in burn patients, 1976 to 1996. Acta Chir Plast 1996; 38:128-31.

21. Cochran A, Edelman LS, Saffle JR, et al. Self-Reported Quality of Life after Electrical and Thermal Injury. J Burn Care Rehabil 2004; 25:61-6.
22. Xie B, Xiao SC, Zhu SH, et al. Evaluation of long term health-related quality of life in extensive burns: A 12-year experience in a burn center. Burns 2012; 38:348-55.
23. Altier N, Malenfant A, Forget R, et al. Long-term adjustment in burn victims: A matched-control study. Psychol Med 2002; 32:677-85.
24. Cromes GF, Holavanahalli R, Kowalske K, et al. Predictors of quality of life as measured by the burn specific health scale in persons with major burn injury. J Burn Care Rehabil 2002; 23:229-34+7+8.
25. Dowda DJ, Li F. Major concerns and issues in burn survivors in Australia. Burns Trauma 2014; 2:84-7.
26. Edgar D, Dawson A, Hankey G, et al. Demonstration of the validity of the SF-36 for measurement of the temporal recovery of quality of life outcomes in burns survivors. Burns 2010; 36:1013-20.
27. Edgar DW, Homer L, Phillips M, et al. The influence of advancing age on quality of life and rate of recovery after treatment for burn. Burns 2013; 39:1067-72.
28. Fauerbach JA, Lawrence JW, Munster AM, et al. Prolonged adjustment difficulties among those with acute posttrauma distress following burn injury. J Behav Med 1999; 22:359-78.
29. Fauerbach JA, Lezotte D, Hills RA, et al. Burden of burn: A norm-based inquiry into the influence of burn size and distress on recovery of physical and psychosocial function. J Burn Care Rehabil 2005; 26:21-32.
30. Finlay V, Burke K, van de Ruit C, et al. Assessing the impact of missing data in evaluating the recovery of minor burn patients. Burns 2009; 35:1086-91.
31. Jarrett M, Physiotherapy B, McMahon M, et al. Physical outcomes of patients with burn injuries - A 12 month follow-up. J Burn Care Res 2008; 29:975-84.
32. Klein MB, Lezotte DC, Heltshe S, et al. Functional and psychosocial outcomes of older adults after burn injury: Results from a multicenter database of severe burn injury. J Burn Care Res 2011; 32:66-78.
33. Niţescu C, Calotţ DR, Sttncioiu TA, et al. Psychological impact of burn scars on quality of life in patients with extensive burns who received allotransplant. Rom J Mor Embry 2012; 53:577-83.
34. Novelli B, Melandri D, Bertolotti G, et al. Quality of life impact as outcome in burns patients. G Ital Med Lav Ergon 2009; 31:A58-A63.
35. Orwelius L, Willebrand M, Gerdin B, et al. Long term health-related quality of life after burns is strongly dependent on pre-existing disease and psychosocial issues and less due to the burn itself. Burns 2013; 39:229-35.
36. Öster C, Willebrand M, Dyster-Aas J, et al. Validation of the EQ-5D questionnaire in burn injured adults. Burns 2009; 35:723-32.
37. Öster C, Willebrand M, Ekselius L. Health-related quality of life 2 years to 7 years after burn injury. J Trauma Inj Infect Crit Care 2011; 71:1435-41.
38. Öster C, Willebrand M, Ekselius L. Burn-specific health 2 years to 7 years after burn

injury. J Trauma Acute Care Surg 2013; 74:1119-24.

39. Palmu R, Partonen T, Suominen K, et al. Health-related quality of life 6 months after burns among hospitalized patients: Predictive importance of mental disorders and burn severity. Burns 2015; 41:742-8.
40. Pfitzer B, Katona LJ, Lee SJ, et al. Three years after black Saturday: Long-term psychosocial adjustment of burns patients as a result of a major bushfire. J Burn Care Res 2016; 37:e244-e53.
41. Renneberg B, Ripper S, Schulze J, et al. Quality of life and predictors of long-term outcome after severe burn injury. J Behav Med 2014; 37:967-76.
42. Ryan CM, Lee A, Kazis LE, et al. Recovery trajectories after burn injury in young adults: Does burn size matter? J Burn Care Res 2015; 36:118-29.
43. Ryan CM, Schneider JC, Kazis LE, et al. Benchmarks for multidimensional recovery after burn injury in young adults: The development, validation, and testing of the american burn association/shriners hospitals for children young adult burn outcome questionnaire. J Burn Care Res 2013; 34:e121-e42.
44. Tahir SM, Memon MM, Ali SA, et al. Health related quality of life after burns: are we really treating burns? J Ayub Med Coll Abbottabad 2011; 23:152-6.
45. Van Loey NE, van Beeck EF, Faber BW, et al. Health-related quality of life after burns: a prospective multicenter cohort study with 18 months follow-up. J Trauma Acute Care Surg 2012; 72:513-20.
46. Wasiak J, Lee SJ, Paul E, et al. Predictors of health status and health-related quality of life 12 months after severe burn. Burns 2014; 40:568-74.
47. Wasiak J, Lee SJ, Paul E, et al. Female patients display poorer burn-specific quality of life 12 months after a burn injury. Injury 2016.
48. Wasiak J, Mahar P, Lee S, et al. 12-Month generic health status and psychological distress outcomes following an australian natural disaster experience: 2009 Black Saturday Wildfires. Injury 2013; 44:1443-7.
49. Wasiak J, Paul E, Lee SJ, et al. Patterns of recovery over 12 months following a burn injury in Australia. Injury 2014; 45:1459-64.
50. Williams N, Stiller K, Greenwood J, et al. Physical and quality of life outcomes of patients with isolated hand burns - A prospective audit. J Burn Care Res 2012; 33:188-98.
51. Koljonen V, Laitila M, Sintonen H, et al. Health-related quality of life of hospitalized patients with burns - Comparison with general population and a 2-year follow-up. Burns 2013; 39:451-7.
52. Willebrand M, Kildal M. Burn specific health up to 24 months after the burn—a prospective validation of the simplified model of the Burn Specific Health Scale-Brief. J Trauma Acute Care Surg 2011; 71:78-84.
53. Koljonen V, Laitila M, Rissanen AM, et al. Treatment of patients with severe burns-costs and health-related quality of life outcome. J Burn Care Res 2013; 34:e318-e25.
54. Ahuja RB, Mulay AM, Ahuja A. Assessment of quality of life (QoL) of burn patients in India using BSHS-RBA scale. Burns 2016; 42:639-47.

55. Anzarut A, Chen M, Shankowsky H, et al. Quality-of-life and outcome predictors following massive burn injury. Plast Reconstr Surg 2005; 116:791-7.

56. Baker CP, Rosenberg M, Mossberg KA, et al. Relationships between the Quality of Life Questionnaire (QLQ) and the SF-36 among young adults burned as children. Burns 2008; 34:1163-8.

57. Baker CP, Russell WJ, Meyer Iii W, et al. Physical and Psychologic Rehabilitation Outcomes for Young Adults Burned as Children. Arch Phys Med Rehabil 2007; 88:S57-S64.

58. Daigeler A, Langer S, Hüllmann K, et al. A follow-up study of adults with suicidal burns: Psychosocial adjustment and quality of life. J Burn Care Res 2009; 30:844-51.

59. Druery M, Brown TLH, Muller M. Long term functional outcomes and quality of life following severe burn injury. Burns 2005; 31:692-5.

60. Dyster-Aas J, Kildal M, Willebrand M. Return to work and health-related quality of life after burn injury. J Rehabil Med 2007; 39:49-55.

61. Ekeblad F, Gerdin B, Öster C. Impact of personality disorders on health-related quality of life one year after burn injury. Disabil Rehabil 2015; 37:534-40.

62. Elsherbiny OEE, Salem MA, El-Sabbagh AH, et al. Quality of life of adult patients with severe burns. Burns 2011; 37:775-88.

63. Ferreira E, Dantas RAS, Rossi LA, et al. The cultural adaptation and validation of the "Burn Specific Health Scale-Revised" (BSHS-R): Version for Brazilian burn victims. Burns 2008; 34:994-1001.

64. Gandolfi S, Auquit-Auckbur I, Panunzi S, et al. Validation of the French version of the Burn Specific Health Scale-Brief (BSHS-B) questionnaire. Burns 2016; 42:1573-80.

65. Hwang YF, Chen-Sea MJ, Chen CL, et al. Validation of a taiwanese version of the burn-specific health scale-brief. J Burn Care Res 2016; 37:e310-e6.

66. Jonsson CE, Schüldt K, Linder J. Rehabilitative, psychiatric, functional and aesthetic problems in patients treated for burn injuries--a preliminary follow-up study. Acta Chir Plast 1996.

67. Kildal M, Andersson G, Gerdin B. Health status in Swedish burn patients - Assessment utilising three variants of the Burn Specific Health Scale. Burns 2002; 28:639-45.

68. Kildal M, Willebrand M, Andersson G, et al. Personality Characteristics and Perceived Health Problems after Burn Injury. J Burn Care Rehabil 2004; 25:228-35.

69. Kildal M, Willebrand M, Andersson G, et al. Coping strategies, injury characteristics and long-term outcome after burn injury. Injury 2005; 36:511-8.

70. Kimmo T, Jyrki V, Sirpa AS. Health status after recovery from burn injury. Burns 1998; 24:293-8.

71. Knight A, Wasiak J, Salway J, et al. Factors predicting health status and recovery of hand function after hand burns in the second year after hospital discharge. Burns 2017; 43:100-6.

72. Leblebici B, Adam M, Batţţ S, et al. Quality of life after burn injury: The impact of joint contracture. J Burn Care Res 2006; 27:864-8.

73. Li DW, Liu WQ, Wang HM, et al. The Chinese language version of the abbreviated Burn Specific Health Scale: A validation study. Burns 2014; 40:1001-6.

74. Low AJF, Dyster-Aas J, Willebrand M, et al. Psychiatric morbidity predicts perceived burn-specific health 1 year after a burn. Gen Hosp Psychiatry 2012; 34:146-52.

75. Mazharinia N, Aghaei S, Shayan Z. Dermatology life quality index (DLQI) scores in burn victims after revival. J Burn Care Res 2007; 28:312-7.

76. Meirte J, Van Daele U, Maertens K, et al. Convergent and discriminant validity of quality of life measures used in burn populations. Burns 2017; 43:84-92.

77. Moi AL, Nilsen RM. Pathways leading to self-perceived general health and overall quality of life in burned adults. Burns 2012; 38:1157-64.

78. Moi AL, Wentzel-Larsen T, Salemark L, et al. Erratum: Validation of a Norwegian version of the Burn Specific Health Scale (Burns (2003) 29: 6 (563-570)). Burns 2003; 29:872-3.

79. Moi AL, Wentzel-Larsen T, Salemark L, et al. Long-term risk factors for impaired burn-specific health and unemployment in patients with thermal injury. Burns 2007; 33:37-45.

80. Moi AL, Wentzel-Larsen T, Salemark L, et al. Impaired generic health status but perception of good quality of life in survivors of burn injury. J Trauma Inj Infect Crit Care 2006; 61:961-8.

81. Mulay AM, Ahuja A, Ahuja RB. Modification, cultural adaptation and validation of burn specific health scale-brief (BSHS-B) for Hindi speaking population. Burns 2015; 41:1543-9.

82. Müller A, Smits D, Jasper S, et al. Validation of the German version of the Burn Specific Health Scale-Brief (BSHS-B). Burns 2015; 41:1333-9.

83. Murphy ME, Holzer CE, Richardson LM, et al. Quality of life of young adult survivors of pediatric burns using world health organization disability assessment Scale II and burn specific health scale-brief: A comparison. J Burn Care Res 2015; 36:521-33.

84. Noble J, Gomez M, Fish JS. Quality of life and return to work following electrical burns. Burns 2006; 32:159-64.

85. Pallua N, Künsebeck HW, Noah EM. Psychosocial adjustments 5 years after burn injury. Burns 2003; 29:143-52.

86. Palmieri TL, Molitor F, Grace C, et al. Long-term functional outcomes in the elderly after burn injury. J Burn Care Res 2012; 33:497-503.

87. Pavoni V, Gianesello L, Paparella L, et al. Outcome predictors and quality of life of severe burn patients admitted to intensive care unit. Scand J Trauma Resusc Emerg Med 2010; 18:24.

88. Piccolo MS, Gragnani A, Daher RP, et al. Validation of the Brazilian version of the Burn Specific Health Scale-Brief (BSHS-B-Br). Burns 2015; 41:1579-86.

89. Pishnamazi Z, Rejeh N, Heravi-Karimooi M, et al. Validation of the Persian version of the Burn Specific Health Scale-Brief. Burns 2013; 39:162-7.

90. Ricci H, Gonçalves N, Gallani MC, et al. Assessment of the health status in Brazilian burn victims five to seven months after hospital discharge. Burns 2014; 40:616-23.

91. Roh YS, Chung HS, Kwon B, et al. Association between depression, patient scar assessment and burn-specific health

in hospitalized burn patients. Burns 2012; 38:506-12.

92. Rosenberg M, Blakeney P, Robert R, et al. Quality of life of young adults who survived pediatric burns. J Burn Care Res 2006; 27:773-8.

93. Rosenberg M, Ramirez M, Epperson K, et al. Comparison of long-term quality of life of pediatric burn survivors with and without inhalation injury. Burns 2015; 41:721-6.

94. Salvador Sanz JF, Sanchez-Payá J, Marín JR. Spanish version of the Burn-Specific Health Scale. J Trauma Inj Infect Crit Care 1998; 45:581-7.

95. Salvador-Sanz JF, Sanchez-Payá J, Rodriguez-Marín J. Quality of life of the Spanish burn patient. Burns 1999; 25:593-8.

96. Szczechowicz J, Lewandowski J, Sikorski J. Polish adaptation and validation of Burn Specific Health Scale - Brief. Burns 2014; 40:1013-8.

97. Willebrand M, Andersson G, Kildal M, et al. Injury-related fear-avoidance, neuroticism and burn-specific health. Burns 2006; 32:408-15.

98. Zhang LJ, Cao J, Feng P, et al. Influencing factors of the quality of life in Chinese burn patients: Investigation with adapted Chinese version of the BSHS-B. Burns 2014; 40:731-6.

99. Zorita LD, Blanes L, Veiga DF, et al. Health-related Quality of Life and Self-esteem Among Burn Patients. Wounds Compend Clin Res Pract 2016; 28:27-34.

100. Xie B, Xiao SC, Peng XD, et al. Epidemiology and outcome analysis of severe extensive burns: A 12-year summary of 103 cases in a burn center in China. J Burn Care Res 2012; 33:e127-e32.

101. Blalock SJ, Bunker BJ, DeVellis RF. Measuring health status among survivors of burn injury: revisions of the Burn Specific Health Scale. J Trauma Acute Care Surg 1994; 36:508-15.

102. Goncalves N, Ciol MA, Dantas RA, et al. A randomized controlled trial of an educational programme with telephone reinforcement to improve perceived health status of Brazilian burn victims at 6-month post discharge. J Adv Nurs 2016; 72:2508-23.

103. Hashemi F, Rahimi Dolatabad F, Yektatalab S, et al. Effect of Orem Self-Care Program on the Life Quality of Burn Patients Referred to Ghotb-al-Din-e-Shirazi Burn Center, Shiraz, Iran: A Randomized Controlled Trial. Int J Community Based Nurs Midwifery 2014; 2:40-50.

104. Tang D, Li-Tsang CWP, Au RKC, et al. Functional Outcomes of Burn Patients with or Without Rehabilitation in Mainland China. Hong Kong J Occup Ther 2015; 26:15-23.

105. Stavrou D, Haik J, Wiser I, et al. Validation of the Hebrew version of the Burn Specific Health Scale-Brief questionnaire. Burns 2015; 41:188-95.

106. Corry NH, Klick B, Fauerbach JA. Posttraumatic stress disorder and pain impact functioning and disability after major burn injury. J Burn Care Res 2010; 31:13-25.

107. Ling-Juan Z, Jie C, Jian L, et al. Development of quality of life scale in Chinese burn patients: Cross-cultural adaptation process of burn-specific health scale - Brief. Burns 2012; 38:1216-23.

108. Kildal M, Andersson G, Fugl-Meyer AR, et al. Development of a brief version of the Burn Specific Health Scale (BSHS-B). J Trauma Acute Care Surg 2001; 51:740-6.

109. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. Med Care 1992:473-83.

110. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35:1095-108.

111. Szende A, Janssen B, Cabases J. Self-reported population health: an international perspective based on EQ-5D: Springer; 2014.

112. Van Beeck EF, Larsen CF, Lyons RA, et al. Guidelines for the conduction of follow-up studies measuring injury-related disability. J Trauma Acute Care Surg 2007; 62:534-50.

113. Kvannli L, Finlay V, Edgar DW, et al. Using the Burn Specific Health Scale-Brief as a measure of quality of life after a burn—What score should clinicians expect? Burns 2011; 37:54-60.

114. Griffiths C, Guest E, White P, et al. A systematic review of patient-reported outcome measures used in adult burn research. J Burn Care Res 2017; 38:e521-e45.

115. Griffiths C, Armstrong-James L, White P, et al. A systematic review of patient reported outcome measures (PROMs) used in child and adolescent burn research. Burns 2015; 41:212-24.

116. 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; 41:1405-19.

117. Kool MB, Geenen R, Egberts MR, et al. Patients' perspectives on quality of life after burn. Burns 2017.

118. Meirte J, Van Loey NEE, Maertens K, et al. Classification of quality of life subscales within the ICF framework in burn research: identifying overlaps and gaps. Burns 2014; 40:1353-9.

119. Goei H, Hop MJ, van der Vlies C, et al. Return to work after specialised burn care: a two-year prospective follow-up study of the prevalence, predictors and related costs. Injury 2016; 47:1975-82.

3

Predictors of health-related quality of life after burn injuries: a systematic review

I. Spronk , **C.M. Legemate**, J. Dokter, N.E. van Loey, M.E. van Baar, S. Polinder

Critical Care 2018, June 14;22:160

ABSTRACT

Background

Identifying predictors of health-related quality of life (HRQL) following burns is essential for optimization of rehabilitation for burn survivors. This study aimed to systematically review predictors of HRQL in burn patients.

Methods

Medline, Embase, Web of Science, Cochrane, CINAHL and Google Scholar were reviewed from inception to October 2016 for studies that investigated at least one predictor of HRQL after burns. The Quality in Prognostic Studies tool was used to assess risk of bias of included studies.

Results

Thirty-two studies were included. Severity of burns, postburn depression, post-traumatic stress symptoms, avoidant coping, less emotional or social support, higher levels of neuroticism and unemployment postburn were found to predict a poorer HRQL after burns in multivariable studies. In addition, weaker predictors included female gender, pain and a postburn substance use disorder. Risk of bias was generally low in outcome measurement and high in study attrition and study confounding.

Conclusions

HRQL after burns is affected by the severity of burns and the psychological response to the trauma. Both constructs provide unique information and knowledge that is necessary for optimized rehabilitation. Therefore both physical and psychological problems require attention months to years after the burn trauma.

BACKGROUND

Health-related quality of life (HRQL) is an important outcome measure of burns in both the short- and long-term[1,2] and is increasingly studied. HRQL is a multidimensional concept that reflects an individual's perception of how a disease affects his/her physical, psychological, and social well-being[3-5]. Insight into which factors determine HRQL after burns is useful for clinical practice, research, and policy making. Conceptual models have been developed in order to better understand HRQL and the variables that relate to HRQL in general[3,6-8]. According to the 'Revised of Wilson and Clearly model for health-related quality of life', HRQL is influenced by individual and environmental characteristics, biological function, symptoms, functional status and general health perceptions[3]. A recent study confirmed that this model is also applicable to burns[9].

Burns can have a considerable negative impact on daily activities and on both physical and psychosocial functioning[10-12]. HRQL domains are often impaired in the short-term. Most domains of HRQL improve in the longer-term, but also in the longer-term some aspects (e.g. physical and emotional role participation) have poor outcomes[13-15]. Burn injuries are thus associated with a significant physical and psychological burden.

The prediction of an individual's ability to adjust to the consequences of their burn injury is important. Information regarding these predictors may help caregivers in selecting patients who require special attention in rehabilitation and in preparing patient specific care plans[16]. Predictors of HRQL following burns have been examined in individual studies, but predictors of HRQL have not been systematically reviewed in the field of burns. Potential meaningful factors are the patient's age and gender, percentage total body surface area (%TBSA) burned), length of hospital stay, body area affected, time since injury and psychological impact of burns. However, it is not yet clear which predictive factors are most important[17-20]. Earlier recent reviews focussed on the evolution and relevance of one specific HRQL instrument in burns[21], on HRQL outcomes in burns[19] and on HRQL instruments used and recovery patterns of HRQL in burns, without studying predictors. Therefore, the aim of present study is to systematically review predictors of a HRQL following burn injuries.

METHODS

This systematic review was conducted and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Statement[22] and it was registered on PROSPERO (ID=CRD42016048065).

Search strategy and inclusion criteria

The databases Medline, Embase, Web of Science, Cochrane, CINAHL and Google Scholar were systematically searched using terms covering HRQL and burns (Appendix 1) in October 2016. The search strategy was developed in collaboration with an experienced librarian. Original prognostic studies conducted in adult burn patients and focussing on at least one predictor of HRQL after burns were included. Studies had to be published in a peer-reviewed journal and written in English. Studies were required to have used a generic or burn-specific instrument to assess HRQL. Outcomes had to be a regression or correlation coefficient of the relation of a predictor with HRQL. All kind of predictors were considered.

Selection of studies and data extraction

An experienced librarian performed the systematic search. After removal of duplicates, relevant articles were selected on the basis of title by one researcher (IS). Ten percent of the abstracts was independently evaluated by two researchers (IS and CL). Perfect agreement on inclusion was achieved (Cohen's kappa coefficient=1), therefore, one researcher evaluated the remaining abstracts (IS)[23]. In case of any doubt, a title or abstract was screened by a second researcher. Two researchers (IS and CL) independently performed screening of full text and extraction of data. The screening of these three steps was performed using the above-mentioned inclusion criteria. Data extraction included study characteristics (study type, country, sample size, assessment time points, length of follow-up), patient and burn characteristics (age, gender, hospital length of stay (LOS), %TBSA, details on HRQL instruments (type, number, general burn-specific HRQL, proxy) and predictors (number of predictors assessed, univariable and multivariable predictors, statistical methods). Discrepancies arising from decisions around inclusion or extraction of data were discussed with a third researcher (MvB) until resolved.

Risk of bias

The Quality in Prognostic Studies (QUIPS) tool[24] was used to assess the risk of bias of the included studies. Two researchers (IS and CL) independently assessed the risk of bias of the six domains. The domains were rated as either low, moderate or high risk of bias. A low risk was obtained when all items of a domain were scored as 'low risk'[24]. A moderate risk was obtained when at least one and maximum half of the items were rated as high or unknown risk of bias. A high risk was obtained when more than half of the items were rated as high or unknown risk of bias. Disagreements were resolved by discussion with a third researcher (MvB).

Data analysis

First the characteristics and the risk of bias of all studies were tabulated. Then the predictor findings of studies using multivariable analysis were analyzed. Multivariable models were models that included at least two factors to predict HRQL. Predictors were divided into four categories: demographic, environmental, burn-specific, and psychological factors. If it was unclear whether

associations were significant ($p \leq 0.05$), results could not be included in our analysis. When more than one time point was used, the point closest to the most often used time points in other studies was chosen. Given the heterogeneity of predictors, HRQL instruments, and statistical reporting, meta-analyses could not be conducted. Therefore, a more qualitative approach was used: all predictors of each study were summarized on the basis of its direction and statistical significance[25,26]. Predictors were scored having no statistically significant association ($p > 0.05$) with HRQL, a significant association ($p \leq 0.05$) with a subscale of the HRQL instrument, or a significant association with the total HRQL instrument. Associations with the total HRQL instrument were heavier weighted (see Table 3). Due to the wide variety of predictors assessed among the included studies, only those predictors that were studied in more than one study were tabulated (Table 3). Predictors were considered strong when ≥67% of the associations were in the same direction and statistically significant and weak if ≥33-<67% of the associations fulfilled these conditions.

RESULTS

Search results

The initial database search netted 6,173 records, including 3,788 unique articles. Screening of titles and abstracts resulted in 144 potentially relevant articles (Figure 1). Thirty-two of these were eligible after reading the full-text. The main reason for exclusion was not studying predictors.

Study characteristics

Sample sizes varied between 20 and 1,051 patients, with most studies (75%) having a sample size below 200 patients (Table 1). In all except one study[27], more males than females were included. The mean percentage TBSA burned ranged from 8% to 84%. Eleven different HRQL instruments were used in the included studies. The most often used instruments were the Burn-specific Health Scale-Brief (BSHS-B) (n=17) and the Medical Outcome Study Short Form-36 items (SF-36) (n=11). Eighteen studies measured HRQL at one time point, whereas thirteen studies measured HRQL two to six times. One study failed to describe their assessment point. The most used time points were at 3 months (n=6), at 6 months (n=11), at 12 months (n=12) and at 24 months (n=7). Seventeen studies used an assessment point more than one year after the burn injury.

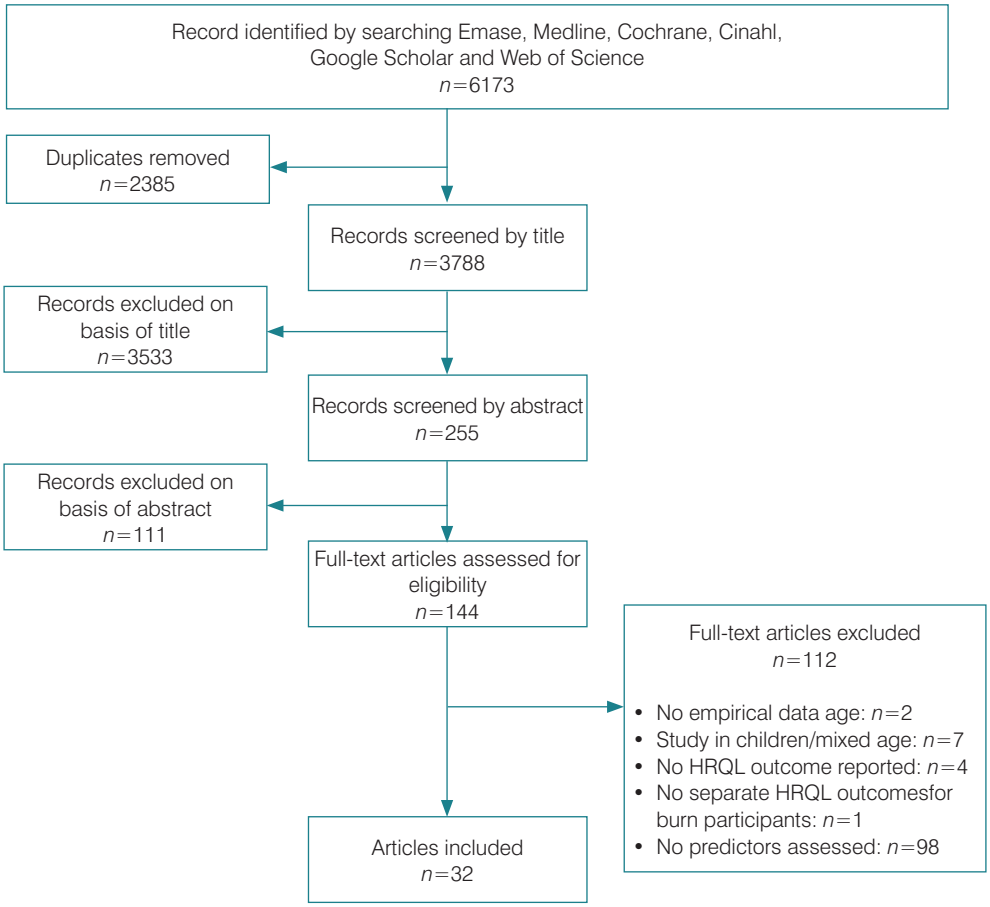


Figure 1. Flowchart selection of studies

Risk of bias

The quality of included studies was in general moderate. In most studies risk of bias was moderate or high on the items 'study attrition' and 'study confounding' (Table 2). Positive aspects of the studies were the low risk of bias on the items 'outcome measurement' and 'statistical analysis and reporting'. None of the studies scored a low risk of bias on all items, one study had a low risk on all but one dimensions[43].

Table 1. Characteristics of included studies (n=32)

Author year (reference)	Country	Study population[1]	%TBSA burned, mean (SD)	HRQL instrument[2]	Assessment time point(s)
Ahuja 2016[27]	India	n=60, (M: 40 %). Age: 18-65yr (median: 28yr)	Median: 30%	BSHS-RBA	Median: 10 months
Anzarut 2005[28]	Canada	n=47, (M: 96%). Age: (mean: 28yr)	64% (2)	BSHS-A, SF-36	≥2 years after discharge
Blalock 1994[29]	USA	n= 254, (M: 74%). Age: (mean: 39yr)	19% (15)	BSHS	Mean: 8-9 months
Corry 2010[30]	USA	n=171, (M: 70%). Age: 18-86yr (mean: 42yr)	15% (13), Range: 1-74%	SF-36	Discharge, 1, 6, 12, and 24 months[$]
Cromes 2002[31]	USA	n=110, (M: 84%). Age: (mean: 38yr)	24%	BSHS	2*, 6* and 12* months
Edgar 2013[17]	Australia	n=1051, (M: 80%). Age: 15-89yr (mean: 37yr)	8% (11), Range: 0-75%	BSHS-B, SF-36	1, 3, 6, 12 and 24 months[$]
Ekeblad 2015[32]	Sweden	n=107, (M: 75%). Age: 19-89yr (mean: 43yr)	23% Range: 1-80%	BSHS-B, EQ-5D, SF-36	12 months
Finlay 2014[33]	Australia	n=927, (M: 73%). Age: 16-83 yr (mean: 32yr)	7% (10)	BSHS-B	Discharge, 1, 3*, 6, 12, and 24 months
Finlay 2015[34]	Australia	n=224, (M:83%). Age: 16-84yr (median: 36yr)	Median: 4%, range 1-60%	BSHS-B	n.a.
Kildal 2001[35]	Sweden	n=248, (M:80%). Age: (mean: 37yr)	23% (16)	BSHS-B	Mean 9.3yr (SD: 4.8yr)
Kildal 2004[36]	Sweden	n=166, (M: 80%). Age: (mean: 50yr)	25% (16)	BSHS-B	Mean: 11.4yr, range: 3-19yr
Kildal 2005[37]	Sweden	n=161, (M: 79%). Age: 17-79yr (mean: 48yr)	24% (16) Range: 1-85%	BSHS-B	Mean: 9.2yr, range: 1-18yr
Knight 2017[38]	Australia	n=41, (M: 81%). Age: 19-81yr (mean: 45yr)	8%	BSHS-B	12-24 months
Leblebici 2006[39]	Turkey	n= 22, (M: 64%). Age: (mean: 25yr)	28% (17)	SF-36	Mean: 21 months

Table 1. Continued

Low 2012[40]	Sweden	n=85, (M: 75%). Age: 19-89yr (mean: 45yr)	24% (20), range: 1-80%	BSHS-B	12 months
Moi 2007[41]	Norway	n=95, (M: 82%). Age: (mean: 44yr)	19% (14)	BSHS-A	Mean: 47 months (SD: 24 months)
Moi 2012[9]	Norway	n=95, (M: 82%). Age: (mean: 44yr)	19% (14)	BSHS-A, SF-36, QOLS	Mean: 47 months (SD: 24 months)
Novelli 2009[42]	Italy	n= 30, (M: 60%). Age: (mean: 42yr)	32% (13)	SIP	Discharge, 3* months
Orwelius 2013[43]	Sweden	n=156, (M: 74%). Age: 16-90yr (mean: 46yr)	24% (19), range: 0-80%	SF-36	12* and 24 months
Oster 2011[18]	Sweden	n=89, (M: 77%). Age: (mean: 43yr)	25% (20)	EQ-5D	Admission, 3, 6, 12 months and 2-7* yr
Oster 2013[44]	Sweden	n= 67, (M: 78%). Age: (mean: 43yr)	25% (20)	BSHS-B	6, 12 months and 2-7* yr
Palmu 2015[45]	Finland	n=92, (M: 70%). Age: (mean: 46yr)	10%	15D, EQ-5D, RAND-36	6 months
Renneberg 2014[46]	Germany	n=265, (M: 72%). Age: 16-73yr (mean: 39yr)	14% (14), Range: 1-76%	BSHS-B, SF-12	Admission, 6, 12, 24, and 36 months[$]
Ricci 2014[47]	Brazil	n= 73, (M: 69%). Age: (mean: 38yr)	14% (12)	BSHS-R	5 to 7 months
Roh 2012[48]	South Korea	n= 113, (M: 71%). Age: (mean: 38yr)	14% (12)	BSHS-B	1 month
Tahir 2011[49]	Pakistan	n=99, (M: 68%). Age: 19-57yr (median: 30yr)	19%, range: 5-38%	SF-36	Admission, 5 and 6* months
Van Loey 2012[20]	Netherlands and Belgium	n=244, (M: 73%). Age: (mean: 39yr)	12% (11), Range 1-65%	EQ-5D	3 weeks, 3, 9, and 18 months[$]
Wasiak 2014[50]	Australia	n=99, (M: 75%). Age: (mean: 42yr)	19%	BSHS-B, SF-36	Preburn and 12* months
Willebrand 2006[51]	Sweden	n= 86, (M: 73%). Age: 15-85yr (mean: 43yr)	17% (14)	BSHS-B	Mean: 3.6yr (SD: 1.2yr)

Willebrand 2011[52]	Sweden	n= 94, (M: 76%). Age: 19-90yr (mean: 44yr)	23% (20)	BSHS-B, SF-36	6*, 12* and 24* months
Xie 2012[53]	China	n=20, (M: 70%). Age: (mean: 43yr)	84% (10)	BSHS-B, SF-36	≥2 years after discharge
Zhang 2014[54]	China	n=208, (M: 77%). Age: (mean: 42yr)	42% (27)	BSHS-B	≥2 years after discharge

[1]Study population: n=sample size; M=males; n.a=not applicable, [2]15D=15-dimensional health-related quality of life instrument, BSHS=Burn-specific Health Scale, BSHS-A=Burn-specific Health Scale-Abbreviated, BSHS-B=Burn-specific Health Scale-Brief, BSHS-RBA=Burn-specific Health Scale Revised, Brief and Adapted, EQ-5D=EuroQol five dimensions, RAND-36=RAND 36-item health survey, SIP=Sickness Impact Profile, SF-10=Medical Outcome Study Short Form-10 items, SF-12=Medical Outcome Study Short Form-12 items, SF-36=Medical Outcome Study Short Form-36 items, QOLS=Quality of Life Scale. *Measurement point used for predictor analysis in studies with ≥1 measurement point. [§]All measurement points were used as the dependent variable was long-term recovery pattern

Table 2. Risk of bias assessment according to the Quality of Prognostic Studies (QUIPS) tool (n=32)

Study	Study Participation	Study Attrition	Prognostic Factor Measurement	Outcome Measurement	Study Confounding	Statistical Analysis and Reporting	Total score
Ahuja 2016	Low	Low	Moderate	Low	Moderate	Low	8
Anzarut 2005	Moderate	Moderate	Moderate	Moderate	High	Moderate	13
Blalock 1994	Moderate	High	Moderate	Low	High	Low	11
Corry 2010	Moderate	High	Low	Low	Moderate	Low	10
Cromes 2002	Moderate	High	Moderate	Low	High	Moderate	13
Edgar 2013	Low	Low	Moderate	Low	Moderate	Low	8
Ekeblad 2015	Low	Moderate	High	Low	High	Low	11
Finlay 2014	Low	Moderate	Low	Low	Moderate	Low	8
Finlay 2015	Low	Moderate	Low	Low	Moderate	Low	8
Kildal 2001	Low	Moderate	Moderate	Low	Moderate	Low	9
Kildal 2004	Low	Moderate	Moderate	Low	Moderate	Low	9
Kildal 2005	Low	High	Moderate	Low	Moderate	Low	10
Knight 2017	Moderate	High	Low	Low	Low	Low	9
Leblebici 2006	Moderate	High	Moderate	Low	Low	Low	10
Low 2012	Low	Moderate	Low	Low	Moderate	Low	8
Moi 2007	Low	Moderate	Low	Low	Moderate	Low	8
Moi 2012	Low	Moderate	Moderate	Low	Moderate	Low	9
Novelli 2009	High	High	Moderate	Low	High	Moderate	14
Orwelius 2013	Low	Moderate	Low	Low	Low	Low	7
Oster 2011	Low	Moderate	Moderate	Low	Moderate	Low	9
Oster 2013	Low	Moderate	Moderate	Low	Moderate	Low	9
Palmu 2015	Low	Moderate	Low	Low	Moderate	Moderate	9
Renneberg 2014	Moderate	High	Low	Low	Low	Low	9

Ricci 2014	Moderate	High	Moderate	Low	Moderate	Low	11
Roh 2012	Moderate	High	Low	Low	Low	Low	9
Tahir 2011	Low	High	Moderate	Low	High	Moderate	12
Van Loey 2012	Low	High	Low	Low	Moderate	Low	9
Wasiak 2014	Low	High	Low	Low	Moderate	Low	9
Willebrand 2006	Low	High	Moderate	Low	Moderate	Low	10
Willebrand 2011	Low	Moderate	Moderate	Low	Moderate	Low	9
Xie 2012	Moderate	Moderate	Low	Low	Low	Low	8
Zhang 2014	Low	Moderate	Moderate	Low	Moderate	Moderate	10

Note. The total score was composed of the sum of the domain scores: low risk=1, moderate risk=2, high risk=3.

Predictors of HRQL

Twenty studies used multivariable analysis. One study[32] did not indicate significant ($p \leq 0.05$) predictors and was therefore not included in our analyses. Three studies applied two different HRQL instruments, resulting in 22 different prediction studies. Eleven of these studies were based on four cohorts. Due to the low number of studies, all of these studies were included in the examination. The studies investigated between five and 42 predictors. Overall, 114 different predictors were investigated, of which 38 were studied in more than one study (Figure 2). These were sixteen burn-specific, twelve psychological, six demographic, and four environmental factors (Table 3).

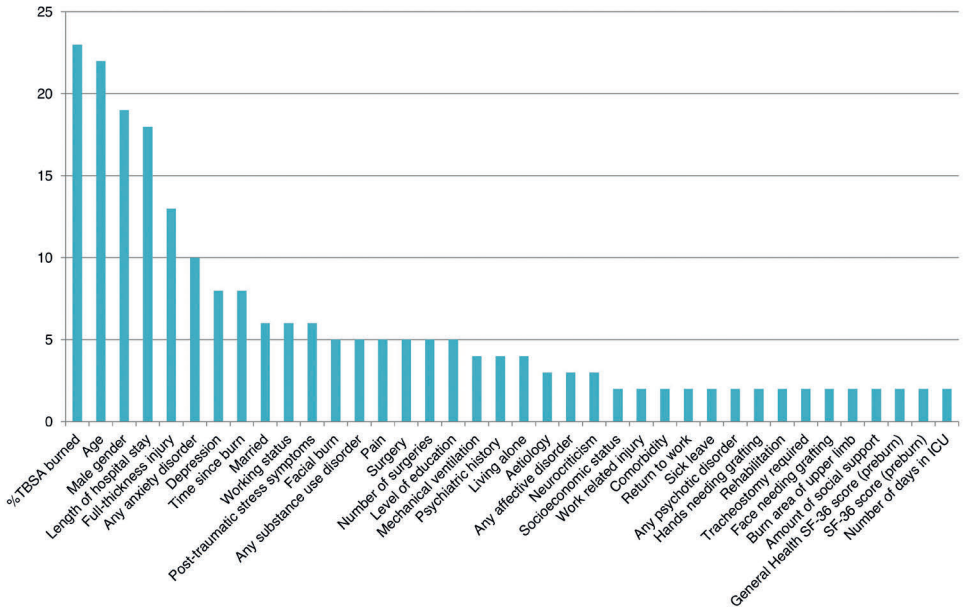


Figure 2. Predictors investigated in more than one multivariable predictive study

Demographic factors

The most studied demographic factors were age (n=21) and gender (n=21). The studies were inconsistent whether age is a predictor for HRQL. Among the studies that studied gender, eleven found that male gender was associated with a better HRQL and three reported an association but failed to describe the direction. Marital status, living alone, rehabilitation and level of education had no significant association with HRQL.

Environmental factors

The only environmental factor that showed an association with HRQL was postburn working status[18,44]. Four studies reported that having a job postburn was related with a better HRQL, and

two did not found an association. Preburn working status was only found to relate to a better HRQL in one of the four studies examining this predictor and none of the studies found a relation between socioeconomic status or work-related injury and HRQL.

Burn-specific factors

Percentage TBSA burned is the most often studied burn-specific predictor (n=18). Twelve studies found no association with HRQL, whereas five found a lower HRQL in more severely burned patients and one failed to describe the direction of the association. Somewhat more evidence exists on the LOS. Seven out of the thirteen studies reported a lower HRQL after a longer LOS. Both surgery and number of surgeries were studied as predictors. Two studies reported a positive association between surgery and HRQL, whereas one study reported a negative association and one did not find an association. A higher number of surgeries resulted in a decreased HRQL in two studies. Three other studies, however, found no statistically significant association. Five individual predictors (LOS, %TBSA burned, full-thickness injury, surgery and number of surgeries) are all indicators of the burn severity. The cluster burn severity is a significant predictor of a diminished HRQL in 13 out of the 18 studies that investigated this predictor. Having pain as predictor was investigated in five studies. Two found that patients that reported pain had a lower HRQL and three did not found an association. Evidence on other burn factors, including full-thickness injury, time since burn, hand burns, face needing grafting, upper limb burns, and mechanical ventilation was inconsistent. Studies found no association between either aetiology, hands needing grafting, facial burns or tracheostomy required and HRQL.

Psychological factors

Postburn depression or depressive symptoms and any preburn psychiatric disorder were most often studied (n=6). Four out of the six studies that investigated postburn depression reported an association with impaired HRQL. Also, evidence exist on higher levels of neuroticism and avoidant coping as predictors. The three studies that investigated these predictors all reported associations with poorer HRQL. Posttraumatic stress symptoms and less emotional or social support were also associated with diminished HRQL in the majority of studies. There was less evidence on preburn psychological factors (any psychiatric disorder, depression, substance use disorder and anxiety disorder) and HRQL. Studies were inconsistent on postburn substance use disorder as predictor and no association was found between any postburn psychiatric disorder and HRQL.

Table 3. Summary of 19 multivariable predictive studies of HRQL in adult burn patients

QUIPS score	8	8	8	8	8	9	9	9	9	9	9	9	9	9	9	9	10	10	10	12	13	13
	Edgar 2013 (SF-36)[1]	Edgar 2013 (BSHS-B)[1]	Finlay 2015 (BSHS-B)[1]	Low 2012 (BSHS-B)[4]	Xie 2012 (SF-36)	Renneberg 2014 (SF-12)	Van Loey 2012 (EQ-5D)	Kildal 2004 (BSHS-B)[2]	Wasiak 2014 (SF-36)[3]	Wasiak 2014 (BSHS-B)[3]	Moi 2012 (QOLS)	Palmu 2015 (RAND-36)	Oster 2011 (EQ-5D index)[4]	Oster 2011 (EQ-5D VAS)[4]	Oster 2013 (BSHS-B)[4]	Knight 2016 (BSHS-B)	Zhang 2014 (BSHS-B)	Kildal 2005 (BSHS-B)[2]	Willebrand 2006 (BSHS-B)	Tahir 2011 (SF-36)	Cromes 2002 (BSHS)	Anzarut 2005 (SF-36)
Demographic factors																						
Increasing age	-	+	0	+-	+	0	--	?	--	--	0	?	0	0	0	0	0	?	-	0		-
Male gender	+	++	++	+	0	+	++	?	+	++	0	?	0	0	0	++	++	?	+	0		0
Married					0							0					0					
Living alone											0		0	0	0							
Low level of education					0							0	0	0	0		0					
Rehabilitation					0												0					
Environmental factors																						
Low SES					0												0					
Work related injury																	0					0
Preburn working status												0	0	++	0							
Working status postburn					+						0		++	++	+		0					
Burn-specific factors																						
High %TBSA burned	--	--	0	-	0		0		0	0	0	?	0	0	0	0	--		-	0		0
Full-thickness injury			0	-	0				--	0	0		0	0	0		0					-
Longer length of stay				-	+				0	0	0	0	--	0	--	--	0			--		0
Surgery	+	++	--														0					
Number of surgeries							--		0	0	0									--		
Burn area								?				0						?		--		

Hand burns					0							-	0	0	+		--					0
Hands needing grafting					0																	0
Facial burns					0								0	0			0					0
Face needing grafting					+																	0
Upper limb burn			++														0					
Mechanical ventilation					0				0	0							--					0
Tracheostomy required					0						0											0
Pain											0		--	0			--				0	
Aetiology					0							0								0		
Longer time since burn	0	++			0			?					0	0	0		0	?	0			0
Psychological factors																						
Any preburn psychiatric disorder				0								0	0	0	0	--						
Any postburn psychiatric disorder				0								0	0	0								
Post-traumatic stress disorder or symptoms							--					?	0	--	-							
Preburn depression				-								?	0	0								
Postburn depression or depressive symptoms						-	--					--	0	0	-							
Preburn substance use disorder				-								?	0	0	0							
Postburn substance use disorder												?	0	--								
Preburn anxiety disorder				-									0	0	0							
Avoidant coping						-												-	-			
Emotional/social support																	0	+				+
Neuroticism						-		--											-			
Body image											0											0

Note. Studies are ordered according to QUIPS score and number of patients. ++ positive significant correlation (p≤0.05) with HRQL, + positive significant correlation with a domain(s) of HRQL only, 0 no significant correlation (p>0.05) with HRQL,-- negative significant correlation with HRQL, - negative significant correlation with a domain(s) of HRQL only, ? direction of correlation not reported, %TBSA=percentage total body surface area, [1,2,3,4]Based on the same dataset.

3

DISCUSSION

This study aimed to systematically review predictors of HRQL following burn injuries. Thirty-two studies were included and 114 predictors were investigated in 19 studies using multivariable analysis. Among burn patients, burn severity factors and psychological factors and to a lesser extent demographic and environmental factors are related to HRQL. Severity of burns, postburn depression, posttraumatic stress symptoms, avoidant coping, less emotional or social support, higher levels of neuroticism and unemployment postburn were found to predict poorer HRQL after burns. In addition, some weaker predictors, including female gender, pain and a postburn substance use disorder were identified. Other demographic and environmental factors showed in general no significant association with HRQL and the evidence was inconclusive on other burn-specific and psychological factors. The quality of these studies was in general moderate.

This review clearly indicates that the severity of burns is a strong predictor of HRQL following burns. More severe burns result in general in a poorer HRQL. It is however not yet clear which individual severity predictor (e.g. LOS, %TBSA burned, number of surgeries) is best to indicate the severity of burns. By studying the multivariable results, the most optimal predictor becomes visible. The optimal predictor differed among the studies. The most consistent severity indicators for the prediction of HRQL seem to be LOS and number of surgeries. In the general trauma population, LOS is also a predictor of HRQL[55,56] and there are some indications that surgical procedures predict a diminished HRQL[57]. The evidence regarding burn size was inconclusive; %TBSA burned was found to be negatively associated with HRQL in a minority (29%) of the studies. The other studies did not report a statistical significant association. Remarkable is that three out of the five larger studies (>200 patients) reported a negative association, suggesting that %TBSA burned is a predictor of diminished HRQL after burns. However, it is questionable whether %TBSA burned is a good proxy for the severity of burns. It reflects the sum of the estimated percentage of full and partial thickness burns; it does not distinguish between deep and superficial wounds. Other burn-specific factors, including LOS or number of surgeries, may be better predictors[20]. Or possibly a combination of severity indicators may be the best predictor. There are also indications that having pain is a predictor for having a poorer HRQL after burns[18,54]. It is known from other fields that patients who have severe continuing pain often also have a low HRQL[58,59]. Other burn-specific factors, including body region burned, aetiology and longer time since burn did in general not seem to influence HRQL to a large extend.

Psychological factors are also important predictors for HRQL following burns. Five of the seven strong predictors are psychological factors, including postburn depression, posttraumatic stress symptoms, avoidant coping, less emotional or social support and higher levels of neuroticism. These psychological factors are also predictors in other trauma populations[60-63]. Also a postburn substance use disorder seems to be a predictor of an impaired HRQL, although evidence

regarding this factor is weaker, both for burns and for trauma in general[64]. The often traumatic nature of burns may result in induced psychopathological responses[65], which is related to a poorer HRQL. Psychological burden can be caused by pain, grief, change of body image, self-blame, feelings of guilt, social isolation during hospital admission or permanent physical disabilities[65]. In addition, earlier studies showed an association between psychological and physical burden. Psychological burden was associated with delayed wound healing[66], with greater physical impairment and role disruption[67], with slower physical recovery[67] and with poorer postburn adjustment[68]. The underlying reason of this relation is not yet clear. On the one hand, psychological distress might be influenced by physical problems[69]; those who appraise their injuries as more severe might have an increased risk of psychological problems. On the other hand, individuals with psychological problems might appraise their condition as worse and their recovery as less complete, and might have a decreased intention to be involved in rehabilitation[67]. Regardless of the underlying reasons of this relation, increased psychological burden may result in an impaired HRQL.

The only demographic predictor of HRQL after burns was gender. Females reported a poorer HRQL after burns compared to males. This finding was also found in a recent study focussing on gender differences in HRQL outcomes in burn patients[70]. Reasons for females experiencing an impaired HRQL after burns are not clear. An explanation might be that females willingness to report problems is greater[71] or that women find it harder to live with a mutilated body. Females also reported higher levels of fatigue and higher mortality rates after burn injuries[71,72]. Besides, poorer outcomes in females have been shown in injury studies in general[73,74]. No strong conclusion could be drawn on the impact of age on HRQL after burns. Some studies reported better HRQL in younger adults, whereas others reported no relation or an adverse relation. These inconsistent results are also seen in the general trauma population[55,62,64,75].

Theoretically you would expect burn-specific instruments to be more sensitive to the consequences of burns. Thus, more statistically significant associations with HRQL measured by a burn-specific instrument would be expected. This was seen in present study. Burn-specific instruments had a higher proportion of significant associations in multivariable studies. Forty-nine (47%) significant associations out of the 104 studied associations were found when HRQL was measured with a burn-specific instrument. For generic instruments, 45 (28%) out of the 163 studied associations were significant. The burn-specific instruments thus seem to be more sensitive compared to the generic instruments used. This finding is in line with the results of an earlier study that compared the BSHS-B against the SF-36. That study showed that SF-36 summary scores were less sensitive than the BSHS-B total score. The domain scores of the SF-36, however, were more sensitive that the domain scores of the BSHS-B[76]. Most included studies in present review used SF-36 summary scores and BSHS-B domain scores.

The risk of bias of included studies was in general moderate. It was remarkable that none of the studies had an overall low risk of bias. In general, the risk of bias was moderate. A moderate or high risk of bias was in particular seen in the domains 'study attrition' and 'study confounding'. Only a minority of the studies set hypotheses before testing predictors and only a few underpinned their search for predictors with the available literature. Most studies did not report how missing data was handled. Besides, confounders were often not defined, attempts to collect information on patients who dropped out were not described and key characteristics on those lost to follow-up were not reported. Future studies should include these factors in order to decrease the risk of bias and improve the overall study quality. Another issue was the use of multivariable analysis in 20 of the 32 included studies, indicating that 38% only used univariable analysis. As HRQL is a multifactorial concept, it is likely that HRQL is influenced by several factors and therefore multivariable analysis seems indicated. Univariable analyses are not very informative due to relations among the predictors.

Strengths and limitations

A strength of this study is that it presents a comprehensive overview of predictors of a HRQL following burn injuries. Relevant literature databases were searched by an experienced librarian and quality was assessed using the wide applied QUIPS tool. A limitation is the exclusion of studies written in other languages than English, which might have resulted in missed studies published in other languages. Another limitation is the absence of a formal meta-analysis. Due to variation in instruments, time points and data presentation in combination with the low number of studies, it was not possible to formally pool the results using meta-analysis. The examination of predictors on the basis of its direction and statistical significance that we applied does not take into account the sample size of the study nor the strength of predictors. However, we have checked that our main outcomes were not conditioned on sample size, risk of bias or studies on the same dataset (Table 3). Due to the wide variation of assessment time points and the limited availability of short-term predictive studies, we were unable to study whether predictors differ in the short- and long-term.

CONCLUSIONS

HRQL after burn injuries is particularly affected by the severity of burns and the psychological response of an individual to the trauma. Both constructs provide unique information and knowledge that is necessary for optimized follow-up treatment and rehabilitation. Therefore, a comprehensive approach, including both physical and psychological care is indicated in the aftermath of burns. Screening of patients during follow-up is valuable to identify those patients who are in need of extra rehabilitation care. Patient-oriented treatment should be given and information on HRQL should be used to enhance patient-centred decision making.

To gain further insight in individual predictors and how they are correlated with each other, future studies should be based on the best available literature or on a theoretical framework, use larger sample sizes and ensure high methodological quality. As it is hard to collect large samples in burns, combining several existing datasets is highly recommended.

Funding

The Dutch Burn Foundation supported in this study, grant number: 15.102

Authors' contributions

IS conceptualized and designed the study, collected, analyzed and interpreted the data, and drafted the manuscript. CML conceptualized and designed the study, collected, analyzed and interpreted the data, and reviewed and revised the manuscript. JD conceptualized and designed the study and reviewed and revised the manuscript. NEvL, SP and MEB conceptualized and designed the study, analyzed and interpreted data, and reviewed and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors gratefully acknowledge Wichor Bramer (Biomedical information specialist, medical Library, Erasmus MC) for performing the database search.

REFERENCES

1. Falder S, Browne A, Edgar D, et al. Core outcomes for adult burn survivors: a clinical overview. Burns 2009;35:618-41.
2. Pereira C, Murphy K, Herndon D. Outcome measures in burn care: is mortality dead? Burns 2004;30:761-71.
3. Ferrans CE, Zerwic JJ, Wilbur JE, et al. Conceptual model of health-related quality of life. J Nurs Scholarsh 2005;37:336-42.
4. Guyatt GH, Jaeschke R, Feeny DH, et al. Measurements in clinical trials: choosing the right approach. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 21996. p. 41-8.
5. WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): position paper. Soc Sci Med 1995;41:1403-9.
6. World Health Organization. International Classification of Functioning, Disability and Health: ICF. 2001.
7. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. JAMA 1995;273:59-65.
8. Roop J, Payne J,et al. Theories and conceptual models to guide quality of life research. Practice 2011:45-51.
9. Moi AL, Nilsen RM. Pathways leading to self-perceived general health and overall quality of life in burned adults. Burns 2012;38:1157-64.
10. Van Baar M, Essink-Bot M-L, Oen I, et al. Functional outcome after burns: a review. Burns 2006;32:1-9.
11. Klinge K, Chamberlain DJ, Redden M, et al. Psychological adjustments made by postburn injury patients: an integrative literature review. J Adv Nurs 2009;65:2274-92.
12. Jasper S, Rennekampff H, de Zwaan M. Psychiatric co-morbidity, body image problems and psychotherapeutic interventions for burn survivors: a review. Psychother Psychosom Med Psychol 2013;63:423-8.
13. Dyster-Aas J, Kildal M, Willebrand M. Return to work and health-related quality of life after burn injury. J Rehabil Med 2007;39:49-55.
14. Moi AL, Wentzel-Larsen T, Salemark L, et al. Impaired generic health status but perception of good quality of life in survivors of burn injury. J Trauma Inj Infect Crit Care 2006;61:961-8.
15. Pfitzer B, Katona LJ, Lee SJ, et al. Three years after black Saturday: Long-term psychosocial adjustment of burns patients as a result of a major bushfire. J Burn Care Res 2016;37:e244-e53.
16. Sackett D, Haynes R, Guyatt G, et al. Clinical epidemiology: a basic science for clinical medicine. 2 ed. Boston: Little, Brown & Company; 1991.
17. Edgar DW, Homer L, Phillips M, et al. The influence of advancing age on quality of life and rate of recovery after treatment for burn. Burns 2013;39:1067-72.
18. Öster C, Willebrand M, Ekselius L. Health-related quality of life 2 years to 7 years after burn injury. J Trauma Inj Infect Crit Care 2011;71:1435-41.
19. Stavrou D, Weissman O, Tessone A, et al. Health related quality of life in burn patients–a review of the literature. Burns 2014;40:788-96.

20. Van Loey NE, van Beeck EF, Faber BW, et al. Health-related quality of life after burns: a prospective multicenter cohort study with 18 months follow-up. J Trauma Acute Care Surg 2012;72:513-20.
21. Yoder LH, Nayback AM, Gaylord K. The evolution and utility of the burn specific health scale: a systematic review. Burns 2010;36:1143-56.
22. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
23. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.
24. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280-6.
25. Varni JW, Burwinkle TM, Sherman SA, et al. Health-related quality of life of children and adolescents with cerebral palsy: hearing the voices of the children. Dev Med Child Neurol 2005;47:592-7.
26. Meirte J, Van Loey NEE, Maertens K, et al. Classification of quality of life subscales within the ICF framework in burn research: identifying overlaps and gaps. Burns 2014;40:1353-9.
27. Ahuja RB, Mulay AM, Ahuja A. Assessment of quality of life (QoL) of burn patients in India using BSHS-RBA scale. Burns 2016;42:639-47.
28. Anzarut A, Chen M, Shankowsky H, et al. Quality-of-life and outcome predictors following massive burn injury. Plast Reconstr Surg 2005;116:791-7.
29. Blalock SJ, Bunker BJ, DeVellis RF. Measuring health status among survivors of burn injury: revisions of the Burn Specific Health Scale. J Trauma Acute Care Surg 1994;36:508-15.
30. Corry NH, Klick B, Fauerbach JA. Posttraumatic stress disorder and pain impact functioning and disability after major burn injury. J Burn Care Res 2010;31:13-25.
31. Cromes GF, Holavanahalli R, Kowalske K, et al. Predictors of quality of life as measured by the burn specific health scale in persons with major burn injury. J Burn Care Rehabil 2002;23:229-34+7+8.
32. Ekeblad F, Gerdin B, Öster C. Impact of personality disorders on health-related quality of life one year after burn injury. Disabil Rehabil 2015;37:534-40.
33. Winthrop AL. Health-related quality of life after pediatric trauma. Curr Opin Pediatr 2010;22:346-51.
34. Finlay V, Phillips M, Allison G, et al. Towards more efficient burn care: identifying factors associated with good quality of life post-burn. Burns 2015;41:1397-404.
35. Kildal M, Andersson G, et al. Development of a brief version of the Burn Specific Health Scale (BSHS-B). J Trauma Acute Care Surg 2001;51:740-6.
36. Kildal M, Willebrand M, Andersson G, et al. Personality Characteristics and Perceived Health Problems after Burn Injury. J Burn Care Rehabil 2004;25:228-35.
37. Kildal M, Willebrand M, Andersson G, et al. Coping strategies, injury characteristics and long-term outcome after burn injury. Injury 2005;36:511-8.
38. Knight A, Wasiak J, Salway J, et al. Factors predicting health status and recovery

of hand function after hand burns in the second year after hospital discharge. Burns 2017;43:100-6.

39. Leblebici B, Adam M, Bağiş S, et al. Quality of life after burn injury: The impact of joint contracture. J Burn Care Res 2006;27:864-8.

40. Low AJF, Dyster-Aas J, Willebrand M, et al. Psychiatric morbidity predicts perceived burn-specific health 1 year after a burn. Gen Hosp Psychiatry 2012;34:146-52.

41. Moi AL, Wentzel-Larsen T, et al. Long-term risk factors for impaired burn-specific health and unemployment in patients with thermal injury. Burns 2007;33:37-45.

42. Novelli B, Melandri D, Bertolotti G, et al. Quality of life impact as outcome in burns patients. G Ital Med Lav Ergon 2009;31:A58-A63.

43. Orwelius L, Willebrand M, et al. Long term health-related quality of life after burns is strongly dependent on pre-existing disease and psychosocial issues and less due to the burn itself. Burns 2013;39:229-35.

44. Öster C, Willebrand M, Ekselius L. Burn-specific health 2 years to 7 years after burn injury. J Trauma Acute Care Surg 2013;74:1119-24.

45. Palmu R, Partonen T, Suominen K, et al. Health-related quality of life 6 months after burns among hospitalized patients: Predictive importance of mental disorders and burn severity. Burns 2015;41:742-8.

46. Renneberg B, Ripper S, Schulze J, et al. Quality of life and predictors of long-term outcome after severe burn injury. J Behav Med 2014;37:967-76.

47. Ricci H, Gonçalves N, Gallani MC, et al. Assessment of the health status in Brazilian burn victims five to seven months after hospital discharge. Burns 2014;40:616-23.

48. Roh Y, Chung H, et al. Association between depression, patient scar assessment and burn-specific health in hospitalized burn patients. Burns 2012;38:506-12.

49. Tahir SM, Memon MM, Ali SA, et al. Health related quality of life after burns: are we really treating burns? J Ayub Med Coll Abbottabad 2011;23:152-6.

50. Wasiak J, Lee SJ, Paul E, et al. Predictors of health status and health-related quality of life 12 months after severe burn. Burns 2014;40:568-74.

51. Willebrand M, Andersson G, Kildal M, et al. Injury-related fear-avoidance, neuroticism and burn-specific health. Burns 2006;32:408-15.

52. Willebrand M, Ekselius L. Health-related quality of life 2 years to 7 years after burn injury. Journal of Trauma and Acute Care Surgery 2011;71:1435-41.

53. Xie B, Xiao S, Zhu S, et al. Evaluation of longterm health-related quality of life in extensive burns: A 12-year experience in a burn center. Burns 2012;38:348-55.

54. Zhang L, Cao J, Feng P, et al. Influencing factors of the quality of life in Chinese burn patients: Investigation with adapted Chinese version of the BSHS-B. Burns 2014;40:731-6.

55. Taylor R, Sander J, Taylor R, et al. Predictors of health-related quality of life and costs in adults with epilepsy: A systematic review. Epilepsia 2011;52:2168-80.

56. Hoogewerf CJ, van Baar ME, Middelkoop E, et al. Impact of facial burns: relationship between depressive symptoms, self-esteem and scar severity. Gen Hosp Psychiatry 2014;36:271-6.

57. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: Reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient populations. Med Care 2001;39:800-12.

58. Becker N, Thomsen A, Olsen A, et al. Pain epidemiology and health related quality of life in chronic non-malignant pain patients referred to a Danish multidisciplinary pain center. Pain 1997;73:393-400.

59. Neria Y, Nandi A, Galea S. Post-traumatic stress disorder following disasters: a systematic review. Psychol Med 2008;38:467-80.

60. Batailler P, Hours M, et al. Health status recovery at one year in children injured in a road accident:A cohort study.Accid Anal Prev 2014;71:267-72.

61. Davidson TN, Bowden ML, Tholen D, et al. Social support and post-burn adjustment. Arch Phys Med Rehabil 1981;62:274-8.

62. Kaugars AS, Silverman A, Kinservik M, et al. Families' perspectives on the effect of constipation and fecal incontinence on quality of life. J Pediatr Gastroenterol Nutr 2010;51:747-52.

63. Moi AL, Haugsmyr E, Heisterkamp H. Long-term study of health and quality of life after burn injury. Ann Burns Fire Disasters 2016;29:295.

64. Rumsey N, Clarke A, White P, et al. Altered body image: Appearance-related concerns of people with visible disfigurement. J Adv Nurs 2004;48:443-53.

65. Van Loey NEE, Van Son MJM. Psychopathology and psychological problems in patients with burn scars. Am J Clin Dermatol 2003;4:245-72.

66. Walburn J, Vedhara K, et al. Psychological stress and wound healing in humans: a systematic review and meta-analysis. J Psychosom Res 2009;67:253-71.

67. Fauerbach JA, Lezotte D, Hills RA, et al. Burden of burn: a norm-based inquiry into the influence of burn size and distress on recovery of physical and psychosocial function. J Burn Care Res 2005;26:21-32.

68. Fauerbach J, Lawrence J, Haythornthwaite J, et al. Pre-injury psychiatric illness and postinjury adjustment in adult burn survivors. Psychosomatics 1996;37:547-55.

69. Cohen S, Rodriguez MS. Pathways linking affective disturbances and physical disorders. Health Psychol 1995;14:374.

70. Wasiak J, Lee SJ, Paul E, et al. Female patients display poorer burn-specific quality of life 12 months after a burn injury. Injury 2017;48:87-93.

71. Gabbe BJ, Cleland H, Watterson D, et al. Predictors of moderate to severe fatigue 12 months following admission to hospital for burn: Results from the Burns Registry of Australia and New Zealand (BRANZ) Long Term Outcomes project. Burns 2016;42:1652-61.

72. Moore E, Pilcher D, Bailey M, et al. The Burns Evaluation and Mortality Study (BEAMS): predicting deaths in Australian and New Zealand burn patients admitted to intensive care with burns. J Trauma Acute Care Surg 2013;75:298-303.

73. Polinder S, van Beeck E, Essink-Bot M, et al. Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma Acute Care Surg 2007;62:133-41.

74. Holbrook TL, Hoyt DB, Anderson JP. The importance of gender on outcome after major trauma: functional and psychologic outcomes in women versus men. J Trauma Acute Care Surg 2001;50:270-3.

75. Eiser C, Morse R. Can parents rate their child's health-related quality of life? Results of a systematic review. Qual Life Res 2001;10:347-57.

76. Edgar D, Dawson A, Hankey G, et al. Demonstration of the validity of the SF-36 for measurement of the temporal recovery of quality of life outcomes in burns survivors. Burns 2010;36:1013-20.

4

Health-related quality of life in children after burn injuries: a systematic review

I. Spronk, **C.M. Legemate**, S. Polinder, M.E. van Baar

J Trauma Acute Care Surg. 2018 Dec;85(6):1110-1118

ABSTRACT

Background

Through improved survival of burns, more children have to deal with consequences of burns. Health-related quality of life (HRQL) measurement is important to qualify the perceived burden of burns in children. No systematic study of this outcome in children exists. Therefore, our objective was to review study designs, instruments, methodological quality, outcomes and predictors of HRQL in children after burns.

Methods

A systematic literature search was conducted in CINAHL, Embase, Google Scholar, Medline, The Cochrane library and Web of science (PROSPERO ID=CRD42016048065). Studies examining HRQL in pediatric burn patients were included. The risk of bias was assessed using the Quality in Prognostic Studies (QUIPS) tool.

Results

Twenty-seven studies using twelve HRQL instruments were included. The Burns Outcome Questionnaire 0-4 and 5-18 years old were most often applied. All longitudinal studies showed improvement of HRQL over time. However, problems were reported on the longer term on the domains '(parental) concern' and 'appearance'. Parental proxy scores were in general comparable to children's self-ratings. Severity of burns, facial burns, hand burns, comorbidity and short time since burn predicted an impaired HRQL. The risk of bias of the studies was in general moderate.

Conclusions

HRQL in children after burns increases over time. Domains and patient groups that require special attention are identified. However, due to lack of comparability of studies, the available information could not be used optimally. To further improve our understanding of HRQL, consensus on design, data-analysis and data presentation is needed.

BACKGROUND

Children, especially those under five years of age, are at the highest risk of hospitalization from burns[1-4]. Due to the improved survival of burns, more children have to deal with consequences of burn injuries[5-10]. This increases the importance of investigating outcomes of burns in children[10]. An important outcome is health related quality of life (HRQL)[11]. HRQL includes a child's or its parent's perception on the child's health status and assesses how burns affect the child's psychological, physical and social abilities[12].

Pediatric HRQL is often assessed by questionnaires. For young children, a parent often completes the questionnaire (a parent-proxy score). Older children sometimes answer questions themselves, or both the child and/or a parent completes the questionnaire[13]. This also depends on which questionnaire is used; not all questionnaires have child-report versions. Instruments to assess HRQL in burns are either generic (applicable to children with all kinds of conditions) or burn-specific (applicable to burned children). A generic instrument facilitates comparison among different illnesses, however, it does not take specific consequences of burns into account[14].

It is important to gain insight in HRQL after burns to qualify the impact of burns on children. As the maturation of scars and expression of psychological consequences may take several years, it is important to assess HRQL also at the longer-term. Information on HRQL can be used to support patient-centered decision making, measure provider performance, inform quality improvement in treatment and improve aftercare of pediatric burn patients[15].

Despite the prominent role of trauma in mortality and morbidity in children, relatively little research has been done in terms of HRQL in burned children. Earlier HRQL reviews focused on adults only[16-19]. Therefore, the current review aims to present an overview of study designs, instruments, methodological quality, outcomes and predictors of HRQL in children after burns.

METHODS

The protocol of this review is available on PROSPERO (ID=CRD42016048065; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=48065) and this review is conducted according to the PRISMA Statement[20].

Search strategy and eligibility criteria

The search strategy was developed in collaboration with a medical librarian, included terms covering HRQL and burns (Appendix 1) and was performed in relevant databases on 12 February 2018. Databases were searched from the earliest record until February 2018. We included

studies on HRQL in pediatric burn patients (<18 years old) written in English and published in a peer-reviewed journal. Studies in other pediatric patients that not presented HRQL outcomes for the burn population separately were excluded. Studies including both children and adults were only included if separate characteristics and outcomes were reported for children.

Study selection and data extraction

The systematic search was performed by an experienced librarian. Duplicates were removed and irrelevant articles excluded based on screening of the titles by one researcher (IS). Ten percent of the remaining abstracts were independently appraised by two researchers (IS, CL). As no disagreement existed, the remaining abstracts were evaluated by one researcher (IS). All full text articles were screened by both researchers independently. Data included study and patient characteristics, details on HRQL instruments, and outcomes and was independently extracted by two researchers (IS, CL). Disagreements regarding eligibility and data extraction were resolved by consulting a third researcher (MvB).

Risk of bias assessment

We assessed the risk of bias of the studies with the Quality in Prognostic Studies (QUIPS) tool[21]. The domains 'study participation', 'study attrition', 'outcome measurement' and 'statistical analysis and presentation' were assessed in all studies. The domains 'prognostic factor measurement' and 'study confounding' were only assessed in studies investigating predictors as these domains are specific for prognostic studies. The domains were rated as 'low' bias (all items 'low risk'), 'moderate' bias (max. 50% items with high or unknown risk of bias) or 'high' risk of bias (>50% items high of unknown risk of bias)[21]. Two researchers (IS, CL) independently assessed 26% of the studies. As there was only a slight disagreement between researchers (<10%), the remaining studies were assessed by one researcher (IS). In case of any doubt, a domain was assessed by a second researcher.

Data analysis

Outcomes of the most applied instruments were assessed. Children's and parent-proxy scores were compared. In case of unequal age groups for children and proxy scores, data was requested. We evaluated predictors of HRQL when a regression coefficient was presented. All kind of predictors were considered and both uni- and multivariable predictors of domains of HRQL as for total HRQL were tabulated. Predictors were divided into three categories: patient, burn-specific, and psychological factors. The predictors were summarized based on their direction and statistical significance[22,23]. Predictors were considered strong when ≥67% of the associations were in the same direction and statistically significant and weak if ≥33 - <67% of the associations fulfilled these conditions.

RESULTS

Identification and selection of studies

Our literature search resulted in 4,423 unique records. After title and abstract screening, 165 records were selected. Most studies (n=126) were conducted with adults and therefore excluded (Figure 1). Twenty-seven studies met our inclusion criteria.

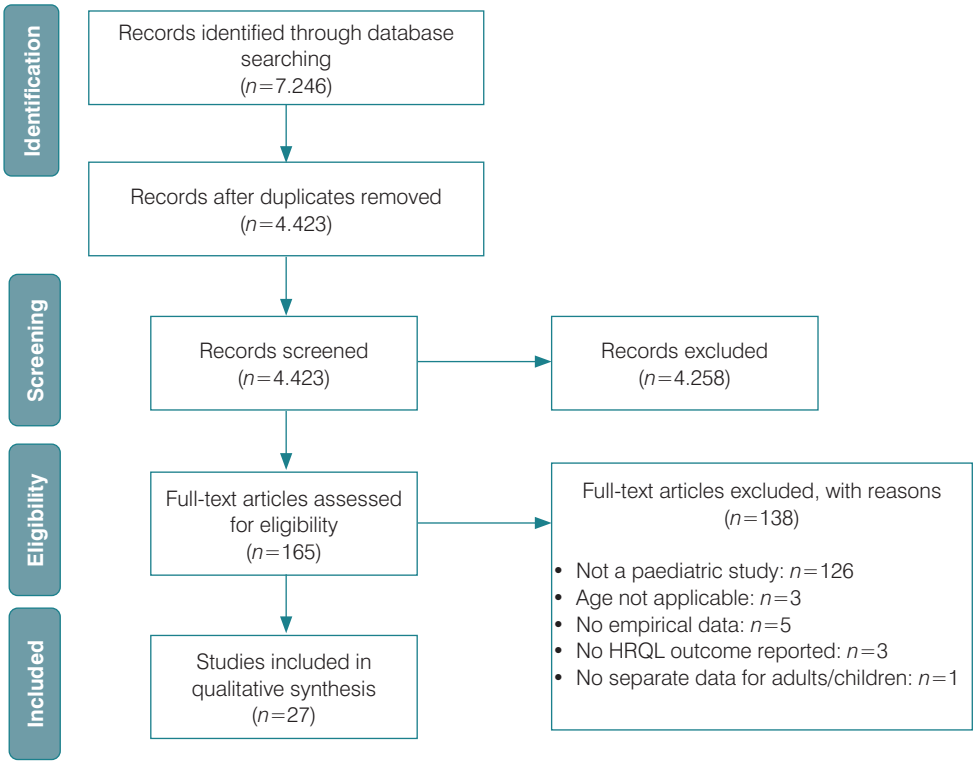


Figure 1. Flowchart outlining selection of studies

Study characteristics

The sample sizes ranged between 31 and 1,140 pediatric burn patients (Table 1). Twenty studies (77%) included more boys than girls. Seventeen studies were based on five cohorts. Thirteen studies were conducted in Europe and eleven in the USA. Most (67%) were published after 2010. Six studies were conducted in children aged 0-4 years old[24-29], eleven in children ≥5 years old[28,30-39], five in adolescents ≥11 years old[40-44], and five included all ages between 0 and 18[45-49]. The mean percentage total body surface area (%TBSA) burned ranged between 6-53% with almost half of the studies (44%) having a mean %TBSA burned of more than 20%[26,30,34-36,39-41,44-47]. Parents completed the questionnaires in twelve studies[24-27,29,33,37-39,45,48,50,] both a parent and the child were asked in ten studies[28,30,34-36,40,42,46,47,49], and five studies only surveyed children[31,32,41,43,44].

Table 1. Characteristics of studies

First author, year	Country	Study population[1]	%TBSA burned, mean (SD)	HRQL instrument(s)[2]	Assessment time point(s)
Daltroy 2000[1]	USA	n= 186, (M: 71 %). Mean age: 10yr	22%	BOQ (5-18)	NA
Dodd 2010[2]	USA	n= 145, (M: NA). Mean age: NA	16% (14) no hand burn, 23% (26) hand burn	BOQ (0-4), BOQ (5-18)	Admission, first visit after discharge, 3, 6, 12, 18, 24 months
Kazis 2002[3]	USA	n= 184, (M: 54%). Mean age: 3yr	17% (18)	BOQ (0-4)	Admission and 6 months
Kazis 2012[4]	USA	n= 1140, (M: 68%). Mean age: 7yr	33% (23)	BOQ (0-4), BOQ (5-18)	Baseline, 3, 6, 12, 18, 24, 36 and 48 months
Kazis 2016[5]	USA	n= 336, (M: 59%). Mean age: 2yr	18% (18)	BOQ (0-4)	at discharge, 1, 3, 6, 12, 18, 24, 36, and 48 months.
Laitakari 2015[6]	Finland	n= 44, (M: 64%). Mean age: 7yr	10%	17D	Median 6.3yr
Landgraf 2013[7]	The Netherlands	n= 194, (M: 54%). Mean age: 3yr	6%	ITQOL	Mean: 18 months (SD: 10 months)
Landolt 2002[8]	Switzerland	n= 105, (M: 65%). Mean age: 11yr	18% (11)	TACQOL	1-13 years postburn
Landolt 2009[9]	Switzerland	n= 43, (M: 65%). Mean age: 10yr	13% (15)	TACQOL	Mean: 4.4yr (SD: 2.0yr)
Maskell 2013[10]	Australia and New Zealand	n= 66, (M: 25%). Mean age: 13yr	23% (20)	PedsQL	Mean: 7.3yr (SD: 4.7yr)
Maskell 2014[11]	Australia and New Zealand	n= 63, (M: 25%). Mean age: 13yr	23% (20)	PedsQL	Mean: 7.3yr (SD: 4.7yr)
Meyer 2012[12]	USA	n= 355, (M: 79%). Mean age: 13yr	35%	BOQ (5-18)	Baseline, 3, 6, 12, 18, 24, 36 and 48 months
Nicolosi 2013[13]	Brazil	n= 63, (M: 30%). Mean age: 16yr	24%	BSHS-R	NA
Palmieri 2012[14]	USA	n= 438, (M: 88%). Mean age: 2yr	61%	BOQ (0-4)	Baseline, 3, 6, 12, 18, 24, 36 and 48 months
Pan 2015[15]	Netherlands and Belgium	n= 54, (M: 74%). Mean age: 14yr	28% (22)	BOQ (5-18)	6 months and 18 months.
Patrick 2007[16]	USA and UK	n= 104, (M: 53%). Mean age: NA	NA	YQOL-FD, YQOL-R	≥2 years postburn
Pope 2007[17]	UK	n= 36, (M: 36%). Mean age: 15yr	23% (17)	YQOL	Mean: 11.8yr
Rosenberg 2013[18]	USA	n=31, (M: 81%). Mean age: 14yr	58% (12) intervention, 49% (8) control	CHQ	Discharge and 3 months
Spuijbroek 2011[19]	The Netherlands	n= 194, (M: 54%). Mean age: 3yr	6%	ITQOL	Mean: 18 months (SD: 10 months)

Stubbs 2011[20]	USA	n= 390, (M: 69%). Mean age: 7yr	36% (23)	BOQ (0-4), BOQ (5-18)	3, 6, 12, 18 and 24 months
Sveen 2012[21]	Sweden	n= 70, (M: 67%). Mean age: 10yr	11% (13)	BOQ (5-18)	Mean 5.4yr (SD: 2.4yr)
Sveen 2014[22]	Sweden	n= 109, (M: 64%). Mean age: 3yr (group0-4yr), 10 yr (group 5-18 yr)	7% (4) (0-4yr), 11% (13) (5-18 yr)	BOQ (0-4), BOQ (5-18)	Mean 1.8yr (SD: 0.9yr) (group 0-4 yr) Mean 5.4yr (SD: 2.4yr) (group 5-18 yr)
Van Baar 2006[23]	The Netherlands	n= 194, (M: 54%). Mean age: 3yr	6%	BOQ (0-4), ITQOL	Mean: 18 months (SD: 10 months)
Van Baar 2006[24]	The Netherlands	n= 145, (M: 65%). Mean age: 9yr	6% (2)	BOQ (5-18), CHQ, EQ-5D	Mean: 21 months (SD: 10 months)
Van Baar 2011[25]	The Netherlands	n= 132, (M: 61%). Mean age: NA	NA	BOQ (5-18), EQ-5D	Mean: 9 and 24 months
Warner 2012[26]	USA	n= 678, (M: 72%). Mean age: 9yr	35% (24)	BOQ (5-18)	First visit after discharge, 6, 12, 18, 24, 36 and, 48 months
Weedon 2011[27]	South-Africa	n= 70, (M: NA). Mean age: 4yr	11% (10)	PedsQL	Discharge and 3 months

[1]Study population: n=sample size; M= males; NA= not applicable. [2]BOQ=Burns Outcome Questionnaire, CHQ=Child Health Questionnaire, EQ-5D=EuroQol five dimensions questionnaire, ITQOL=Infant Toddler Quality of Life Questionnaire, PedsQL=Pediatric Quality of Life Inventory, TACQOL=TNO AZL Child Quality Of Life, 17D=17-dimensional health-related quality of life instrument, BSHS-R=Burn Specific Health Scale – Revised, YQOL=Youth Quality of Life Questionnaire, YQOL-FD=Youth Quality of Life Questionnaire - Facial Differences, YQOL-R=Youth Quality of Life Instrument Research Version–Adolescent

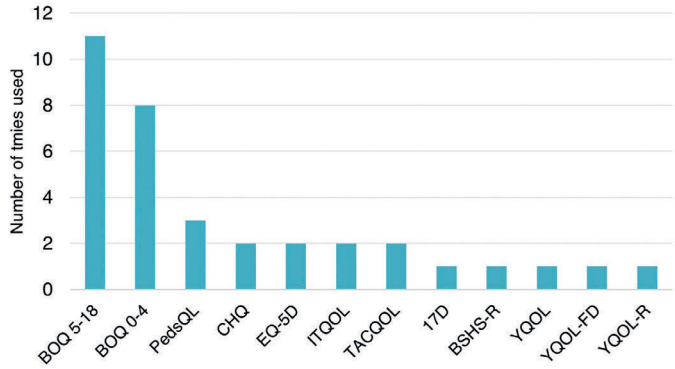


Figure 2. Instruments used to measure health-related quality of life in burn patients

BOQ = Burns Outcome Questionnaire, CHQ = Child Health Questionnaire, EQ-5D = EuroQol five dimensions questionnaire, ITQOL = Infant Toddler Quality of Life Questionnaire, PedsQL = Pediatric Quality of Life Inventory, TACQOL = TNO AZL Child Quality Of Life, 17D = 17-dimensional health-related quality of life instrument, BSHS-R = Burn Specific Health Scale - Revised, YQOL = Youth Quality of Life Questionnaire, YQOL-FD = Youth Quality of Life Questionnaire - Facial Differences, YQOL-R = Youth Quality of Life Instrument Research Version-Adolescent

HRQL measurement

Twelve HRQL instruments were used; nine generic and three burn-specific instruments. The Burns Outcome Questionnaire (BOQ) 0-4 years old (n=8)[29] and the BOQ 5-18 years old (n=11)[30] were most often used (Figure 2). The BOQ are multidimensional instruments. The 0-4 instrument includes 10 domains: play, language, fine motor, gross motor, behavior, family, pain/itching, appearance, satisfaction, concern/worry. The 5-18 instrument consists of 12 domains: upper extremity function, physical function and sports, transfers and mobility, pain, itch, appearance, compliance, satisfaction with current state, emotional health, family disruption, parental concern, school re-entry. Four studies[28,38,43,50] used more than one instrument to assess HRQL (Table 1). Three studies used both a burn specific and a generic questionnaire. HRQL was most often measured at three, six, twelve, eighteen and twenty-four months (Figure 3).

Quality assessment

Almost all studies had a low risk of bias on the items 'outcome measurement' (96%) and 'statistical analysis and reporting' (85%) (Figure 4). On the contrary, all studies had a high or moderate risk of bias for 'study attrition'. Most of the prediction studies had a moderate risk of bias on both 'prognostic factor measurement' and 'study confounding'. Seven studies had a low risk of bias on all except one domain[25,27,28,32,37,48,50] (Appendix 2).

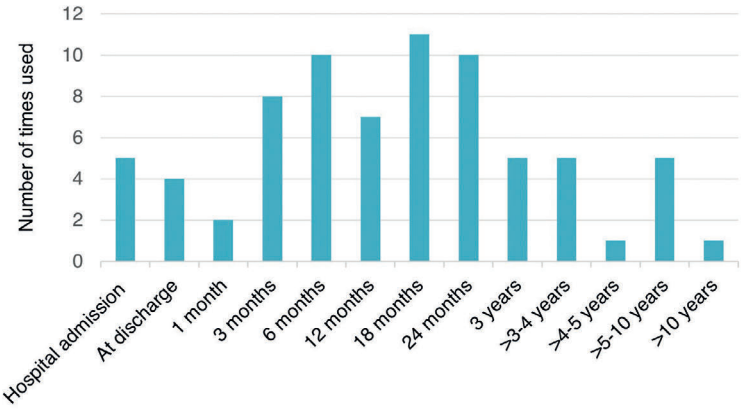


Figure 3. Time point at which health-related quality of life was assessed

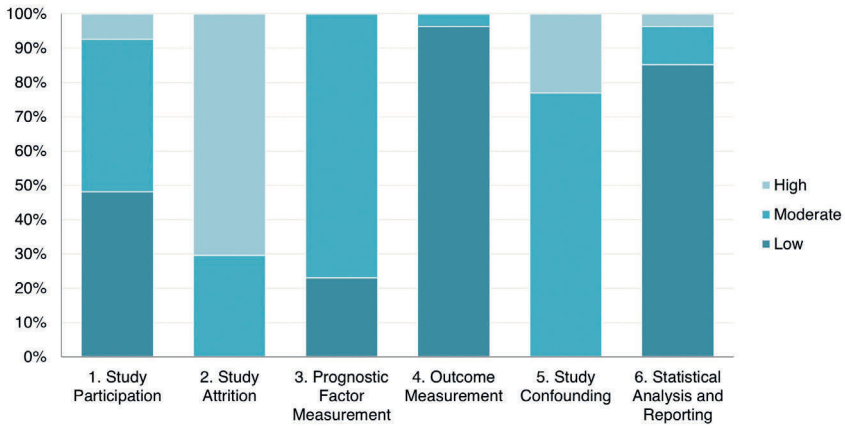


Figure 4. Risk of bias assessed with the Quality in Prognostic Studies (QUIPS) risk of bias tool.

Note. Domain 3 and domain 5 were only assessed for the thirteen prediction studies.

HRQL outcomes

Due to a large number of studies based on the same datasets, differences in instruments and data presentation, and the use of two norm scores for the same instrument, we were unable to pool the data. Therefore, we examined the outcomes of the individual longitudinal studies. All reported improvement of HRQL over time[24,26,29,39,40,42,45-47], especially in the first six months after burns. Afterwards improvement continues slowly (at least up to 48 months), but levels off gradually[24,26,40,46]. There are some domains that seem to improve less: 'concern/worry' and 'appearance' for young children (BOQ 0-4) and 'appearance' and 'parental concern' for older children (BOQ 5-18).

Parental proxy versus children's scores

Although eleven studies examined children and parent scores, only five reported these scores from the same children. Three studies used the BOQ 5-18[40,42,50] and two the PedsQL[34,35]. Three studies compared the scores at several time points[35,40,42]. In total, 14 sets of data comprising 142 (domain) scores were compared. Absolute scores were in general comparable (Appendix 3a+3b). Whether differences were statistically significant was examined in three studies. One study found no significant differences[50], the two others reported significant differences on three of the 12 domains[42] and six of the 72 domains[40]. Both found significant differences on the domains 'appearance' and 'family disruption'[40,42]. Children reported better scores on 'appearance' in both studies. Results on family disruption were inconsistent; in one study children reported better scores, in the other parents did.

Predictors of HRQL

Thirteen studies investigated predictors of HRQL. Seven studies used univariable and six multivariable analysis (Appendix 4). Between one and 16 predictors were investigated per study. In total, 39 different predictors were studied. These were 22 patient, 15 burn-specific and two psychological factors. Predictors (n=13) assessed by at least two studies are listed in Table 2.

Patient characteristics

Studies were inconsistent whether gender is a predictor. Two studies reported a better HRQL in boys, one in girls, and four did not found an association. More evidence exists on young age at burn. Three out of five studies reported a better HRQL when children were burned at a young age. Age at time of assessment showed inconsistent results. Comorbidity was associated with a poorer HRQL in two studies and one study found no association.

Burn-specific factors

The three predictors covering burn severity (%TBSA burned, total full-thickness injury and LOS) were all related to an impaired HRQL. Also, localization of burns was studied. Five of the six studies found that facial burns were associated with a lower HRQL. Hand burns were also found to predict an impaired HRQL in three of the four studies. Less evidence was found on the predictor visible scars. A longer time since burn was related to a better HRQL in eight of the nine studies investigating this predictor. The two studies that investigated etiology reported both no relation with HRQL.

Table 2. Predictors of HRQL in pediatric burn patients

Predictor	No of studies that found positive associations	No of studies that found negative associations	No of studies that found no associations	Conclusion*
Patient characteristics				
Young age at burn	3	0	2	Weak predictor
Increasing age	2	2	0	No predictor
Male gender	2	1	4	No predictor
Socioeconomic status	0	0	3	No predictor
Comorbidity	0	2	1	Weak predictor
Burn-specific factors				
%TBSA burned	0	7	2	Strong predictor
Total full-thickness injury	0	4	0	Strong predictor
Longer hospital stay	0	3	2	Weak predictor
Visible scars	0	1	2	No predictor
Facial burns	0	5	1	Strong predictor
Hand burns	0	3	1	Strong predictor
Time since burn/injury	8	2	1	Strong predictor
Aetiology	0	0	2	No predictor

*Predictors are considered strong when ≥67% of the associations were in the same direction and statistically significant and weak if ≥33 - <67% of the associations fulfilled these conditions.

DISCUSSION

This review provides a comprehensive overview of instruments, outcomes and predictors of HRQL in children after burns. Twenty-seven studies using twelve HRQL instruments were included. The BOQ 0-4 and BOQ 5-18 were most widely applied. Twelve studies had a longitudinal design with different lengths and time assessment points. All of these reported an improved HRQL over time, however, problems were reported in the longer term on the domains '(parental) concern' and 'appearance'. Children reported comparable HRQL outcomes as their parent(s). Severity of burns, facial burns, hand burns, comorbidity and short time since burn were found to predict a poorer HRQL.

The relatively low number of studies on HRQL in children after burns is striking. Especially HRQL of young children ≤4 years old is hardly studied, though this is the most prevalent group of burn victoms[4]. This might be because it is harder to assess HRQL in young children. Until the age of five, proxy-reports are necessary as children are not able to provide self-reports. Obtaining valid information on HRQL of these children is crucial to provide good aftercare and should therefore be applied in the future.

Children ≥5 years old are able to reliably report their HRQL[51-53]. Children self-report is recommended as this is in line with the subjective nature of HRQL and parental proxy scores might result in observational bias. Parent views can be influenced by parental post-traumatic stress, concerns and their perspectives on pre-burn functioning of their child[42,54-56]. Our study showed, however, reassuring results of parental proxy scores; children's self-reported scores were comparable on most domains. Self-report in children aged 5-8 years old is challenging[51]. The use of toys or drawings might help to elicit valuable information[51]. Children >8 years old should in general be able to answer themselves. Only one of the studies[49] used self-reported data from children ≥5 years old. The most applied burn-specific instrument only has a child-report version for children ≥11 years old[30]. For future studies it is valuable to consider assessing self-reported HRQL from children ≥5 years old. To do this, a burn-specific child-report version for children ≥5 years old should be created. If proxy reports are used, it is important to be aware that outcomes might be different. Differences are especially prevalent in psychosocial domains[57].

Another important issue in the current literature is the lack of comparability among studies. The wide variations in instruments, differences in data presentation, and different norm scores for the same instrument, impeded us to pool and analyze the available data. Uniform data is needed to improve our global insight and understanding of HRQL in burned children. A uniform way of presenting results and a norm score of non-burned children that is widely available is necessary.

The most often used and only burn-specific HRQL instruments are the two BOQ instruments[29,30]. These instruments are not freely available, and there are two manners of scoring, which hamper comparison of outcomes. These disadvantages have led to initiatives to develop alternative burn-specific instruments[58-60]. However, these are unfortunately not yet available and researchers for who the BOQ instruments are not available have to use a generic instrument when assessing pediatric HRQL. More generic instruments are available, but no consensus exists on the most optimal instrument or instruments; a total of nine different instruments were used in the fourteen studies assessing generic HRQL. The choice for a specific generic instrument depends on several factors, including the age of the study population, the availability of an instrument in a specific language and whether the aim is to assess children's own HRQL or a parent proxy of their children's HRQL. Given these considerations, we are not able to recommend a specific instrument. Ideally, to capture the full impact of a health condition, both a generic and burn-specific instrument should be used[61]. However, a combination of instruments was only used by three studies and the burn-specific instruments are thus only available to a limited group of researchers. We recommend to achieve consensus on the best generic instrument(s), and a freely available, sensitive and reliable burn-specific instrument should be created. If both types of instruments are available; it is recommended to use both types in future studies to assess the full impact of burns on HRQL in children.

HRQL improves over time, however, parental concern remained an issue in the longer term. Parents keep express concerns regarding the recovery and development of their child years after the injury, stressing the importance of parent counselling and the inclusion of parents in the aftercare, regardless of the child's age. Attention should also be paid to appearance in the aftercare, as it was shown that children, regardless of their age, have problems with their change in appearance due to the burn injury.

Children with more severe burns, or with a facial or hand burn have a poorer HRQL. Having a more severe injury is also related to HRQL in burned adults and children with other injuries[18,62-64]. There is less evidence on the impact of facial and hand injuries in other populations[18,65-68]. The five out of six included studies that found a relation were all proxy reports. The only study that found no relation was based on children's self-report. It might be that parents are constantly reminded on the trauma by seeing the scar, which results in having more concerns. Patients do not see their scar constantly, which might have less influence on their HRQL. Other strong predictors were time since burn and comorbidity, indicating that HRQL improves over time and that children with comorbidities have a poorer HRQL. These factors are also found to influence HRQL in children with other injuries[69,70]. Age at time of burn was a less strong predictor. Young children might not remember the trauma and might adjust easier to scars while growing older and developing a body image[10,33]. The low attention for psychological predictors is remarkable as both in burned adults as in children with other injuries these factors are found to predict HRQL[18,32,71]. Children have an increased risk of post-traumatic stress after burns[32,72] and investigating psychological factors is worthwhile.

A strength of this review is the comprehensive overview of HRQL in burned children, based on six relevant databases and the identification of domains of HRQL and patient groups that need special attention in aftercare. A limitation is, however, that we only included studies written in English; we might have missed studies written in other languages. Another limitation is the large number of studies on the same datasets, the variation in study designs, instruments, and data presentation which impeded us to pool the data. As a consequence, the comparison of parent and children's scores was based on differences as reported by authors. Predictors were summarized based on their association and statistical significance. The sample size of a study was not taken into account, however, we have checked that our results were not conditioned on risk of bias or sample size (Table 2). Another limitation is the risk of bias of included studies which was in general moderate and varied widely. None of the studies had a low risk of bias on all domains. The most alarming was the high risk of bias on study attrition. Most studies failed to include a non-responder analyses and only few adequately report attempts to collect information on patients who dropped out. Also, the in general high or moderate risk of study confounding needs attention; only a few studies set hypotheses before conducting analyses and only a minority described to use available literature for their search for predictors. Besides, most

studies failed to report how missing data was handled and confounders were often not defined. In future articles it is important to include (description of) all of these factors in order to lower the risk of bias and improve the overall study quality.

CONCLUSION

This review shows that HRQL of burned children increases over time and it revealed domains and patient groups that require special attention in the aftermath of burns. Short- and long-term aftercare is indicated and should be available for children and their parents. Special attention is needed for children with major burns, with facial or hand burns, and for those with comorbidity. And parental concern and children's appearance are important topics within aftercare.

Besides, to improve our understanding of HRQL and further facilitate the comparability of data across studies, consensus on the most optimal generic instrument(s), on a widely available norm score and on data presentation is required. Moreover, a freely available, sensitive and reliable burn-specific instrument should be created, with a child-report version for children ≥5 years old.

Author contribution

Inge Spronk conceptualized and designed the study, collected data, analyzed and interpreted data, drafted the initial manuscript, and reviewed and revised the manuscript.
Nine Legemate conceptualized and designed the study, collected data, analyzed and interpreted data, and reviewed and revised the manuscript.
Dr Suzanne Polinder and Dr Margriet van Baar conceptualized and designed the study, analyzed and interpreted data, and reviewed and revised the manuscript.
All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Acknowledgements

The authors gratefully acknowledge Wichor Bramer (Biomedical information specialist, Medical Library, Erasmus MC) for performing the database search.

REFERENCES

1. Åkerlund E, Huss FR, Sjöberg F. Burns in Sweden: an analysis of 24538 cases during the period 1987-2004. Burns 2007; 33:31-6.
2. Dokter J, Vloemans A, Beerthuizen G, et al. Epidemiology and trends in severe burns in the Netherlands. Burns 2014; 40:1406-14.
3. Wasiak J, Spinks A, Ashby K, et al. The epidemiology of burn injuries in an Australian setting, 2000–2006. Burns 2009; 35:1124-32.
4. Peck MD. Epidemiology of burns throughout the world. Part I: Distribution and risk factors. Burns 2011; 37:1087-100.
5. Merrell S, Saffle J, Sullivan J, et al. Increased survival after major thermal injury: a nine year review. Am J Surg 1987; 154:623-7.
6. Tompkins RG, Remensnyder JP, Burke JF, et al. Significant reductions in mortality for children with burn injuries through the use of prompt eschar excision. Ann Surg 1988; 208:577.
7. Duke JM, Rea S, Boyd JH, et al. Mortality after burn injury in children: a 33-year population-based study. Pediatrics 2015; 135:e903-e10.
8. Bakker A, Maertens K, Van Son M, et al. Psychological consequences of pediatric burns from a child and family perspective: A review of the empirical literature. Clin Psychol Rev 2013; 33:361-71.
9. Van Baar M, Essink-Bot M-L, Oen I, et al. Functional outcome after burns: a review. Burns 2006; 32:1-9.
10. Sheridan RL, Hinson MI, Liang MH, et al. Long-term Outcome of Children Surviving Massive Burns. JAMA 2000; 283:69-73.
11. Ravens-Sieberer U, Erhart M, Wille N, et al. Generic health-related quality-of-life assessment in children and adolescents. Pharmacoeconomics 2006; 24:1199-220.
12. Guyatt GH, Jaeschke R, Feeny DH, et al. Measurements in clinical trials: choosing the right approach. Quality of life and pharmacoeconomics in clinical trials. p. 41-8.
13. Drotar D. Measuring health-related quality of life in children and adolescents: implications for research and practice: Psychology Press; 2014.
14. Coons SJ, Rao S, Keininger DL, et al. A comparative review of generic quality-of-life instruments. Pharmacoeconomics 2000; 17:13-35.
15. Varni JW, Burwinkle TM, Sherman SA, et al. Health-related quality of life of children and adolescents with cerebral palsy: hearing the voices of the children. Dev Med Child Neurol 2005; 47:592-7.
16. Yoder LH, Nayback AM, Gaylord K. The evolution and utility of the burn specific health scale: a systematic review. Burns 2010; 36:1143-56.
17. Stavrou D, Weissman O, Tessone A, et al. Health related quality of life in burn patients–a review of the literature. Burns 2014; 40:788-96.
18. Spronk I, Legemate C, Dokter J, et al. Predictors of health related quality of life after burn injuries: a systematic review. Crit Care 2018; 22:160.
19. Spronk I, Legemate C, Oen I, et al. Health related quality of life in adults after burn

injuries: a systematic review. PLoS One 2018; 13:e0197507.

20. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 6:e1000097.
21. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med 2013; 158:280-6.
22. Hedges L, Olkin I. Book Review: Meta Analysis: A Review and A New View. Educ Res 1986; 15:14-6.
23. Taylor R, Sander J, Taylor R, et al. Predictors of healthţrelated quality of life and costs in adults with epilepsy: A systematic review. Epilepsia 2011; 52:2168-80.
24. Kazis LE, Lee AF, Rose M, et al. Recovery curves for pediatric burn survivors advances in patient-oriented outcomes. JAMA Pediatr 2016; 170:534-42.
25. Landgraf JM, Vogel I, Oostenbrink R, et al. Parent-reported health outcomes in infants/toddlers: measurement properties and clinical validity of the ITQOL-SF47. Qual Life Res 2013; 22:635-46.
26. Palmieri TL, Nelson-Mooney K, Kagan RJ, et al. Impact of hand burns on health-related quality of life in children younger than 5 years. J Trauma Acute Care Surg 2012; 73:S197-204.
27. Spuijbroek AT, Oostenbrink R, Landgraf JM, et al. Health-related quality of life in preschool children in five health conditions. Qual Life Res 2011; 20:779-86.
28. Van Baar ME, Essink-Bot ML, Oen IMMH, et al. Reliability and validity of the Health Outcomes Burn Questionnaire for infants and children in The Netherlands. Burns 2006; 32:357-65.
29. Kazis L, Liang M, et al. The development, validation, and testing of a health outcomes burn questionnaire for infants and children 5 years of age and younger: American Burn Association/Shriners Hospitals for Children. J Burn Care Res 2002; 23:196-207.
30. Daltroy LH, Liang MH, Phillips CB, et al. American burn association/shriners hospitals for children burn outcomes questionnaire: Construction and psychometric properties. J Burn Care Rehabil 2000; 21:29-39.
31. Laitakari E, Koljonen V, et al. The long-term healthrelated quality of life in children treated for burns as infants 5-9 years earlier. Burns 2015; 41:1186-92.
32. Landolt MA, Buehlmann C, Maag T, et al. Brief report: Quality of life is impaired in pediatric burn survivors with posttraumatic stress disorder. J Pediatr Psychol 2009; 34:14-21.
33. Landolt MA, Grubenmann S, Meuli M. Family impact greatest: Predictors of quality of life and psychological adjustment in pediatric burn survivors. J Trauma Inj Infect Crit Care 2002; 53:1146-51.
34. Maskell J, Newcombe P, et al. Psychosocial functioning differences in pediatric burn survivors compared with healthy norms. J Burn Care Res 2013; 34:465-76.
35. Maskell J, Newcombe P, et al. Psychological and psychosocial functioning of children with burn scarring using cosmetic camouflage: A multi-centre prospective randomised controlled trial. Burns 2014; 40:135-49.
36. Rosenberg M, Celis MM, Meyer Iii W, et al. Effects of a hospital based wellness and exercise program on quality of life of children with severe burns. Burns 2013; 39:599-609.

37. Sveen J, Huss F, Sjöberg F, et al. Psychometric properties of the swedish version of the burn outcomes questionnaire for children aged 5 to 18 years. J Burn Care Res 2012; 33:e286-e94.
38. Van Baar ME, Polinder S, Essink-Bot ML, et al. Quality of life after burns in childhood (5-15 years): Children experience substantial problems. Burns 2011; 37:930-8.
39. Warner P, Stubbs TK, Kagan RJ, et al. The effects of facial burns on health outcomes in children aged 5 to 18 years. J Trauma Acute Care Surg 2012; 73:S189-96.
40. Meyer 3rd WJ, Lee AF, Kazis LE, et al. Adolescent survivors of burn injuries and their parents' perceptions of recovery outcomes: do they agree or disagree? J Trauma Acute Care Surg 2012; 73:S213-20.
41. Nicolosi JT, de Carvalho VF, Sabatés AL, et al. Assessment of health status of adolescent burn victims undergoing rehabilitation: a cross-sectional field study. Plast Surg Nurs 2013; 33:185-91.
42. Pan R, Egberts MR, Nascimento LC, et al. Health-Related Quality of Life in adolescent survivors of burns: Agreement on self-reported and mothers' and fathers' perspectives. Burns 2015; 41:1107-13.
43. Patrick DL, Topolski TD, Edwards TC, et al. Measuring the quality of life of youth with facial differences. Cleft Palate Craniofac J 2007; 44:538-47.
44. Pope SJ, Solomons WR, Done DJ, et al. Body image, mood and quality of life in young burn survivors. Burns 2007; 33:747-55.
45. Dodd AR, Nelson-Mooney K, Greenhalgh DG, et al. The effect of hand burns on quality of life in children. J Burn Care Res 2010; 31:414-22.
46. Kazis LE, Lee AF, Hinson M, et al. Methods for assessment of health outcomes in children with burn injury: the Multi-Center Benchmarking Study. J Trauma Acute Care Surg 2012; 73:S179-88.
47. Stubbs TK, James LE, Daugherty MB, et al. Psychosocial impact of childhood face burns: A multicenter, prospective, longitudinal study of 390 children and adolescents. Burns 2011; 37:387-94.
48. Sveen J, Sjöberg F, Öster C. Health-related quality of life in Swedish pediatric burn patients and associations with burn and family characteristics. Burns 2014; 40:987-94.
49. Weedon M, Potterton J. Socio-economic and clinical factors predictive of paediatric quality of life post burn. Burns 2011; 37:572-9.
50. Van Baar ME, Essink-Bot ML, Oen IMMH, et al. Reliability and validity of the dutch version of the American burn association/ Shriners hospital for children burn outcomes questionnaire (5-18 years of age). J Burn Care Res 2006; 27:790-802.
51. Arbuckle R, Abetz-Webb L. "Not just little adults": qualitative methods to support the development of pediatric patient-reported outcomes. Patient 2013; 6:143-59.
52. Matza LS, Swensen AR, Flood EM, et al. Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. Value Health 2004; 7:79-92.
53. Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life?:

An analysis of 8,591 children across age subgroups with the PedsQL™ 4.0 Generic Core Scales. Health Qual Life Outcomes 2007; 5:1.

54. Egberts MR, Van de Schoot R, Boekelaar A, et al. Child and adolescent internalizing and externalizing problems 12 months postburn: the potential role of preburn functioning, parental posttraumatic stress, and informant bias. Eur Child Adolesc Psychiatry 2016; 25:791-803.

55. Kassam-Adams N, Garcia-Espana JF, Miller VA, et al. Parent-child agreement regarding children's acute stress: The role of parent acute stress reactions. J Am Acad Child Adolesc Psychiatry 2006; 45:1485-93.

56. Kaugars AS, Silverman A, Kinservik M, et al. Families' perspectives on the effect of constipation and fecal incontinence on quality of life. J Pediatr Gastroenterol Nutr 2010; 51:747-52.

57. Eiser C, Morse R. Can parents rate their child's health-related quality of life? Results of a systematic review. Qual Life Res 2001; 10:347-57.

58. 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; 41:1405-19.

59. Griffiths C, Armstrong-James L, White P, et al. A systematic review of patient reported outcome measures (PROMs) used in child and adolescent burn research. Burns 2015; 41:212-24.

60. Tyack Z, Kimble R, McPhail S, et al. Psychometric properties of the Brisbane Burn Scar Impact Profile in adults with burn scars. PloS One 2017; 12:e0184452.

61. Van Beeck EF, Larsen CF, Lyons RA, et al. Guidelines for the conduction of follow-up studies measuring injury-related disability. J Trauma Acute Care Surg 2007; 62:534-50.

62. Polinder S, Meerding WJ, Toet H, et al. Prevalence and prognostic factors of disability after childhood injury. Pediatrics 2005; 116:e810-e7.

63. Sturms LM, van der Sluis CK, Groothoff JW, et al. The health-related quality of life of pediatric traffic victims. J Trauma Acute Care Surg 2002; 52:88-94.

64. Batailler P, Hours M, et al. Health status recovery at one year in children injured in a road accident: A cohort study. Accid Anal Prev 2014; 71:267-72.

65. Rumsey N, Clarke A, White P, et al. Altered body image: Appearance‐related concerns of people with visible disfigurement. J Adv Nurs 2004; 48:443-53.

66. Sainsbury DCG. Body image and facial burns. Adv Skin Wound Care 2009; 22:39-44.

67. McLean LM, Rogers V, et al. The patient–body relationship and the "lived experience" of a facial burn injury: a phenomenological inquiry of early psychosocial adjustment. J Multidiscip Healthc 2015; 8:377.

68. Hoogewerf C, van Baar M, et al. Impact of facial burns: relationship between depressive symptoms, self-esteem and scar severity. Gen Hosp Psychiatry 2014; 36:271-6.

69. DeMatteo CA, Hanna SE, Yousefi R, et al. Quality-of-life after brain injury in childhood: Time, not severity, is the significant factor. Brain Inj 2014; 28:114-21.

70. Winthrop AL, Brasel KJ, Stahovic L, et al. Quality of life and functional outcome after pediatric trauma. J Trauma Acute Care Surg 2005; 58:468-74.

71. Winthrop AL. Health-related quality of life after pediatric trauma. Curr Opin Pediatr 2010; 22:346-51.

72. Drake JE, Stoddard Jr FJ, Murphy JM, et al. Trauma severity influences acute stress in young burned children. J Burn Care Res 2006; 27:174-82.

PART II
CLINIMETRIC STUDIES ON OUTCOMES

5

Evaluation of measurement properties of health-related quality of life instruments in burn patients: a systematic review

C.M. Legemate, I. Spronk, L.B. Mokkink, S. Polinder, M.E. van Baar, C.H. van der Vlies

J Trauma Acute Care Surg. 2020 Apr;88(4):555-571

ABSTRACT

Background

Health-related quality of life (HRQL) is a key outcome in the evaluation of burn treatment. HRQL instruments with robust measurement properties are required to provide high quality evidence to improve patient care. The aim of this review was to critically appraise the measurement properties of HRQL instruments used in burns.

Methods

A systematic search was conducted in in EMBASE, MEDLINE, CINAHL, Cochrane, Web of Science and Google scholar to reveal articles on the development and/or validation of HRQL instruments in burns. Measurement properties were assessed using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) methodology. A modified GRADE analysis was used to assess risk of bias (Prospero ID: CRD42016048065).

Results

Forty-three articles covering 15 HRQL instruments (12 disease-specific and 3 generic instruments) were included. Methodological quality and evidence on measurement properties varied widely. None of the instruments provided enough evidence on their measurement properties to be highly recommended for routine use, however two instruments had somewhat more favorable measurement properties. The Burn Specific Health Scale Brief (BSHS-B) is easy to use, widely accessible and demonstrated sufficient evidence for most measurement properties. The Brisbane Burn Scar Impact Profiles (BBSIPs) were the only instruments with high quality evidence for content validity.

Conclusion

The BSHS-B (burn specific HRQL) and the BBSIP (burn scar HRQL) instruments have the best measurement properties. There is only weak evidence on the measurement properties of generic HRQL instruments in burn patients. Results of this study form important input to reach consensus on a universally used instrument to assess HRQL in burn patients.

INTRODUCTION

Due to substantial advances in surgical and critical care management, the number of people surviving burns has increased during the past few decades[1-3]. As a result, more patients have to deal with lifelong disabilities and disfigurements which are frequently a consequence of burn injury[5]. This has led to a shift in attention from clinician-led short-term outcomes, such as improvement of survival, to longer-term patient-centred outcomes of burn care focusing progressively on physical and psychological sequelae[5-7]. Therefore, perceived health-related quality of life (HRQL) of burn patients has become a key outcome in burn treatment[8,9].

Patient-reported outcome (PRO) measurement of HRQL offers an assessment of the patients' perspectives on burn care outcomes and is therefore useful in decision-making. Along with the variations in defining and operationalizing HRQL a variety of measurement instruments is currently available[9,10]. Measurement instruments to assess HRQL after burn injury are either generic (assessing general aspects of health) or disease specific (covering aspects that are specifically relevant for burn patients), with benefits and disadvantages to the use of either type. Generic instruments allow comparison with the general population and other diseases, whereas burn-specific instruments include disease specific items and may thus be better targeted to burn patients. Within burns a subtype of a burn specific instruments has been introduced: instruments that assess the influence of burn scarring on HRQL.

Selecting the best instrument to evaluate HRQL after burn injury requires the evaluation of specific instrument characteristics, feasibility of use (e.g. availability, patient compliance), and measurement properties. Measurement properties are quality aspects of a measurement instrument, such as reliability, validity, or responsiveness, and provide information whether the results obtained by an instrument can be trusted. HRQL instruments with robust measurement properties in burn patients are required to draw valid conclusions about HRQL outcomes and, ultimately, to provide high quality evidence to improve patient care. In this systematic review, the COnsensus-based Standards for the Selection of health Status Measurement INstruments (COSMIN) methodology and guidelines[11-13] are used to critically appraise the measurement properties of HRQL instruments used in burn patients.

METHODS

This review was conducted according to the COSMIN methodology and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Statement[11,14]. The protocol was registered a priori in the International Prospective Register of Systematic Reviews (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=48065).

Literature search

A systematic literature search (no date or language restriction) was conducted in EMBASE, MEDLINE, CINAHL, Cochrane, Web of Science and Google scholar on February 12, 2018. A medical librarian optimized the search strategy and performed the systematic search. The search strategy combined terms covering HRQL and the target population (patients with burn injury) (Appendix 1). A combined library of the retrieved articles was created using Endnote and duplications were excluded. The reference lists of included studies were hand searched for additional articles.

Article selection and data extraction

Articles were included if they met the following criteria: (1) written in English, (2) published as full-text papers in a peer-reviewed journal, and (3) their purpose was the development and/or evaluation of the measurement properties of instruments that measure the construct HRQL in burn patients. Relevant articles were selected on the basis of title by one researcher (IS). Two researchers (CL and IS) independently screened a random sample of ten percent of the abstracts. As there was no disagreement between the reviewers, one reviewer (IS) appraised the remaining abstracts. At a second stage, two reviewers (CL and IS) assessed all full-texts independently to identify studies evaluating measurement properties. Conflicts were resolved by consensus of the two reviewers and, if necessary, discussion with a third reviewer (MvB). Data on characteristics of included studies and instruments, and results on measurement properties were extracted independently by two reviewers (CL and IS) and cross-checked. Evidence tables were used to summarize data.

Assessment of methodological quality of included studies

Two researchers (CL and IS) independently scored all quality assessment steps described below. Any discrepancies were discussed and, if necessary, resolved with a third reviewer (MvB). The COSMIN taxonomy was used to select which measurement properties of an instrument were evaluated15 (Table 1). As there is no gold standard for HRQL, 'criterion validity' was not considered. Individual articles may comprise more than one study if they evaluate more than one measurement property or the same measurement property for more than one HRQL instrument The COSMIN Risk of Bias checklist was used to assess the methodological quality for each study11-13. Studies were stratified as having very good, adequate, doubtful, or inadequate methodological quality. More detailed information on the COSMIN Risk of Bias checklist and can be found elsewhere (http://www.cosmin.nl).

Table 1. Definitions and criteria for good measurement properties

Measurement property	Definition	Rating	Criteria
Reliability	The degree to which the measurement is free from measurement error		
Reliability (extended)	The extent to which scores for patients who have not changed are the same for repeated measurements under several conditions		
Internal consistency	The degree of the interrelatedness among the items	+	At least low evidence for sufficient structural validity AND Cronbach's alpha(s)≥0.70 for each unidimensional (sub)scale
		?	Criteria for "at least low evidence for sufficient structural validity" not met
		-	At least low evidence for sufficient structural validity AND Cronbach's alpha(s)≥0.70 for each unidimensional (sub)scale
Reliability	The proportion of the total variance in the measurements which is due to 'true' differences between patients	+	ICC or weighted Kappa≥0.70
		?	ICC or weighted Kappa not reported
		-	ICC or weighted Kappa<0.70
Measurement error	The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured	+	SDC or LoA <MIC
		?	MIC not defined
		-	SDC or LoA > IC
Validity	The degree to which a HRQL instrument measures the construct(s) it purports to measure		
Content validity	The degree to which the content of a HRQL instrument is an adequate reflection of the construct to be measured		
Relevance	The degree to which items in a HRQL instrument are relevant for the construct of interest within a specific population and context of use	+	≥85% of the items of the instrument fulfil the criterion*
		?	Not(enough) information available or quality of (part of a) the study inadequate
		-	<85% of the items of HRQL instrument fulfil the criterion
Comprehensiveness	The degree to which key aspects of the construct are missing		Idem relevance*
Comprehensibility	The degree to which terms are understood by patients as intended		Idem relevance*
Construct validity	The degree to which the scores of a HRQL instrument are consistent with hypotheses based on the assumption that the HRQL instrument validly measures the construct to be measured		

Table 1. Continued

Structural validity	The degree to which the scores of a HRQL instrument are an adequate reflection of the dimensionality of the construct to be measured	+	CTT: CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08. IRT/Rasch: No violation of unidimensionality: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08 *AND* no violation of local independence: residual correlations among the items after controlling for the dominant factor <0.20 OR Q3's <0.37 *AND* no violation of monotonicity: adequate looking graphs /item scalability >0.30 *AND* adequate model fit: IRT: -2 >0.01, Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z-standardized values > -2 and <2
		?	CTT: Not all information for '+' reported. IRT/Rasch: Model fit not reported
		-	Criteria for '+' not met
Hypotheses testing	Idem construct validity	+	At least 75% of the result is in accordance with the hypotheses[2] or no differences between groups reported[3]
		?	No correlations with instrument(s) measuring related construct(s) or no differences between groups reported[4]
		-	Criteria for '+' not met
Cross-cultural validity	The degree to which the performance of items of a translated/ culturally adapted HRQL instrument are an adequate reflection of the performance of items of the original version	+	No important differences found between group factors (such as age) in multiple group factor analysis OR no important DIF for group factors (McFadden's R2 < 0.02)
		?	No multiple group factor analysis/DIF analysis performed
		-	Important differences between group factors/DIF found
Criterion validity	The degree to which the scores of a HRQL instrument are an adequate reflection of a 'gold standard'	+	Correlation with gold standard ≥0.70 or AUC ≥ 0.70
		?	Not all information for '+' reported
		-	Correlation with gold standard <0.70 or AUC <0.70
Responsiveness	The ability of a HRQL instrument to detect change over time in the construct to be measured	+	The result is in accordance with the hypothesis / AUC ≥ 0.70
		?	No hypothesis defined (by the review team)
		-	Result is not in accordance with the hypothesis/ AUC <0.70

*Criteria on relevance, comprehensiveness and comprehensibility can be found on www.cosmin.nl. AUC = area under the curve, CFA = confirmatory factor analysis, CFI = comparative fit index, CTT = classical test theory, DIF = differential item functioning, ICC = intraclass correlation coefficient, IRT = item response theory, LoA = limits of agreement, MIC = minimal important change, RMSEA: Root Mean Square. Error of Approximation, SEM = Standard Error of Measurement, SDC = smallest detectable change, SRMR: Standardized Root Mean Residuals, TLI = Tucker-Lewis index."+" = sufficient, "–" = insufficient, "?" = indeterminate; 2. Correlations with instruments measuring the same construct >0.50 OR at least 75% of the results are in accordance with the hypotheses; 3. Known-groups were based on factors determining burn severity: Percentage Total Body Surface Area burned, Length of Stay, Surgery yes or no; 4. No hypotheses defined

Assessment of measurement property results

The result of each study on a measurement property was rated against criteria for good measurement properties: sufficient (+), insufficient (-), or indeterminate (?) (Table 1). Evidence on relevance, comprehensiveness and comprehensibility (aspects of content validity) was derived from development and content validity studies in which patients and/or professionals were involved. This was done first based on the methods and results of the instrument development study; second, based on each available content validity study of the specific instrument; and third, based on the reviewer's own rating of the content of the instrument (i.e. assessment of coverage of burn specific consequences, which was a subjective assessment of both reviewers on all items in each included HRQL instrument because no precedent exists)[16]. If instruments were not freely available, developers of the instrument were contacted. If they were not willing to distribute the instrument, the review team could not evaluate the content. Regarding hypothesis testing and responsiveness, we predefined that correlations with (domain scores of) other outcome measurements that aim to measure related constructs should be ≥0.30[17] and there should be significant differences in scores between relevant subgroups. Subgroups were based on the results of a previous systematic review on predictors of HRQL in burn patients and involved factors determining burn severity: percentage Total Body Surface Area burned (%TBSA), length of hospital stay and the necessity of surgery[18].

Synthesis of evidence and recommendations

All results per measurement property of each HRQL instrument were checked for consistency, and were qualitatively summarized. These summarized results were evaluated against the criteria for good measurement properties to produce an overall rating ((sufficient (+), insufficient (-), inconsistent (±), or indeterminate (?)) for each measurement property of each HRQL instrument. The focus was on the HRQL instrument specifically, while in the previous steps the focus was on the single studies. The GRADE approach (Grading of Recommendations, Assessment, Development, and Evaluation) was used to grade the quality of the evidence, determining the trustworthiness of the summarized results. For content validity, the evidence quality could be downgraded because of risk of bias (as determined using the COSMIN Risk of Bias checklist), inconsistency of results across studies and indirectness (i.e. evidence from different populations) (Appendix 5). For the other measurement properties, the evidence quality could be downgraded because of risk of bias, imprecision (i.e. low sample size), inconsistency, and indirectness (Appendix 5)[11]. To come to an evidence-based and transparent recommendation the instruments were categorized in three categories[11]. According to the COSMIN guidelines, instruments with sufficient content validity and sufficient internal consistency can be recommended for use (category A), PROMS can have the potential to be recommended for use (category B) and PROMS with high quality evidence for an insufficient measurement property should not be recommended for use (category C)[11]. The COSMIN guidelines indicate that if all instruments fall in category B, the most important property of a measurement instrument is content validity,

followed by structural validity and internal consistency. Subsequently, the results of the other measurement properties should be considered. In addition, information on feasibility was appraised to determine the feasibility of use, so recommendations would not only be based on the measurement properties. Important aspects of feasibility were defined as length of the instrument, completion time, and ease of score calculation and access fee of an instrument.

RESULTS

Of the 7246 records identified, 43 articles were considered eligible for assessment (Figure 1, Table 2). These 43 articles evaluated 15 different HRQL instruments. Most articles studied more than one measurement property; the included articles comprised 118 separate studies. Of the HRQL instruments identified, 3 were generic, and 12 were disease-specific. Of these 12 instruments, 4 instruments measured the impact of burn scarring on HRQL (Table 3). Six instruments were specifically developed for the use in children (1 generic, 5 disease specific, of which 3 on burn scarring). The most frequently appraised instruments were all burn-specific HRQL instruments: the Burns Specific Health Scale Brief (BSHS-B[17,19-37]), Burn Specific Health Scale Abbreviated (BSHS-A[38-42]), Burn Specific Health Scale Revised (BSHS-R[43-46]) (Table 2; Supplemental Digital Content). Of the instruments that were specifically for the use in children, the Burn Outcome Questionnaire 5-18 (BOQ 5-18y) was the most frequently appraised [47-49] (Table 2; Supplemental Digital Content).

General characteristics of the included articles and instruments are summarized in Table 2 and Table 3, respectively[17,19-60]. Table 3 also includes feasibility aspects of each HRQL instrument.

The most commonly assessed measurement properties were internal consistency, hypotheses testing, and reliability. No study assessed measurement error. Methodological quality and evidence on measurement properties were variable (Supplemental Digital Content).

Table 4 presents the results of the best evidence syntheses. All instruments were categorized as level B instruments: PROMs that have the potential to be recommended based on their measurement properties.

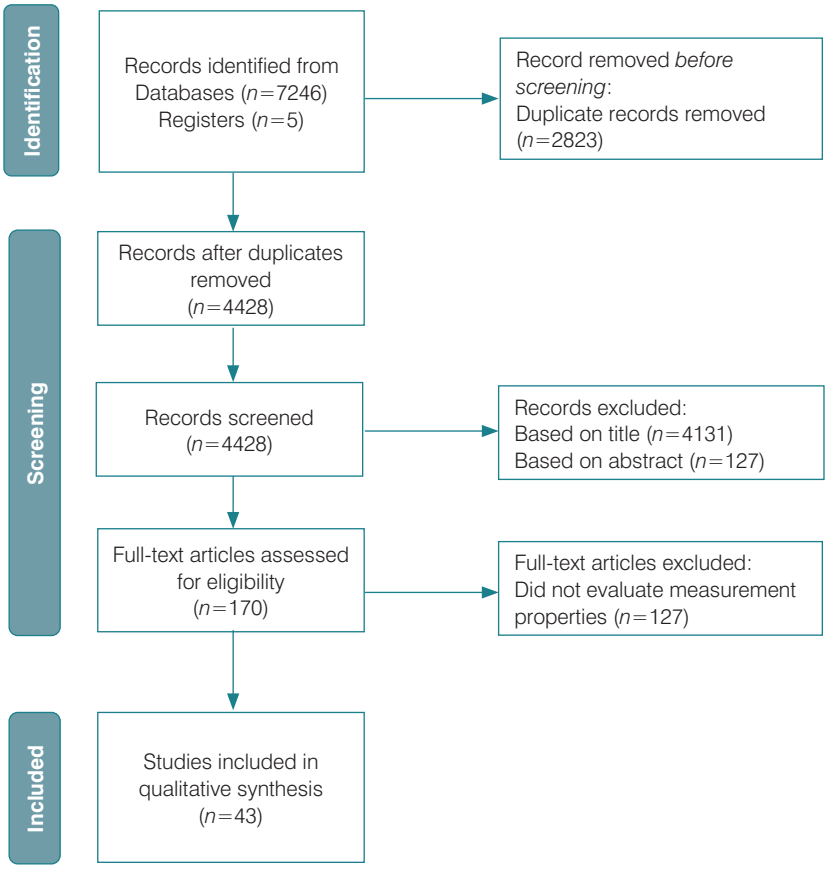


Figure 1. PRISMA flow diagram demonstrating the identification and screening of studies for inclusion

Table 2. Characteristics of the included articles

Author, year	Country (Language)	Measurement properties[1]	Study population[2]	Adults/ children	Time post burn	%TBSA M (SD)
BBSIP (adults)						
Tyack 2015	Australia (English)	CV	n=10, M:60%.Age: mean:49y	≥8y	Median 12.5 months	14 (20)
Tyack 2017	Australia (English)	IC, R, HT RE	n=118, M:74%. Age: median 34y	Adults (>18y)	Baseline, wound healing, 1-2 weeks, 1-month	Median 4
BBSIP (care-giver 0-8y)						
Tyack 2015	Australia (English)	CV	n=9, M:44%. Age: mean: 2.5y	Children (0-8y)	Mean 8 months	8 (14)
BBSIP (care-giver 8-18y)						
Tyack 2015	Australia (English)	CV	n=11, M:64%. Age: mean:13y	Children (8-18y)	Median 10 months	7.5 (18)
BBSIP (child 8-18y)						
Tyack 2015	Australia (English)	CV	n=11, M:64%. Age: mean:13y	Children (8-18y)	Median 10 months	7.5 (18)
BOQ 0-4y						
Kazis 2002	USA (English)	SV, IC, R, HT	n=184, M:54%. Age: mean: 2.6y	Children (0-5y)	Baseline, 6 months	17
van Baar 2006	Netherlands (Dutch)	CV, IC, R, HT	n=194, M:54%. Age:37months	Children (0-4y)	Mean 17.5 months	6 (0-66)
BOQ 5-18y						
Daltroy 2000	USA (English)	IC, HT	n=86, M:71%. Age: mean:10y	Children (5-18y)	NR	22 (NR)
van Baar 2006	Netherlands (Dutch)	CV, IC, R, HT	n=145, M:65%. Age: mean:8.8y	Children (5-15y)	Mean 21.1 months	6.0 (2.0)
Sveen 2012	Sweden (Swedish)	CV, IC, R, HT	n=70, M:67. Age: mean:9.5y	Children (5-18y)	Mean 5.4 years	10.5 (12.7)
BSHS						
Blades 1982	USA (English)	IC	n=40, M:NR.Age: mean:32.1y	Adults	35 weeks	30 (11-80)
BSHS-A						
Adam 2009	Turkey (Turkish)	IC, R, HT	n=53, M:81%.Age: mean:33.7y	≥ 16 y	2 weeks	19.9 (12.5)
Li 2014	China (Chinese)	CV, IC, R, HT	n=457, M:70%.Age: mean:36.7y	Adults (≥18y)	13.6 months	39.6 (27.3)
Moi 2003	Norway (Norwegian)	CV, IC, R, HT	n=95, M:82%.Age: mean:43.7y	Adults (≥18y)	Mean 47.0 months	18.5 (14.2)
Munster 1987	USA (English)	IC, R	n=70, M:NR.Age: mean: NR	Adults	NR	NR
Salvador Sanz 1998	Spain (Spanish)	CV, IC, R, HT	n=115, M:54%.Age: mean:40.5y	Adults (16-73y)	Mean 783 days	14 (11.6)
BSHS-B						
Finlay 2014	Australia (English)	SV, IC, HT	n=224, M:83%.Age: mean:36y	≥ 16y	1 and ≥6 months	4 (1-60)
Gandolfi 2016	France (French)	IC, R, HT	n=53, M:66%.Age: mean:46.4y	Adults (18-70y)	Range 2-4 years	26.9 (15.9)
Goudarzian 2017	Iran (Persian)	SV, IC	n=410, M:100%.Age: mean:39.4y	Pregnant women	NR	19.1 (2.16)
Hwang 2016	Taiwan (Chinese Taiwanese)	IC, R, HT	n=108, M:64%.Age: mean:42.1y	Adults	Mean 565 days	23.3 (25.4)

Kildal 2001	Sweden (Swedish)	SV, IC, HT	n=248, M:80%.Age: mean:36.8y	Adults (≥18y)	Mean 9,3 years	23.1 (16.2)
Ling-Juan 2012	China (Chinese)	CV, SV, IC	n=208, M:77%. Age: mean:40.4y	Adults (≥18y)	Mean 37.12 months	40.1 (27.4)
Meirte 2012	Netherlands and Belgium (Dutch)	HT	n=184, M:71%. Age: mean:39.0y	Adults	9 months	11.8 (10.2)
Mulay 2015	India (Hindi)	CV, SV, IC, R	n=20, M:40%. Age: mean:31.0y	Adults (18-65y)	6 to 12 months	39.8(20-60)
Muller 2015	Germany (German)	SV, IC, HT	n=141, M:65%. Age: mean:49.6y	Adults (≥18y)	Mean 45.01 months	12.9 (10.3)
Piccolo 2015	Brazil (Brazilian Portuguese)	CV, IC, R, HT	n=92, M:52%. Age: mean:37.1y	Adults (≥18y)	Mean 4.38 years (range 1-30 years)	19.1 (18.8)
Pishnamazi 2013	Iran (Persian)	CV, SV, IC, R, HT	n=200, M:38%. Age: mean:25y	Adults (≥18y)	NR	34.9 (2.0)
Sideli 2010	Italy (Italian)	IC, HT	n=50, M:72%. Age: mean:40.1y	Adults	One month	31.8 (17.6)
Sideli 2014	Italy (Italian)	IC, HT	n=131, M:53%. Age: mean:40.2y	Adults (18-65y)	<6 months	16.8 (12.2)
Stampolidis 2012	Greece (Greek)	IC, R, HT	n=40, M:74%. Age: mean:52.2y	Adults	During admission	15.6 (13.0)
Stavrou 2015	Israel (Hebrew)	CV, IC, R, HT	n=86, M:79%. Age: mean:38.0y	Adults(≥18y)	12.7 (13.3) months	11.1 (11.6)
Stolle 2017	Germany (German)	SV, IC, R, HT	n=364, M:69%. Age: mean:44.7y	Adults(≥18y)	Range 1-50 year	9.5 (12.2)
Szczechowicz 2014	Poland (Polish)	CV, IC, R, HT	n=190, M:66%. Age: mean:46.5y	Adults(≥17y)	NR	23.6 (15.7)
Van Loey 2013	Swedish (S) and Dutch (D)	CCV	S:n=231, M:73%. Age: mean:45.6y D:n=275, M:74%. Age: mean:39.3y	S: 16-93y D: 18-88y	S: 9 months D: 12 months	S:21.6 (18) D:11.8 (11)
Willebrand 2008	Sweden (Swedish)	SV, IC	n=334, M:78%. Age: mean:46.4y	Adults (≥18y)	Mean 7.9 (4.8) year	21.6 (16.0)
Willebrand 2011	Sweden (Swedish)	SV, IC, HT	n=94, M:73%. Age: mean:44.4y	Adults (≥18y)	6, 12, 24 months	23.4 (19.6)
BSHS-R						
Blalock 1992	USA (English)	CV	n=38, M:82%. Age: mean:43.2y	Adults (≥18y)	Mean 362 days	26.5 (14.4)
Blalock 1994	USA (English)	SV, IC, HT	n=254, M:74%. Age: mean:39.3y	Adults	Mean 313 days	19.2 (15.1)
Ferreira 2008	Brazil (Brazilian Portuguese)	SV, IC, HT	n=115, M:66%. Age: mean:31.8y	Adults (≥18y)	NR	19.3 (16.0)
Nicolosi 2013	Brazil (Brazilian Portuguese)	IC	n=63, M:30%. Age: mean:16.0y	12-20y	NR	23.8
DLQI						
Finlay 1994	UK (English)	CV	n=120, M:42%. Age: median:42y	Adults (15-17y)	NA	NA
Mazharinia 2007	Iran (Persian)	SV, IC	n=109, M:44%. Age: mean:28.9y	Adults (≥16y)	NR	NR
EQ-5D						
Oster 2009	Sweden (Swedish)	HT	n=78, M:78%. Age: mean:43.6y	Adults (≥18y)	Baseline, 3, 6, 12 months	24.3 (19.7)
Meirte 2017	Netherlands and Belgium (Dutch)	HT	n=184, M:71%. Age: mean:39.0y	Adults	9 months	11.8 (10.2)
ITQOL-SF47						
Landgraf 2013	Netherlands (Dutch)	SV, IC	n=194, M:54%. Age: 37 months	Children	Mean 17.5 months	6 (0-66)

Table 2. Continued

SF-36						
Edgar 2010	Australia (English)	HT	n=280, M:81%. Age: mean:37.4y	≥16y	1, 3, 6, 12 months	8.9 (11)
Meirte 2017	Netherlands and Belgium (Dutch)	HT	n=184, M:71%. Age: mean:39.0y	Adults	9 months	11.8 (10.2)
YABOQ						
Ryan 2013	USA (English)	SV, IC, R	n=153, M:73%. Age: mean:24.7y	Adults (19-30y)	Baseline, 2 weeks, 6, 12 months	11 (14)

[1]CV=Content validity; CCV=Cross-cultural validity; HT=Hypothesis testing; IC=Internal consistency; R=Reliability; RE=Responsiveness; SV=Structural validity. [2]n=sample size; M=males. BBSIP=Brisbane Burn Scar Impact Profile, BDI=Beck Depression Inventory, BOQ = Burn Outcome Questionnaire, BSHS = Burn Specific Health Scale, BSHS-A = Burn Specific Health Scale-Abbreviated, BSHS-B = Burn Specific Health Scale-Brief, BSHS-R=Burn Specific Health Scale-Revised, *EQ-5D* = EuroQol 5 dimensions, ITQOL=Infant Toddler Quality of Life Questionnaire, NA=Not applicable, NR=Not reported, SF-36=Short Form 36, YABOQ =Young Adult Burn Outcome Questionnaire

Measurement properties of generic HRQL instruments

The three generic HRQL measurement instruments include the EuroQol 5 dimensions (EQ-5D), the 47-item short form Infant Toddler Quality of Life Questionnaire (ITQOL-SF47), and the 36-item Short Form survey (SF-36). There was only weak evidence on the measurement properties of generic HRQL instruments in burns. The comprehensiveness of all of these instruments was rated insufficient because these instruments did not cover all the aspects of HRQL that are relevant to patients with burn injury (e.g. problems related to scarring). There was high quality evidence for sufficient hypotheses testing for construct validity of the EQ-5D and SF-36, but studies on other measurement properties in burns were lacking[17,55,59]. Both scales are widely available and especially the SF-36 is widely applied within the field of burns[9]. In terms of feasibility, a limitation of the SF-36 is the license fee. Structural validity and internal consistency of the ITQOL-SF47 were studied but both rated as indeterminate[58].

Measurement properties of burn-specific HRQL instruments

The twelve disease-specific HRQL instruments were: The Brisbane Burn Scar Impact Profile (BSSIP) for adults, BBSIP for children 8-18 years, BBSIP for caregivers of children <8 years, BBSIP for caregivers of children aged 8-18 years, BOQ 0-4, BOQ 5-18 years, Burn Specific Health Scale (BSHS), BSHS-A, BSHS-B, BSHS-R, Dermatology Life Quality index (DLQI), and the Young Adult Burn outcome Questionnaire (YABOQ) (Table 3). The different versions of the BSSIP focus on the impact of burn scarring on HRQL and are the only instruments with moderate to high quality evidence for sufficient content validity, which is the most important measurement property according to the COSMIN guideline. The BSHS-B is the only instrument with high quality evidence for internal consistency, which is (together with structural validity) the second important measurement property according to the COSMIN guideline. Therefore, these instruments will be discussed in more detail. Regarding the other instruments, it is of note that these are not necessarily inadequate but that their measurement properties are merely not or scarcely investigated in literature.

BBSIP

The BBSIP was developed in 2013 to assess burn scar specific HRQL in burn patients at risk of, or with burn scars[51]. Multiple versions were developed for different age groups (Table 3). International scar management experts and patients were involved in the development of the items and cognitive interviews were done to understand how patients interpreted the items[51]. Nevertheless, the overall rating of comprehensiveness was judged to be doubtful for all versions because patients were not asked about the comprehensiveness of the final developed forms. Other content validity studies were not encountered.

Table 3a. Characteristics of the Health-Related Quality of Life instruments

Name	No items	Adults/ children	HRQL construct definition	Subscales	Number/type of options	Scoring algorithm	Feasibility	Language version	Administration costs
Generic instruments									
EQ-5D	5	Both	Health-related quality of life	5 subscales: Mobility, self-care, usual activities, pain/discomfort and anxiety/depression and a visual analogue scale (VAS) for general health	Three response levels	3 levels per dimension, a VAS score from 0-100, and index 'utility' score	Few minutes	176 translations*	Dependents on project, funding, etc.
ITQOL-SF47	47	Children 0-5 years	'health': a complete physical, mental and social wellbeing	2 subscales comprising 13 concepts: Infant (38 items); physical abilities, growth and development, bodily pain, temperament and moods, behaviour. Parent (9 items); Emotional impact, Impact time, family cohesion	5-point Likert scale	Sum score from 47 items and transformation to a scale from 0 (worst health) - 100 (best health)	10 min	50 languages****	Varies according to use
SF-36	36	Adults	'health', eight concepts	8 subscales: Physical functioning, role limitations physical, bodily pain, general health, vitality, social functioning, role limitations emotional, and mental health	3-point, 5 point, 6 point Likert scale	Transformed mean domain scores (0 [worst] to 100 [best]).	5-10 min	>170 translations	Licensing fee dependents on type of organization
Disease-specific instruments									
BBSIP (adults)	66	Adults	Impact of scarring on a person's life experience	7 subscales: Overall impact of scars; impact of itch, pain and other sensations; work and daily activities (mobility and daily activities items); friendship and social interaction; appearance; emotional reactions; and physical symptoms	7-point Likert scale Dichotomous / Numeric	The total score is the summed score of individual items divided by the number of applicable items.	ND	English	Free

BBSIP (caregivers of children 0-8y)	58	Children	Impact of scarring on a person's life experience	8 subscales: Overall impact of scars; impact of itch, pain and other sensations; school, play and daily activities (mobility and daily activities items); friendships and social interaction; appearance; emotional reactions; physical symptoms; and parent and family concerns	5 point Likert scale Dichotomous / Numeric	The total score is the summed score of individual items divided by the number of applicable items.	ND	English	Free
BBSIP (caregivers of children 8-18y)	62	Children	Impact of scarring on a person's life experience	8 subscales: Overall impact of scars; impact of itch, pain and other sensations; school, play and daily activities (mobility and daily activities items); friendships and social interaction; appearance; emotional reactions; physical symptoms; and parent and family concerns	5-point Likert scale Dichotomous / Numeric	The total score is the summed score of individual items divided by the number of applicable items.	ND	English	Free
BBSIP (children 8-18y)	58	Children	Impact of scarring on a person's life experience	7 subscales: Overall impact of scars and treatment; impact of itch, pain and other sensations; daily activities; friendship and social interaction; appearance; emotional reactions; physical symptoms	5-point Likert scale Dichotomous / Numeric	The total score: summed score of items divided by the number of applicable items.	ND	English	Free

Table 3b. Characteristics of the Health-Related Quality of Life instruments

Name	No items	Adults/ children	HRQL construct definition	Subscales	Number/type of options	Scoring algorithm	Feasibility	Language version	Administration costs
BOQ <5y	55	Children <5y	'Health status'; no definition given	10 subscales: Play, language, fine motor, gross motor, behaviour, family, pain/itching, appearance, satisfaction, concern/worry	3 point Likert scale / 5 point Likert scale	Domain scores (0 [worst] – 100 [best])	16 min	English, Dutch	ND
BOQ 5-18y	53	Children 5-18y	Function, physical appearance and 'other relevant outcomes'	12 subscales: Upper extremity function, physical function and sports, transfers and mobility, pain, itch, appearance, compliance, satisfaction with current state, emotional health, family disruption, parental concern, school re-entry	Numeric/ Likert scales/ Dichotomous	Domain scores (0 [worst] – 100 [best])	Parents: 30 min; Child: 45 min	English, SwedishDutch	ND
BSHS	114	Adults	Dysfunction and distress/Health related quality of life	6 subscales: Physical health, body image, psychological health, sexual health, physical activities and family/social relationships	ND	ND	ND	English, Spanish	ND
BSHS-A	80	Adults	Health related quality of life; no definition given	8 subscales: Mobility and Self-Care, Family/Friends, Body Image, Affective, Hand Function, Sexual Activity, Role Activities, General Functioning	Ordinal score (0-4)	Dividing the total score for a domain by the total possible score: range from 0.00 (best) to 1.00 (worst).	31 min	English, Chinese, Norwegian, Turkish	Free
BSHS-B	40	Adults	Health related quality of life; no definition given	9 subscales: Simple Abilities, Interpersonal Relationships, Body Image, Affect, Hand Function, Sexuality, Heat Sensitivity, Treatment Regimens, Work	Ordinal score (0-4)	Mean scores per domain. Higher scores reflect a higher perceived functioning.	10-15 min	14 languages	Free

BSHS-R	31	Adults	The impact of burn injury	6 subscales: Simple Functional Abilities, Interpersonal Relationships, Body Image/ Affect, Heat Sensitivity, Treatment Regimens, Work	Ordinal score (1-5)	Mean scores per domain and sum score (31 [worst] - 155 [best])	ND	English, Brazilian, Portuguese	Free
DLQI[§]	10	Adults	Quality of life: no definition given	6 subscales: Symptoms and feelings, daily activities, leisure, work and school, personal relationships, treatment	4-point Likert scale	Sum score (range 30 [best] – 0 [worst])	2 min	115 languages***	Free (if not external funded).
YABOQ[4]	47	Adults <30y	Long-term burn recovery	14 subscales: Physical Function, Fine Motor Function, Pain, Itch, Social Function Limited by Physical Function, Perceived Appearance, Social Function Limited by Appearance, Sexual Function, Emotion, Family Function, Family Concern, Satisfaction with Symptom Relief, Satisfaction with Role, Work Reintegration, and Religion.	Likert scales / numeric / dichotomous	For each domain, scores are standardized to a mean of 50 (reference group), a higher score denotes a better health.	10 min	English	Free

ND = Not Determined, [§]Disease-specific for all patients with a skin disease *www.euroqol.org; **number of items and items modified from original version; ***www.cardiff.ac.uk/dermatology/quality-of-life/dermatology-quality-of-life-index-dlqi/ ; ****https://www.healthactchq.com/translation/itqol

BBSIP adult version

The adult versions of the BBSIP consists of 66 items divided into 7 subscales (Table 3). One study reported that Cronbach's alpha was ≥0.7 for all subscales[50], but the quality of the study was rated doubtful and the overall rating of internal consistency was indeterminate because there were no studies on structural validity (Table 4; Supplemental Digital Content). Reliability and hypotheses testing for construct validity were sufficient, however, the evidence was graded as moderate as a consequence of downgrading for risk of bias (i.e. only one study of adequate quality was available). One study provided high quality evidence for sufficient responsiveness.

BBSIP child versions

The version of the BBSIP for children aged 8-18y consists of 58 items divided into 7 subscales. The BBSIP for caregivers of children <8 years and the BBSIP for caregivers of children 8-18 years both comprise an extra subscale to measure 'parent and family concerns' and consist of 58 and of 62 items, respectively[51]. No studies on other measurement properties of the child or caregiver versions of the BBSIP were revealed in our systematic search. Regarding the feasibility of the different versions of the BBSIP, currently, all versions are only available in English, but validated translation studies may emerge in future. In order to reach the level A status, it is vital that structural validity is assessed to determine if the item on the scales sufficiently measure the same construct.

BSHS-B

The 40-item BSHS-B was derived from items of the BSHS and Revised BSHS (BSHS-R) in 2001[23,43,61]. Despite a development process that involved patients and featured a pilot study, comprehensibility was the only aspect of content validity that was rated sufficient (Table 4).

Relevance and comprehensiveness of the BSHS-B were rated inconsistent as a result of conflicting results of multiple studies (Supplemental Digital Content)[19-35,37]. The BSHS-B consists of 9 subscales that have been confirmed in one study that used confirmatory factor analysis with an adequate sample size, which was therefore of very good quality[35]. Nevertheless, some studies that were of inferior quality because they used exploratory factor analysis and/or had an inadequate sample size showed other results and therefore the overall quality of structural validity was graded moderate[21,26]. The BSHS-B carries high quality evidence for sufficient internal consistency, reliability, and very low-quality evidence for sufficient cross-cultural validity. Furthermore, moderate quality evidence for sufficient hypotheses testing for construct validity was found.

Table 4. Evidence synthesis (rating[§] and quality[¥] of the evidence) on measurement properties of HRQL after burn injury

	Content validity			Internal structure			Construct validity			
	Relevance	**Comprehensiveness**	**Comprehensibility**	**Structural validity**	**Internal consistency**	**Reliability**	**Hypotheses testing**	**Cross-cultural validity**	**Responsiveness**	**Category****
Generic instruments										
EQ-5D	+ Very low	- Very low	+ Very low				+ High			B
SF-36	+ Very low	- Very low	+ Very low				+ High			B
ITQOLSF-47*	+ Very low	- Very low	+ Very low	? Moderate	? High					B
Disease specific instruments										
BBSIP (adults)	+ High	+ Moderate	+ High		? Low	+ Moderate	+ Moderate		+ High	B
BBSIP (caregivers 0-8y)*	+ High	+ Moderate	+ High							B
BBSIP (caregivers 8-18y)*	+ High	+ Moderate	+ High							B
BBSIP (children 8-18y)*	+ High	+ Moderate	+ High							B
BOQ 0-4*	± Moderate	+ Very low	+ Very low	? Moderate	? Moderate	+ Moderate	- Moderate			B
BOQ 5-18*	+ Moderate	± Low	± Moderate		? High	± Moderate	- Low			B
BSHS					? Very low					B
BSHS-A	+ Very low	+ Very low	± Low		? Low	+ High	+ High			B
BSHS-B	± Moderate	± Low	+ Moderate	+ Moderate	+ High	+ High	+ Moderate	+ Very low		B

Table 4. Continued

BSHS-R	+	-	+	?	?		+	B
	Very low	Very low	Very low	Moderate	High		Moderate	
DLQI**	±	+	±	?	?			B
	Very low	Very low	Very low	Moderate	Low			
YABOQ	±	+	+	?	?	?		B
	Very low	Very low	Very low	Very low	Low	Very low		

*Developed for the use in children with burns

§ Results were qualitatively summarized in an overall conclusion that was either sufficient (+), insufficient (-), inconsistent (±), or indeterminate (?).

¥ The quality of the evidence contributing to rating of results was graded according to the modified Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach adapted for this type of review into: ''high,'' ''moderate,'' ''low,'' or ''very low''.

** A: PROMs that have the potential to be recommended as the most suitable PROM for the construct and population of interest (HRQL instruments with evidence for sufficient content validity (any level) and at least low quality for sufficient internal consistency). B: PROMs that may have the potential to be recommended, but further validation studies are needed (HRQL instruments categorized not in A or C). C: PROMs that should not be recommended (HRQL instruments with high quality evidence for an insufficient measurement property)[11].

"+" = sufficient, "–" = insufficient, "±" = moderate, "?" = indeterminate

The BSHS-B carries the best evidence for sufficient measurement properties (Table 4). It has been studied extensively (Table 1), resulting in good quality evidence for sufficient structural validity, internal consistency and cross-cultural validity. The instrument is relatively short, and freely available in 14 languages. Nevertheless, there is only low to moderate evidence on sufficient content validity (which is the most important measurement property according to the COSMIN guidelines). Of note is that especially relevance and comprehensives of the BSHS-B should therefore be investigated further.

DISCUSSION

This systematic review provides a comprehensive overview of all available studies on measurement properties of instruments used to assess HRQL in burn patients. Recently updated, consensus-based standards, developed by the COMSIN initiative[11,12,15], were used to ascertain sufficient quality of this review. This review comprised 118 different studies on the measurement properties of 15 different instruments. The methodological quality of the studies varied widely. Most of the measurement properties reported in the studies were rated sufficient; only 11 (9%) were rated insufficient (Supplemental Digital Content), which might indicate publication bias as positive results are more likely to be published.

According to the COSMIN guidelines, PROMs with evidence for sufficient content validity and at least low quality evidence for sufficient internal consistency can be recommended for use and results obtained with these PROMS can be trusted[11]. None of the instruments provided enough evidence on their measurement properties to be highly recommended for routine use. All instruments were categorized as level B instruments: PROMs that have the potential to be recommended based on their measurement properties. Further validation studies are needed before one instrument can be highly recommended. Though, two instruments (the BSHS-B and the different versions of the BBSIP) currently have favorable measurement properties compared to the rest.

The BSHS-B was studied most and possessed the strongest evidence for sufficient quality of most of the measurement properties assessed. Moreover, it seemed the most feasible instrument as is relatively short and freely available in 14 languages. However, the analysis of content validity showed that adding items or item refinement seems necessary before the BSHS-B can be highly recommended. Inconsistency in the results of content validity studies made it difficult to define the 'true' gaps in the content of its items. Further validation of the content should therefore be obtained by systematically asking patients and professionals (e.g. clinicians, researchers) about the relevance and comprehensiveness of the items. Also, data on measurement error of the BSHS-B is lacking and should be investigated to determine if the measurement errors are small

enough to obtain important differences in change scores and to determine the importance of (change) scores in an individual.

The four versions of the BBSIP were more recently developed than the other HRQL instruments. Hereby, the developers of these instruments were the only one able to use modern-state-of-the-art methods to develop the instruments (Supplemental Digital Content)[51,62]. This may have contributed to the fact that these were the only instruments that met the high standards for high-quality PROM development and content validity. It is of note to mention that the BBSIP versions were developed to measure HRQL for people at risk of or with burn scars; all questions are asked in relation to scarring whilst domains like work and daily activities or emotional reactions may be also influenced by other trauma-related factors and not all patients may only suffer from scarring[16,50,51]. The BBSIP versions have to be translated and validated further before they can be highly recommended based on their measurement properties. The outcomes of the questionnaires are the sum score of all items divided by the number of completed items. Future studies should preferably focus on structural validity to determine if this method allows for a meaningful interpretation of scores and to identify whether or not treatment effects are influenced by some scales or items and not others.

All other instruments showed moderate to very low-quality evidence for the aspects of content validity. This was likely the result of poorly performed development studies (no patient involvement or insufficiently sized qualitative interview groups) and a general paucity of studies that analyzed the content of the instruments. Regarding the other measurement properties of the other instruments, it is of note that these are not necessarily inadequate but they are merely not or scarcely investigated in literature.

The generic instruments EQ-5D and SF-36 are helpful for making a comparison with population norms and other patient groups[9]. Both instruments score moderate to high quality evidence for sufficient hypotheses testing which suggests that these instruments can adequately determine differences between groups that differ in burn severity[17,57,59]. However, they seem to miss important content that is relevant for patients after burns; items related to scarring (self-esteem, stigmatization, physical appearance) are missing. Therefore, it cannot be assured that the patient's perspective on HRQL is comprehensively captured in the outcomes.

Burn injury comprises a wide range of patients with mild to severe injury and can affect all domains of physical, psychological and social functioning[16,63]. Unfortunately, there is no consensus on what items should be included in an instrument to measure HRQL after burn injury[16]. Apart from further studies on the measurement properties of the identified instruments, there is a need to reach consensus on the definition of HRQL for burn patients, as well as on the best instrument to measure HRQL. In a broader perspective, it would be valuable to come to worldwide consensus

on a core outcome set (COS) (agreed minimum set of outcomes that should be measured) that should be measured in burn patients. Recently, the development of a COS for clinical trials in burns has been initiated by Young et al., proposing HRQL as one of the outcomes[64]. The combination of the COSMIN Risk of Bias checklist and criteria for good measurement properties to form a summary of the evidence base for each PROM is crucial to determine which outcome measurement instruments should be included in a COS. Results of current review can therefore guide these recommendations[65].

Limitations

The COSMIN risk of bias checklist and criteria for good measurement properties are strict, require high standards for reporting and call for distinct reporting of results. Some of the studies may be of higher quality than rated in this review as a result of incomplete reporting, even though researchers may perform extensive studies. In addition to the quality of measurement instruments, the specific construct as measured by the measurement instrument, feasibility and interpretability are important aspects when selecting the most suitable measurement instruments. In Table 3 we described the completion time and aspects of feasibility but the assessment of interpretability (e.g. floor- and ceiling effects, minimal important changes) went beyond the scope of this review. Current review focussed on instruments that aimed to measure HRQL. As a consequence, other PROMS that may assess only specific aspects of HRQL have not been included. For example, the LIBRE (Life Impact Burn Recovery Evaluation) profile that aims to measure social participation includes items on social role and personal relationships, which may also be important to measure HRQL[66].

CONCLUSION

This is the first systematic review to critically appraise instruments that measure HRQL after burn injury taking into account the quality of the studies in which the measurement properties were assessed using internationally accepted standards. It showed that the BSHS-B (burn specific HRQL) and the BBSIP (burn scar HRQL) instruments have the best measurement properties and that there is only weak evidence on the measurement properties of generic HRQL instruments in burns. This systematic review provides guidance on the HRQL instrument with the best measurement properties. There is a need for consensus on what specific symptoms or aspects are relevant and need to be included in an instrument to comprehensively assess HRQL after burn injury. The overview provided in this review forms important input to reach consensus on a universally used instrument to assess HRQL in burns. In time, this will ultimately provide high quality evidence to improve patient-centred care.

Acknowledgements

The authors gratefully acknowledge Wichor Bramer (Biomedical information specialist, Medical Library, Erasmus MC) for performing the database search.

Disclosure

For all authors, no conflicts are declared. This study was funded by the Dutch Burn Foundation (grant numbers 15.101 and 15.102). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the article.

Supplemental Digital Content

Available online: https://links.lww.com/TA/B547

REFERENCES

1. Bloemsma GC, Dokter J, Boxma H, et al. Mortality and causes of death in a burn centre. Burns 2008; 34:1103-7.
2. Brusselaers N, Hoste EA, Monstrey S, et al. Outcome and changes over time in survival following severe burns from 1985 to 2004. Intensive Care Med 2005; 31:1648-53.
3. Tompkins RG. Survival from burns in the new millennium: 70 years' experience from a single institution. Ann Surg 2015; 261:263-8.
4. Ryan CM, Schneider JC, Kazis LE, et al. Benchmarks for multidimensional recovery after burn injury in young adults: the development, validation, and testing of the American Burn Association/Shriners Hospitals for Children young adult burn outcome questionnaire. J Burn Care Res 2013; 34:e121-42.
5. Van Baar M, Essink-Bot M-L, Oen I, et al. Functional outcome after burns: a review. Burns 2006; 32:1-9.
6. Pereira C, Murphy K, Herndon D. Outcome measures in burn care: is mortality dead? Burns 2004; 30:761-71.
7. Klein MB, Goverman J, Hayden DL, et al. Benchmarking outcomes in the critically injured burn patient. Ann Surg 2014; 259:833-41.
8. Stavrou D, Weissman O, Tessone A, et al. Health related quality of life in burn patients–a review of the literature. Burns 2014; 40:788-96.
9. Spronk I, Legemate C, Oen I, et al. Health related quality of life in adults after burn injuries: a systematic review. PLoS One 2018; 13:e0197507.
10. Spronk I, Legemate CM, Polinder S, et al. Health-related quality of life in children after burn injuries: A systematic review. J Trauma Acute Care Surg 2018; 85:1110-8.
11. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018; 27:1147-57.
12. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res 2018; 27:1171-9.
13. 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; 27:1159-70.
14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 6:e1000097.
15. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of Life Research 2010; 19:539-49.
16. Kool MB, Geenen R, Egberts MR, et al. Patients' perspectives on quality of life after burn. Burns 2017.
17. Meirte J, Van Daele U, Maertens K, et al. Convergent and discriminant validity of quality of life measures used in burn populations. Burns 2017; 43:84-92.

18. Spronk I, Legemate C, Dokter J, et al. Predictors of health related quality of life after burn injuries: a systematic review. Crit Care 2018; 22:160.
19. Finlay V, Phillips M, Wood F, et al. Enhancing the clinical utility of the Burn Specific Health Scale-Brief: not just for major burns. Burns 2014; 40:328-36.
20. Gandolfi S, Auquit-Auckbur I, Panunzi S, et al. Validation of the French version of the Burn Specific Health Scale-Brief (BSHS-B) questionnaire. Burns 2016; 42:1573-80.
21. Goudarzian AH, Taebei M, Soleimani A, et al. Burn Specific Health Scale-Brief (BSHS-B) in Pregnant Burned Women: Translation and Psychometric Evaluation of the Persian Version. International Journal of Pediatrics-Mashhad 2017; 5:5391-400.
22. Hwang YF, Chen-Sea MJ, Chen CL, et al. Validation of a taiwanese version of the burn-specific health scale-brief. J Burn Care Res 2016; 37:e310-e6.
23. Kildal M, Andersson G, Fugl-Meyer AR, et al. Development of a brief version of the Burn Specific Health Scale (BSHS-B). J Trauma Acute Care Surg 2001; 51:740-6.
24. Ling-Juan Z, Jie C, Jian L, et al. Development of quality of life scale in Chinese burn patients: Cross-cultural adaptation process of burn-specific health scale - Brief. Burns 2012; 38:1216-23.
25. Mulay AM, Ahuja A, Ahuja RB. Modification, cultural adaptation and validation of burn specific health scale-brief (BSHS-B) for Hindi speaking population. Burns 2015; 41:1543-9.
26. Müller A, Smits D, Jasper S, et al. Validation of the German version of the Burn Specific Health Scale-Brief (BSHS-B). Burns 2015; 41:1333-9.
27. Piccolo MS, Gragnani A, Daher RP, et al. Validation of the Brazilian version of the Burn Specific Health Scale-Brief (BSHS-B-Br). Burns 2015; 41:1579-86.
28. Pishnamazi Z, Rejeh N, Heravi-Karimooi M, et al. Validation of the Persian version of the Burn Specific Health Scale-Brief. Burns 2013; 39:162-7.
29. Sideli L, Di Pasquale A, Prestifilippo A, et al. Validation of the Italian version of the burn specific health scale-brief. Burns 2014; 40:995-1000.
30. Sideli L, Prestifilippo A, Di Benedetto B, et al. Quality of life, body image, and psychiatric complications in patients with a burn trauma: preliminary study of the italian version of the burn specific health scale-brief. Ann Burns Fire Disasters 2010; 23:171-6.
31. Stampolidis N, Castana O, Nikiteas N, et al. Quality of life in burn patients in Greece. Ann Burns Fire Disasters 2012; 25:192-5.
32. Stavrou D, Haik J, Wiser I, et al. Validation of the Hebrew version of the Burn Specific Health Scale-Brief questionnaire. Burns 2015; 41:188-95.
33. Szczechowicz J, Lewandowski J, Sikorski J. Polish adaptation and validation of Burn Specific Health Scale - Brief. Burns 2014; 40:1013-8.
34. Willebrand M, Kildal M. Burn specific health up to 24 months after the burn—a prospective validation of the simplified model of the Burn Specific Health Scale-Brief. J Trauma Acute Care Surg 2011; 71:78-84.
35. Stolle A, Ripper S, Magdanz J, et al. Validation of the Ludwigshafen German

Version of the Burn Specific Health Scale-Brief. J Burn Care Res 2018; 39:252-60.

36. Van Loey NE, Van de Schoot R, Gerdin B, et al. The Burn Specific Health Scale-Brief: measurement invariant across European countries. J Trauma Acute Care Surg 2013; 74:1321-6.

37. Willebrand M, Kildal M. A simplified domain structure of the burn-specific health scale-brief (BSHS-B): a tool to improve its value in routine clinical work. J Trauma 2008; 64:1581-6.

38. Adam M, Leblebici B, Tarim MA, et al. Validation of a Turkish version of the burn-specific health scale. J Burn Care Res 2009; 30:288-91; discussion 92-3.

39. Li DW, Liu WQ, Wang HM, et al. The Chinese language version of the abbreviated Burn Specific Health Scale: A validation study. Burns 2014; 40:1001-6.

40. Moi AL, Wentzel-Larsen T, Salemark L, et al. Erratum: Validation of a Norwegian version of the Burn Specific Health Scale (Burns (2003) 29: 6 (563-570)). Burns 2003; 29:872-3.

41. Munster AM, Fauerbach JA, Lawrence J. Development and utilization of a psychometric instrument for measuring quality of life in burn patients, 1976 to 1996. Acta Chir Plast 1996; 38:128-31.

42. Salvador Sanz JF, Sanchez-Payá J, Marín JR. Spanish version of the Burn-Specific Health Scale. J Trauma Inj Infect Crit Care 1998; 45:581-7.

43. Blalock SJ, Bunker BJ, DeVellis RF. Measuring health status among survivors of burn injury: revisions of the Burn Specific Health Scale. J Trauma Acute Care Surg 1994; 36:508-15.

44. Blalock SJ, Bunker BJ, Moore JD, et al. The impact of burn injury: a preliminary investigation. J Burn Care Rehabil 1992; 13:487-92.

45. Ferreira E, Dantas RAS, Rossi LA, et al. The cultural adaptation and validation of the "Burn Specific Health Scale-Revised" (BSHS-R): Version for Brazilian burn victims. Burns 2008; 34:994-1001.

46. Nicolosi JT, de Carvalho VF, Sabatés AL, et al. Assessment of health status of adolescent burn victims undergoing rehabilitation: a cross-sectional field study. Plast Surg Nurs 2013; 33:185-91.

47. Daltroy LH, Liang MH, Phillips CB, et al. American burn association/shriners hospitals for children burn outcomes questionnaire: Construction and psychometric properties. J Burn Care Rehabil 2000; 21:29-39.

48. Sveen J, Huss F, Sjöberg F, et al. Psychometric properties of the swedish version of the burn outcomes questionnaire for children aged 5 to 18 years. J Burn Care Res 2012; 33:e286-e94.

49. Van Baar ME, Essink-Bot ML, Oen IMMH, et al. Reliability and validity of the dutch version of the American burn association/Shriners hospital for children burn outcomes questionnaire (5-18 years of age). J Burn Care Res 2006; 27:790-802.

50. Tyack Z, Kimble R, McPhail S, et al. Psychometric properties of the Brisbane Burn Scar Impact Profile in adults with burn scars. PloS One 2017; 12:e0184452.

51. 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; 41:1405-19.

52. Kazis LE, Liang MH, Lee A, et al. The development, validation, and testing of a health outcomes burn questionnaire for infants and children 5 years of age and younger: American Burn Association/ Shriners Hospitals for Children. J Burn Care Res 2002; 23:196-207.

53. Van Baar ME, Essink-Bot ML, Oen IMMH, et al. Reliability and validity of the Health Outcomes Burn Questionnaire for infants and children in The Netherlands. Burns 2006; 32:357-65.

54. Blades B, Munster AM. A burn specific health scale. J Trauma 1982; 22:872-5.

55. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol 1994; 19:210-6.

56. Mazharinia N, Aghaei S, Shayan Z. Dermatology life quality index (DLQI) scores in burn victims after revival. J Burn Care Res 2007; 28:312-7.

57. Öster C, Willebrand M, Dyster-Aas J, et al. Validation of the EQ-5D questionnaire in burn injured adults. Burns 2009; 35:723-32.

58. Landgraf JM, Vogel I, Oostenbrink R, et al. Parent-reported health outcomes in infants/ toddlers: measurement properties and clinical validity of the ITQOL-SF47. Qual Life Res 2013; 22:635-46.

59. Edgar D, Dawson A, Hankey G, et al. Demonstration of the validity of the SF-36 for measurement of the temporal recovery of quality of life outcomes in burns survivors. Burns 2010; 36:1013-20.

60. Ryan CM, Schneider JC, Kazis LE, et al. Benchmarks for multidimensional recovery after burn injury in young adults: The development, validation, and testing of the american burn association/shriners hospitals for children young adult burn outcome questionnaire. J Burn Care Res 2013; 34:e121-e42.

61. Munster AM, Horowitz GL, Tudahl LA. The abbreviated Burn-Specific Health Scale. J Trauma 1987; 27:425-8.

62. Streiner D NG, Cariney J. Health measurement scales: a practival guide to their development and use2015.

63. Falder S, Browne A, Edgar D, et al. Core outcomes for adult burn survivors: a clinical overview. Burns 2009; 35:618-41.

64. Young A, Brookes S, Rumsey N, et al. Agreement on what to measure in randomised controlled trials in burn care: study protocol for the development of a core outcome set. BMJ Open 2017; 7:e017267.

65. Boers M, Kirwan JR, Wells G, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014; 67:745-53.

66. Kazis LE, Marino M, Ni P, et al. Development of the life impact burn recovery evaluation (LIBRE) profile: assessing burn survivors' social participation. Qual Life Res 2017; 26:2851-66.

6

Clinically relevant changes and differences in scar outcome for burn patients measured by the Patient and Observer Scar Assessment Scale

C.M. Legemate, E. Middelkoop, M.E. Carrière, P.P.M. van Zuijlen, M.E.. van Baar, C.H. van der Vlies, the Hycon Study group

Submitted

ABSTRACT

Background

The Patient and Observer Scar Assessment Scale (POSAS) is frequently used to assess scar quality after burns. It is important to be aware of the minimal important change (MIC) and the minimal clinically important difference (MCID) to establish if a POSAS score represents a clinically relevant change or difference. The aim of this study is to explore the MIC and MCID of POSAS version 2.0.

Methods

During a prospective cohort study, POSAS data was obtained at three, six, and 12 months after split skin grafting. At the second and third visits, patients rated the degree of clinical change in scar quality in comparison to the previous visit. At 12 months, they completed the POSAS for a second scar and rated the degree of clinical difference between the two scars. Two anchor-based methods were used to determine the MIC and MCID.

Results

MIC values of the patient POSAS ranged from -0.59 to -0.29 between three and six months and from -0.75 to -0.38 between six and 12 months follow-up. Both had a poor discriminatory value. MCID values ranged from -0.39 and -0.08, with a better discriminatory value.

Conclusion

Our results suggest that patients consider minor differences (less than 0.75 on the 1–10 scale) in POSAS scores as clinically important scar quality changes. MCID values can be used to evaluate the effects of burn treatment and perform sample-size calculations.

INTRODUCTION

The Patient and Observer Scar Assessment Scale version 2.0 (POSAS 2.0) measures scar quality and is a frequently used outcome in burn studies[1-3]. The patient scale gives the POSAS an important extra dimension because the patient's opinion is included, which is a necessary component for a complete scar evaluation[4]. To interpret changes and differences in scores arising from patient reported questionnaires it is important to be aware of the minimal important change (MIC) and the minimal clinically important difference (MCID). The MIC represents the smallest change in score that would be perceived by the patient as important and refers to changes in scores within one patient over time[5]. The MCID represents the smallest difference in score that would be perceived by the patient as important and is used to compare scores between individuals or groups at the same time (i.e. the difference between two trial arms)[6, 7].

If a study finds a difference in POSAS scores that is less than the MIC or MCID, it may not represent a true clinically-relevant change or difference, even if it is statistically significant. On the other hand, if the difference is greater than the MIC or MCID, it might be important to patients even though it is not statistically significant. As both terms define an outcome that is considered important to patients, it may serve as the basis for estimating the necessary sample size in designing future studies.

To our knowledge, the MIC and MCID of the POSAS have never been studied in patients after burn surgery. Therefore, the purpose of our study is three-fold: 1) explore the MIC for a change in POSAS score for scars of skin-grafted burn wounds between three and six months post-surgery, 2) explore the MIC for a change in POSAS score for scars of skin-grafted burn wounds between six and 12 months post-surgery, and 3) explore the MCID for a difference in POSAS score for scars of skin-grafted burn wounds at 12 months post-surgery.

PATIENTS AND METHODS

This study was conducted alongside a prospective multicentre clinical trial that was undertaken in three dedicated burn centres in the Netherlands[8]. The aim of the trial was to compare scar quality after hydrosurgical debridement versus conventional knife-debridement of deep dermal burns before split grafting. The regional Ethics Committee granted approval before initiation of this parallel study (reference number L2016087). Patients were recruited between January 2017 – October 2019. Participants had deep dermal burns with two similar wound areas of at least 25 cm^2 that required excision and grafting. Patient were excluded if they: 1) did not want to complete the questionnaire at the outpatient clinic, 2) did not complete the anchor-based questions, or 3) were unable to understand the study information. All the participants provided written informed consent.

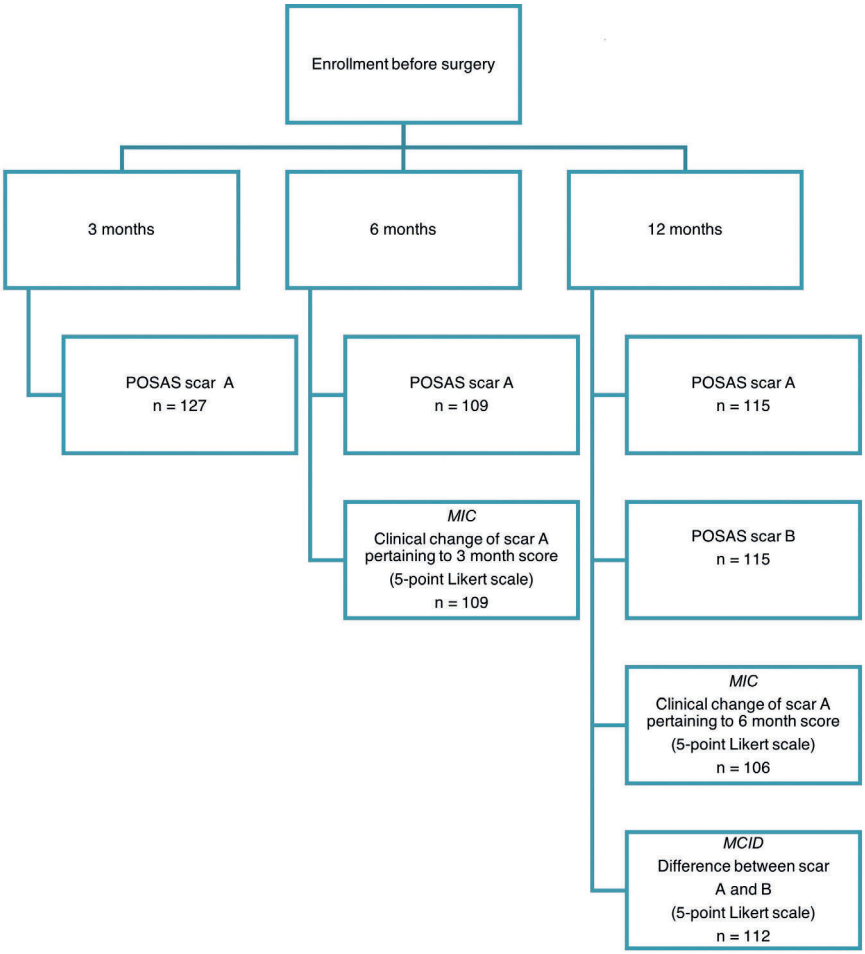


Figure 1. Schedule of assessments

Data collection

Before split skin grafting, the surgeon designated two similar wound areas as study areas A and B. If possible, the study areas were adjacent. Otherwise, a similar burn wound at the contralateral body part, or the nearest comparable burn wound, was chosen.

At three-, six-, and 12-months post-surgery, patients completed the patient part of the Dutch version of the POSAS 2.0 for the scar on study area A (Figure 1). At the second and third visits, patients were asked to indicate the degree of clinical change they had noticed since the previous visit on a five-point Likert scale (much worse, slightly worse, unchanged, slightly better, much

better) (Figure 1). If patients found it difficult to remember what their scar looked like, they were shown a picture from the scar during their earlier visit.

At the third visit, patients additionally completed the POSAS 2.0 for the scar of study area B and were asked to rate the degree of clinical difference on the same five-point Likert scale. The purpose of the Likert scales was to "anchor" the change or difference observed in POSAS scores to patients' perspectives regarding what is clinically important.

POSAS patient scale

The patient scale of the POSAS consists of six descriptive items: pain, itch, colour, thickness, relief, and pliability. Pain and itch are scored between 1 (no pain/itch) and 10 (extreme pain/itch). Each of the other items is scored between 1 (no difference with normal skin) and 10 (very different from normal skin). Patient POSAS scores were calculated as the mean score of the six POSAS items. A lower score in POSAS correlates with a better scar.

Statistical methods

Although there is no consensus about the best method to determine the MIC and MCID, anchor-based approaches are preferred above distribution-based approaches[5, 9]. Using an anchor-based approach, the MIC and MCID were derived by comparing the change or difference in score to another measure that defines a clinically relevant change or difference (i.e. the anchor). To determine the MIC, changes in POSAS score for the scar of study area A were calculated at 2 time points: between the first and second visit, and between the second and third visit (Figure 1). To determine the MCID, the difference between POSAS scores for the scar of study area A and B at 12 months was calculated (Figure 1). Multiple anchor-based approaches exist and recent publications recommend to use multiple methods[5, 10-13]. Two well-researched anchor-based methods were used here:

The mean change method

With the mean change method, the MIC and MCID are defined as the mean difference in POSAS score in patients who considered their scars to be minimally importantly changed or different according to the anchor (e.g. in patients who rate their scar as 'slightly improved' or 'slightly better')[5, 14]. The adequacy of the Likert scale was explored by quantifying the correlation between change in POSAS scores and the anchor-based questions using Spearman's rho. Correlation coefficients were interpreted as negligible correlation (0-0.3), low correlation (0.3-0.5), moderate correlation (0.3-0.5), high correlation (0.7-0.9), or very high correlation (0.9-1.0)[15]. In addition, significant score changes were tested among patients who indicated different categories on the anchor-based questions, using the Kruskal-Wallis test. Non-significant differences among the five-point Likert scale of the anchor-based questions could suggest that the categories were not sufficiently discriminative.

The receiver operating characteristic method

The receiver operating characteristic (ROC) method is similar to the method used in diagnostic test research, whereby the POSAS score was considered the diagnostic test and the anchor served as the gold standard. Using the anchor, the patients were divided into two groups: those who reported their scars as importantly changed or better (i.e. slightly or much improved/better) and those who reported their scars as not changed or not better (i.e. slightly or much deteriorated/worse). Then, the distribution of the change and difference scores of the POSAS in both groups was plotted, and sensitivity and specificity values were calculated. For the ROC curve, sensitivity was plotted against 1 – specificity for all possible cut-off points. The optimal cut off point gives the smallest change/difference score of misclassifying importantly improved/better and not-improved/better scars and is therefore considered the MIC or MCID[5, 16]. The ability to distinguish patient groups was expressed as the area under the curve (AUC) of the ROC, and was considered excellent (AUC ≥0.9), good (0.8 to 0.899), adequate (0.7 to 0.799), or poor (<0.7)[17].

RESULTS

The study population consisted of 127 patients with a mean of 44 years (range 0 – 87) and total body surface area of 10% (range 0.5 – 55). The mean POSAS score at three months was 4.3 (standard deviation (SD) 1.9), at six months was 3.7 (SD 1.9), and at 12 months was 3.3 (SD 1.8).

Minimal important change (MIC)

Most patients reported improvement between the two time points, with 87% improved at six months and 84% at 12 months (Table 1).

Mean change method

With the mean change method, the MIC was -0.29 (95% confidence interval (CI) -0.62, 0.04) between three and six months, and -0.38 (95% CI -0.75, -0.01) between six and 12 months (Table 1). The correlation between the change in POSAS scores and the Likert scale was low at both time points, suggesting insufficient discriminative categories (correlation coefficient 0.16 at six months, 0.08 at 12 months). There were no significant differences in changes in POSAS scores between patients who scored the change as much improvement, slight improvement, no change, slight deterioration, or much deterioration in scar quality ($p = 0.383$ at six months, $p = 0.489$ at 12 months).

Table 1. Mean changes in POSAS scores

	Number of patients Total = 109	POSAS mean item score Change between 3-6 months, Mean (95% CI)	Number of patients Total = 106	POSAS mean item score Change between 6-12 months, Mean (95% CI)
Much improved	46	-0.80 (-1.28 – -0.32)	32	-0.78 (-1.20 – -0.36)
Slightly improved	49	-0.29 (-0.62 – 0.04)	57	-0.38 (-0.75 – -0.01)
Unchanged	8	-0.08 (-1.24 – 1.07)	15	-0.79 (-1.84 – 0.26)
Slightly deteriorated	6	0.36 (-1.79 – 2.50)	2	0.16 (-12.54 – 12.87)
Much deteriorated	0	NA	0	NA

ROC method

Table 2 shows the optimal cut-off points that were found at each follow-up point. Using this method, the MIC was -0.59 between three and six months and -0.75 between six and 12 months. The AUC of the change in POSAS score to differentiate between patients with improvement in scar quality and patients with no improvement was low at 0.59 (95% CI 0.44, 74) and 0.50 (95% CI 0.34, 0.66), indicating poor discrimination.

Table 2. Values from the ROC-curve analysis. Patients were divided into two groups: those who reported their scars as importantly changed or better (i.e. 'slightly' or 'much' improved/better) and those who reported their scars as not changed or not better (i.e. 'slightly' or 'much' deteriorated/worse).

	Cut-off point	AUC (95% CI)	Sensitivity	Specificity
MIC score between 3-6 months	-0.59	0.59 (0.44 – 0.74)	0.46	0.79
MIC score between 6-12 months	-0.75	0.50 (0.34 – 0.66)	0.49	0.59
MCID score at 12 months	-0.08	0.79 (0.71 – 0.87)	0.69	0.81

Minimal Clinically Important Difference (MCID)

Mean change method

With the mean change method, the MCID was -0.39 (95%CI -0.69, -0.08) (Table 3). There was a significant difference in POSAS scores for scar A compared to scar B ($p < 0.001$), suggesting sufficient discriminative categories. The correlation coefficient between the difference in POSAS scores and Likert scale categories was 0.6.

ROC method

Using the ROC curve method, the optimal cut-off point to represent the MCID was -0.08. The AUC of the difference in POSAS score to differentiate between patients that indicated their scar as better and patients with no difference or worse was 0.79 (95% CI 0.71, 0.87) (Table 2), indicating acceptable discrimination.

Table 3. Mean differences in POSAS scores for scars A and B at 12 months post-surgery

Total POSAS	Number of patients Total = 112	POSAS mean item score Difference, Mean (95% CI)
Much better	18	-1.20 (-1.74 – -0.67)
Slightly better	37	-0.39 (-0.69 – -0.08)
No difference	29	-0.07 (-0.21 – 0.08)
Slightly worse	21	0.81 (0.23 – 1.38)
Much worse	7	2.67 (1.14 – 4.19)

DISCUSSION

The Patient Scale of the POSAS 2.0 is a well-accepted measure of patient reported scar quality after burns. This was the first study to explore the MIC and MCID of the POSAS. Two anchor-based methods were used at two follow-up timepoints, and MIC values were found to range from -0.29 to -0.75. POSAS change scores between three and six months seem to adequately increase (i.e. clinically worsen) with deterioration of scar indicated by patients. For improved scars, however, POSAS change scores in the much improved group were the same as the unchanged group, between six and 12 months. This was probably due to the small group size (15 patients) and a large variation in scores, as indicated by the wide confidence interval. Nevertheless, changes in POSAS score did not differ statistically significant between the groups in different anchor categories and there was poor correlation between the changes in POSAS score and anchor categories. This suggests that the categories were not sufficiently discriminative and did not adequately represent the changes in POSAS score.

Although patients were divided into 2 categories in the ROC analyses, which created larger groups, the MICs derived using this method also had low discriminative abilities. The AUC, sensitivity, and specificity were low for the best cut-off points at six and 12 months, which indicates that the cut-off value was poor in distinguishing false positives and false negatives. In this case, false positives represent patients who have a larger change score than the MIC but reported themselves to be unchanged or deteriorated. False negatives represent patients who have a smaller change score than the MIC but reported themselves to be importantly improved. It may be that patients found it difficult to remember what their scar looked like three or six months earlier, even though they had access to pictures of their scar. Acceptance or other psychological factors may have influenced their POSAS ratings overtime.

The MCID determined with the mean change method was -0.39. The group categories of the Likert scale were sufficiently discriminative and correlated adequately with differences in POSAS

scores. The MCID value of -0.08 that was derived using the ROC method also showed adequate discriminatory values.

In accordance with the literature, different anchor-based methods were used to estimate the MIC and MCID[5, 10-13]. The anchor-based approaches that we used are preferred over distribution-based methods. Distribution-based methods use statistical parameters, such as the standard deviation and standard error of measurement to estimate the MIC[5, 10-13]. The advantage of anchor-based methods is that the concept of minimal importance, and therefore the patient's experience, is clearly defined and incorporated in the methods.

However, the methods do not consider the variability of the scores of the instrument in the sample. Also, there is a lack in consensus about the best anchor-based methodology. One drawback of the mean change method described is that the mean value does not reflect a threshold for minimal improvement, as it uses the mean of the group that reported to be a little better on the anchor-based questions. Since all patients in this group reported their scar to be minimal importantly changed on the anchor, the mean change in score on the patient reported outcome measure of interest in this group is higher than the threshold for minimal important change[7]. Also, the method does not consider the values of patients that reported much improvement even though this is also an important change to patients. This could explain the lower cut-off values found in the ROC-analyses. The advantage of the ROC method is that is uses the entire study sample, which results in more reliable estimates than the mean change method. Moreover, it estimates the threshold between not changed/no difference and slightly improved/slightly better.

Strengths & Limitations

A limitation of this study is that only one dataset was used for analysis, which may limit the generalizability of the results. One drawback of the mean change method is that the subgroup of patients who reported to be slightly improved is often small, which results in imprecise estimates[18]. An advantage of the ROC method is that it uses the entire study sample, leading to more reliable estimates. However, values are estimated higher when the proportion improved is greater than 50% and lower when the proportion improved is less than 50%[19]. In the MIC estimates, the group that improved was far more than 50% and results are therefore probably biased. In the MCID estimate, however, the proportion of improved patients was 49%, indicating a low risk of bias. There are currently no guidelines for correlating the results of the mean change and ROC method. Nevertheless, the results of the current study should improve the interpretability of change scores and difference for users of the POSAS and can be used for clinical trial sample size calculations. Currently, a new version of the POSAS (POSAS 3.0) is being developed. Given the lack of patient input in the development of the Patient Scale of the POSAS 2.0[20], this study provides valuable data to fill this gap. As part of this current research initiative, measurement properties like the MIC will be compared between versions 2.0 and 3.0.

CONCLUSION

This study explored the MIC and MCID of the patient scale of the POSAS 2.0 using two anchor-based approaches. MIC values could not be calculated with good discrimination between three to six and six to 12 months. However, the item score of the POSAS ranges from 1 – 10 and an important finding of the analyses was that patients seemed to consider minor differences as important. At 12-months post-surgery, MCID values of -0.39 and -0.08 demonstrated better discrimination. Therefore, if scars are compared at the same time (e.g. after two treatment methods), patients were able to distinguish scars that were at least slightly better. The threshold for a difference in POSAS score that patients would consider important lies between these values.

Acknowledgments

This research was funded by the Dutch Burns Foundation (grant no. 15.101).

Disclosure

Dr. van Zuijlen is the developer of the Patient and Observer Scar Assessment Scale. The authors have no financial interest to declare in relation to the content of this communication.

The Hycon Study Group

Gerard I.J.M. Beerthuizen, Jan Dokter, Kim L.M. Gardien, Harold Goei, Jakob Hiddingh, Helma W.C. Hofland, Kelly A.A. Kwa, Ymke Lucas, Annebeth Meij-de Vries, Marianne K. Nieuwenhuis, Tim H.J. Nijhuis, Irma M.M.H. Oen, Anouk Pijpe, Dorotka T. Roodbergen, Sonja M.H.J. Scholten-Jaegers, Matthea Stoop, Marcel Vlig, Nicole Trommel

REFERENCES

1. Tyack Z, Wasiak J, Spinks A, Kimble R, Simons M. A guide to choosing a burn scar rating scale for clinical or research use. *Burns* 2013;39(7): 1341-1350.
2. Carriere ME, Kwa KAA, de Haas LEM, Pijpe A, Tyack Z, Ket JCF, van Zuijlen PPM, de Vet HCW, Mokkink LB. Systematic Review on the Content of Outcome Measurement Instruments on Scar Quality. *Plast Reconstr Surg Glob Open* 2019;7(9): e2424.
3. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7): 737-745.
4. Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, van Zuijlen PP. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. *Plast Reconstr Surg* 2004;113(7): 1960-1965; discussion 1966-1967.
5. De Vet HCW TC, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge University Press: Cambridge, United Kingdom, 2011.
6. de Vet HC, Ostelo RW, Terwee CB, van der Roer N, Knol DL, Beckerman H, Boers M, Bouter LM. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2007;16(1): 131-142.
7. Terwee CB, Peipert JD, Chapman R, Lai JS, Terluin B, Cella D, Griffith P, Mokkink LB. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Qual Life Res* 2021.
8. Legemate CM, Goei H, Middelkoop E, Oen I, Nijhuis THJ, Kwa KAA, van Zuijlen PPM, Beerthuizen G, Nieuwenhuis MK, van Baar ME, van der Vlies CH. Long-term scar quality after hydrosurgical versus conventional debridement of deep dermal burns (HyCon trial): study protocol for a randomized controlled trial. *Trials* 2018;19(1): 239.
9. van der Willik EM, Terwee CB, Bos WJW, Hemmelder MH, Jager KJ, Zoccali C, Dekker FW, Meuleman Y. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. *Nephrology (Carlton)* 2021;26(5): 391-399.
10. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61(2): 102-109.
11. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes* 2006;4: 70.
12. Wyrwich KW, Metz SM, Kroenke K, Tierney WM, Babu AN, Wolinsky FD. Triangulating patient and clinician perspectives on clinically important differences in health-related quality of life among patients with

heart disease. *Health Serv Res* 2007;42(6 Pt 1): 2257-2274; discussion 2294-2323.

13. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, Croft P, de Vet HC. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010;63(5): 524-534.

14. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10(4): 407-415.

15. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012;24(3): 69-71.

16. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 1986;39(11): 897-906.

17. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60(1): 34-42.

18. Terwee CB, Peipert JD, Chapman R, Lai JS, Terluin B, Cella D, Griffith P, Mokkink LB. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Qual Life Res* 2021;30(10): 2729-2754.

19. Terluin B, Eekhout I, Terwee CB. The anchor-based minimal important change, based on receiver operating characteristic analysis or predictive modeling, may need to be adjusted for the proportion of improved patients. *J Clin Epidemiol* 2017;83: 90-100.

20. van Zuijlen PPM, Mokkink LB, Hoogewerf CJ, de Vet HCW. The official update of the POSAS: An invitation to share experiences to improve the POSAS in 'Project POSAS 3.0'. *Burns* 2017;43(4): 893-894.

PART III
OUTCOMES OF BURN SURGERY

7

Application of hydrosurgery for burn wound debridement: An 8-year cohort analysis

C.M. Legemate, H. Goei, O.F.E. Gostelie, T.H.J. Nijhuis, M.E. van Baar, C.H. van der Vlies, the Dutch Burn Repository Group R3

Burns. 2019 Feb;45(1):88-96

ABSTRACT

Introduction

During the last decade, the Versajet™ hydrosurgery system has become popular as a tool for tangential excision in burn surgery. Although hydrosurgery is thought to be a more precise and controlled manner for burn debridement prior to skin grafting, burn specialists decide individually whether hydrosurgery should be applied in a specific patient or not. The aim of this study was to gain insight in which patients hydrosurgery is used in specialized burn care in the Netherlands.

Methods

A retrospective study was conducted in all patients admitted to a Dutch burn centre between 2009 and 2016. All patients with burns that underwent surgical debridement were included. Data were collected using the national Dutch Burn Repository R3.

Results

Data of 2113 eligible patients were assessed. These patients were treated with hydrosurgical debridement (23.9%), conventional debridement (47.7%) or a combination of these techniques (28.3%). Independent predictors for the use of hydrosurgery were a younger age, scalds, a larger percentage of total body surface area (TBSA) burned, head and neck burns and arm burns. Differences in surgical management and clinical outcome were found between the three groups.

Conclusion

The use of hydrosurgery for burn wound debridement prior to skin grafting is substantial. Independent predictors for the use of hydrosurgery were mainly burn related and consisted of a younger age, scalds, a larger TBSA burned, and burns on irregularly contoured body areas. Randomized studies addressing scar quality are needed to open new perspectives on the potential benefits of hydrosurgical burn wound debridement.

INTRODUCTION

In the last decade, hydrosurgery has become available in burn surgery as an alternative technique for tangential excision alongside the golden standard of conventional tangential excision by guarded knives. The hydrosurgical device used in the treatment of burns is usually known as the Versajet™. The Versajet™ hydrosurgery system (Smith and Nephew, St. Petersburg, FL, USA) was developed in 1997 for soft tissue debridement in various types of wounds. The Versajet™ hydrosurgery system works by producing a high-pressure jet of water across an aperture in an angled hand piece. Through the Venturi principle, the jet creates a suction force that draws tissues into the path of the fluid where they are ablated and sucked into the device together with the irrigation fluid[1,2]. Power settings can be adjusted to control the cutting and aspiration effects, depending on the depth of debridement the surgeon wants to achieve[3]. Although hydrosurgery was introduced for burn wound debridement in Dutch burn care in 2006, it only became widely used in 2008[4].

A report of the National Institute for Health Care and Excellence (NICE) presented an overview of the studies concerning the safety of hydrosurgery[2]. The majority of these studies showed good clinical results with minimal adverse outcomes in both adults and children with acute and chronic wounds[4-11]. Studies on burn wounds showed that the Versajet system may be faster and more precise in obtaining the desirable excision plane. Nevertheless, the Versajet has typically not been favoured in deeper burns due to belief its penetration is less efficient in thick eschar, as it 'bounces' off the hard tissue and causes irregular grooves[2,7]. Burn specialists widely use hydrosurgery as an alternative for conventional debridement prior to skin grafting, however, only two randomized controlled trials comparing hydrosurgical and conventional debridement in patients with burns have been published[7,12]. These studies reported a significant reduction in excision time and better preservation of viable tissue after hydrosurgical debridement. Nevertheless, no significant differences were found on postoperative pain, contracture rates, healing time, graft take, post-operative infection, bacterial burden and scar quality at 6 months post burn. Whether these results influenced the current application of the Versajet™ system is unexplored. To our knowledge, no algorithm is available for burn specialists guiding them whether or not to use hydrosurgery. Due to an absent algorithm and a paucity of studies, the clinical application of hydrosurgery in burn care remains unknown.

The aim of this study was to gain insight in which patients hydrosurgery is used in specialized burn care in the Netherlands and whether the actual application of hydrosurgical application matches the currently available literature. Furthermore, surgical outcomes of different debridement techniques are examined.

METHODS

Study design and population

In this cohort study, all patients with a burn-related admission in one of the burn centres in the Netherlands (Maasstad Hospital in Rotterdam, Martini Hospital in Groningen, and Red Cross Hospital in Beverwijk) between January 2009 and 31 December 2016 were included.

Data collection

Data were obtained from the national burn registry of the three burn centres in the Netherlands (Dutch Burn Repository R3) which started collecting data from 2009 onwards. The database is filled by dedicated burn care professionals, and quality monitoring by a coordinator and improvement is formally organized. Data on patient characteristics, burn, treatment, and outcome were documented (Table 1, Table 4).

Data analysis

Eligible patients were divided into three groups: hydrosurgical debridement, hydrosurgical in combination with conventional debridement and a conventional debridement group (Fig. 1). The proportions of patient and injury related characteristics were compared between the three groups. Patients were divided into an early surgery group (<7 days after burn) and a delayed surgery group (>7 days after burn) to evaluate the effect of timing of surgery on the use of hydrosurgery. A subgroup analysis of patients with only one body part burned was performed to identify the prevalence of the use of hydrosurgery per affected body site.

Statistical analysis

Data were analysed using IBM SPSS Statistics for Windows version 23.0 (IBM Corp., NY, USA). Outcomes were reported as percentages for categorical variables. Continuous variables were summarized as either means with corresponding standard deviations (SD) or medians with interquartile ranges (IQR) depending on normality of distribution. Univariable logistic regression analysis was performed to identify parameters that were associated with the use of hydrosurgery. Parameters that were associated in univariable analysis ($p < 0.10$) were checked for multicollinearity (Spearman's r (rs) >0.75) and subsequently entered into a multivariable logistics regressions analysis (forward stepwise LR). Differences in patient, and injury related characteristics, differences in surgical treatment and outcome between the three groups were compared using the chi-squared (categorical data) or Kruskal–Wallis (continuous data) test. Differences in surgical treatment and outcome between the groups treated with hydrosurgiery alone and the group treated with conventional techniques alone were compared using the chi-squared (categorical data) or Mann–Whitney (continuous data) test. Two-tailed p values below 0.05 were considered statistically significant for all statistical tests.

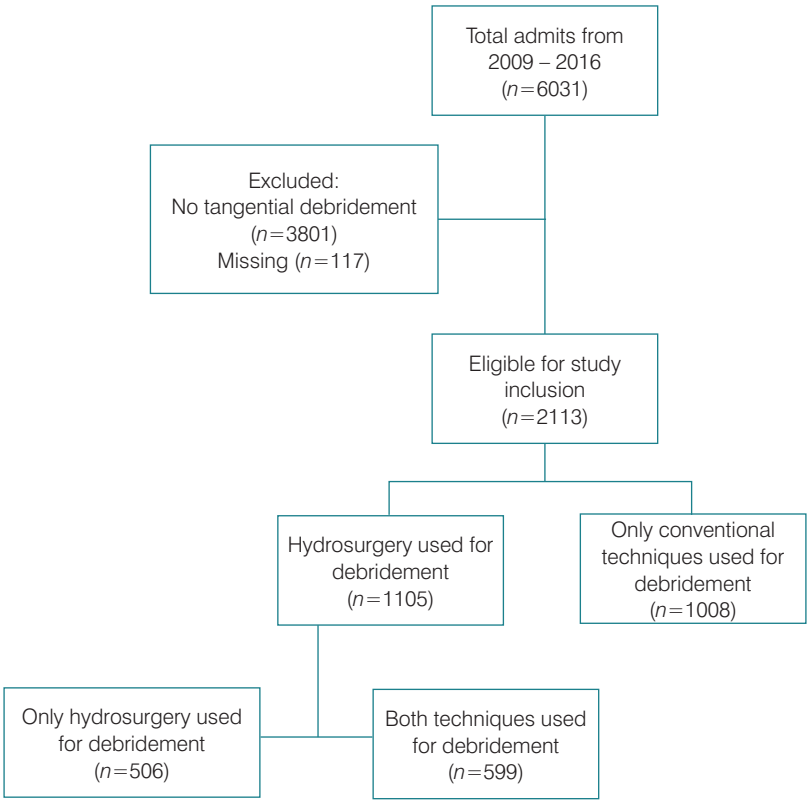


Figure 1. Patient inclusion flowchart

RESULTS

Inclusions

A total of 6031 patients had been admitted in the three Dutch Burn centres between January 2009 and December 2016. In total 63.0% of the patients was excluded because they did not have surgical debridement of their wounds and 1.9% was excluded because of lack of information on the used technique during surgery (Fig. 1). The final study population consisted of 2113 patients (59.5% males) with a median age of 41 years (IQR 36) and median TBSA of 5% (IQR 10). Patient and injury characteristics per group are shown in Table 1.

Prevalence of the use of hydrosurgery and predictors

In 52.3% (n = 1105) of the included patients hydrosurgery was used for debridement of their burn wounds. In 23.9% (n = 506) of these patients hydrosurgery was used exclusively for debridement of their burn wounds and in 28.3% (n = 599) hydrosurgery was used in combination with conventional debridement. The mean prevalence in the period 2009–2016 was 25.3% (Fig.2).

Table 1. Patient and injury characteristics

	Hydrosurgery used for debridement		Only conventional techniques used for debridement	
	Only hydrosurgery (n = 506)	**Both techniques (n = 599)**	**(n = 1008)**	p value*
Total, %	23.9	28.3	47.7	
Median age at injury (IQR)¥	29 (42)	43 (35)	44 (35)	<0.001
Age in categories, %				
0-4y	21.9	7.8	6.7	<0.000
5-17y	13.4	9.8	9.3	0.011
18-65y	54.5	66.6	64.1	<0.001
>65y	10.1	16.4	20.0	<0.001
Gender, %				0.059
Female	37.2	39.0	43.1	
Male	62.8	61.0	56.9	
Aetiology, %¶				
Scald	37.9	21.6	18.8	<0.001
Flame	43.1	44.6	64.0	<0.001
Grease	11.0	9.0	7.9	0.207
Contact	4.8	13.6	3.7	<0.001
Other	3.2	11.1	5.5	<0.001
Median % TBSA burned (IQR)	5.0 (8)	11.0 (17.8)	2.0 (5.5)	<0.001
TBSA burned in categories, %				
<1	6.5	3.8	23.2	<0.001
1-2	12.1	4.7	21.1	<0.001
2-5	25.7	17.0	21.4	<0.001
5-10	28.1	20.5	18.3	0.001
10-20	20.2	25.2	9.9	0.001
>20	7.5	28.7	6.1	<0.001
Body location, %†				
Head and neck	42.5	52.4	22.0	<0.001
Trunk	43.9	62.9	35.8	<0.001
Arms	69.4	78.5	53.5	<0.001
Genitals	9.3	21.7	8.2	<0.001
Legs	47.2	57.4	55	<0.001
Median time to surgery in days (IQR)	15.0 (8)	29.1 (10)	13.0 (9)	<0.001
Time to excision, %				<0.001
≤ 7 days	9.3	29.9	15.6	
> 7 days	90.7	70.1	84.4	

Values are presented as median (IQR) and percentage

¥n = 1 missing, ¶n = 19 missing. †More than one location per patient is possible

IQR = interquartile range, TBSA = Total Body Surface Area

*Between the three different groups

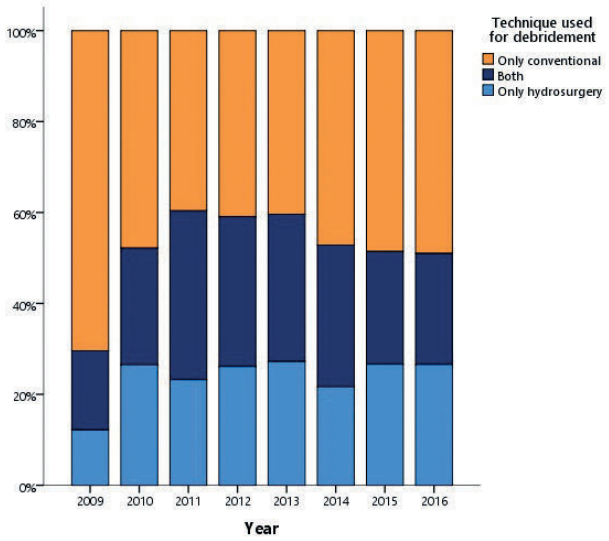


Figure 2. Patient and injury characteristics

The lowest prevalence of patients who received exclusive hydrosurgical debridement was in 2009: 12.2%. Burn severity did not change between 2009 and 2016 (ANOVA, $p = 0.16$).

The median age in the groups in which hydrosurgery was used was lower (29 (IQR 42) years and 28.3 years (IQR 35) vs. 44 years (IQR 35), $p < 0.001$; Table 1). Elderly patients (>65 years) had lower odds of being treated with hydrosurgical debridement compared to all other age categories (univariable analysis; Table 2). There was a trend toward differences in gender ($p < 0.10$; Table 1). Males had a higher likelihood of being treated with hydrosurgical debridement, whether or not in combination with conventional debridement techniques (resp. OR 1.23 95%CI 1.03–1.46, OR 1.23 95%CI (1.03–1.59); Table 2). Scalds were more frequently debrided with hydrosurgery alone (Table 2; OR 2.23 95%CI 1.76–2.83), while contact burns and burns with other causes (e.g. electricity, chemical) were more frequently debrided with conventional excision alone (Table 2, univariable analyses both $p < 0.01$). Median percentage TBSA burned was higher (11.0% (IQR 17.8; Table 1) and time to surgery was longer (29.1 days (IQR 10); Table 1) in the combination group compared to the hydrosurgical (5.0% (IQR 8), 15.0 days (IQR, 8)) and conventional group (2.0% (IQR 5.5), 13.0 days (IQR, 9))). Hydrosurgery was more often used in patients with a higher percentage TBSA burned, although the odds for exclusive hydrosurgical debridement decreased in patients with a TBSA >20% (OR 4.42 95%CI 2.56–.62; Table 2). In addition, patients with a delayed timing of surgery had higher odds of being treated with hydrosurgery alone (OR 1.80 (1.20–2.54); Table 2).

Table 2. Predictors for the use of hydrosurgery for burn wound debridement

	Hydrosurgery used for debridement[e] vs. Only conventional techniques used for debridement		**Only hydrosurgery used for debridement vs. Only conventional techniques used for debridement**	
	Univariable analysis OR (95% CI)	Multivariable analysis OR (95% CI)[c]	Univariable analysis OR (95% CI)	Multivariable analysis OR (95% CI)[d]
Age in categories				
0-4y	3.27 (2.29-4.67)*	2.50 (1.67-3.74)*	6.53 (4.24-10.06)*	4.00 (2.48-6.43)*
5-17y	1.87 (1.33-2.64)*	2.31(1.58-3.39)*	2.85 (1.84-4.42)*	3.29 (2.05-5.29)*
18-65y	1.45 (1.14-1.84)*	1.49 (1.13-1.87)*	1.69 (1.20-2.37)*	1.75 (1.22-2.51)*
>65y	ref.	ref.	ref.	ref.
Gender				
Male	1.23 (1.03-1.46)*		1.23 (1.03-1.59)**	
Aetiology[a]				
Scald	1.38 (1.13-1.69)*	1.45 (1.13-1.87)**	2.23 (1.76-2.83)*	1.80 (1.33-3.21)*
Fire/Flame	1.45 (1.25-1.77)*		0.94 (0.76-1.17)	
Grease	1.04 (0.77-1.40)		1.24 (0.87-1.77)	
Contact	0.28 (0.19-0.39)*		0.32 (0.20-0.50)*	
Other	0.42 (0.30-0.58)*		0.31 (0.19-0.50)*	0.54 (0.32-0.92)*
% TBSA burned				
<1	ref.	ref.	ref.	ref.
1-2	1.75 (1.20-2.57)*	1.77 (1.20-2.62)*	2.04 (1.29-3.24)*	1.98 (1.23-3.21)*
2-5	4.49 (3.18-6.34)*	4.33 (3.03-6.20)*	4.27 (2.79-6.52)*	3.88 (2.48-6.08)*
5-10	6.02 (4.25-8.51)*	4.86 (3.31-7.12)*	5.47 (3.58-8.37)*	3.81 (2.36-6.16)*
10-20	10.57 (7.29-15.34)*	8.55 (5.59-13.08)*	7.23 (4.58-11.42)*	5.44 (3.20-9.23)*
>20	14.39 (9.57-21.63)*	11.21 (6.95-18.09)*	4.42 (2.56-7.62)*	3.50 (1.84-6.64)*
Body location [a,b]				
Head and neck	3.25 (2.69-3.93)*	1.66 (1.31-2.09)*	2.60 (2.08-3.30)*	1.85 (1.38-2.48)*
Trunk	2.12 (1.78-2.53)*	0.71 (0.56-0.91)*	1.40 (1.13-1.74)*	0.57 (0.43-0.77)*
Arms	2.52 (2.10-3.02)*	1.64 (1.32-2.04)*	1.97 (1.57-2.47)*	1.81 (1.34-2.37)*
Genitals	2.13 (1.61-2.80)*		1.41(0.78-1.66)	
Legs	0.92 (0.77-1.09)		0.74 (0.59-0.91)*	
Time to surgery				
> 7 days	0.72 (0.57-0.90)*	1.68 (1.15-2.44)*	1.80 (1.20-2.54)*	

ref = reference group, y = years, TBSA = Total Body Surface Area

*p < 0.01, **p<0.05, [a] Reference group = all others, [b] More than one body location per patients possible, [c] The following variables were included in the multivariable odds: age, gender, scalds, fire/flame burns, contact, other, %TBSA burned, head and neck burns, trunk, arms, genitals and >7 days to surgery, [d] The following variables were included in the multivariable odds: Age, gender, scalds, contact, other, %TBSA burned, head and neck, trunk, arms legs an >7days to surgery, [e] Whether or not in combination with conventional techniques.

Significant independent predictors of the use of hydrosurgery were a younger age, scalds, a larger TBSA burned, head/neck burns and arm burns (multivariable analyses; Table 2).

Prevalence of hydrosurgery for burn wound debridement per affected body region

In patients who were only burned in one body region hydrosurgery was most frequently exclusively used for debridement of the neck (58.3%), followed by the scalp (31.6%) and genitals (31.6%) (Table 3).

Table 3. Details on body region debrided with hydrosurgery[a]

Body region[b]	Only hydrosurgical debridement	Total	%
Head and neck	14	43	32.6
Scalp	6	19	31.6
Face	9	33	27.3
Neck	7	12	58.3
Trunk	14	99	14.1
Ventral	12	83	15.7
Dorsal	4	26	15.4
Upper extremity	103	369	26.0
Arm	55	213	25.8
Hand	70	252	27.8
Genitals	11	40	27.5
Genital area	6	19	31.6
Buttocks	6	27	22.2
Lower extremity	135	675	20.0
Legs	96	486	19.8
Feet	63	279	22.6

[a]The burn centre registration allowed the registration of multiple burn locations per patient and does not differentiate between conservative, conventional and hydrosurgically treated body locations. Therefore, a subgroup analysis of patients with only one body part burned was performed to identify the prevalence of the use of hydrosurgery per body region.

[b] More than one subcategory per body region is possible.

Surgical treatment and clinical outcome

The TBSA excised was higher in both groups in which hydrosurgery was used ($p < 0.001$ Table 4). Patients in the group exclusively treated with hydrosurgery were less often treated with dermal substitutes. Also, they underwent less surgical procedures and had a lower mean volume of blood transfusion.

In the groups of patients in which hydrosurgery was used, whether or not in combination with conventional techniques median length of stay were higher. In the group of patients in which

hydrosurgery was used exclusively, wound infection rates were lower compared to the other groups.

Table 4. Surgical treatment and clinical outcome

	Hydrosurgical H	**Both B**	**Conventional C**	**p-value* Overall**	**p-value** H vs C**
n (%)	506 (23.9)	599 (28.3)	1008 (47.7)		
Surgical management					
Median TBSA Excised (IQR)[a]	2.0 (3.0)	5.0 (11.0)	1.0 (2.5)	<0.001	<0.001
Grafting technique (%)[b]					
SSG	95.5	96.0	93.2	0.031	0.077
MEEK	1.6	19.4	3.1	<0.001	0.083
Homograft	1.8	11.5	3.3	<0.001	0.095
Dermal substitute	0.2	2.2	1.5	0.018	0.021
Mean number of surgical procedures (range, SD)[c†]	1.2 (1-12, 0.8)	2.8 (1-22, 3.1)	1.4 (1-11,1.1)	<0.001	0.019
Mean volume of blood transfusion in ml (range, SD)[d†]	57.2 (0-4400, 361)	821.2 (0-32625, 2480)	156.0 (0-1485, 870)	<0.001	0.036
Clinical outcome					
Re-admission (%)	22.9	26.0	20.3	0.030	0.245
Median length of stay (IQR)	17.0 (16.0)	27.0 (27)	8.0 (20.0)	<0.001	<0.001
Wound infection (%)	1.6	6.7	3.8	<0.001	0.019
Reconstructions (%)	4.7	18.0	5.3	<0.001	0.667

Values are presented as median (IQR) or percentage

† Presentation of range and SD to improve interpretability

SSG = split skin graft, MEEK = Meek micrografting, IQR = inter quartile range, SD = standard deviation

[a]188 missing

[b]More than one surgical technique per patient possible

[c]23 missing

[d]307 missing

DISCUSSION

This multi-centre study appears to be the first evaluation of the application of hydrosurgery for tangential excision in burns in a large cohort over multiple years.

The aim of this study was to gain insight in which patients hydrosurgery is used and whether the actual field of hydrosurgical application matches the current available literature. Our data show that the use of hydrosurgery is substantial, as it has been used, also in combination with conventional debridement techniques, in more than fifty percent of the patients undergoing

tangential excision since 2010. Hydrosurgical excision is described to be specifically useful for the debridement of superficial and deep dermal burns since full thickness burns are not as easily debrided hydrosurgically[2,7]. Therefore, conventional techniques have to be used next to hydrosurgery for sufficient debridement of burn wounds with mixed depths[1,13].

Our study identified a younger age, scalds, a higher percentage TBSA burned, head/neck burns and arm burns (including hands) as independent predictors for the use of hydrosurgery. Our finding that scalds are predictors for the use of hydrosurgery might be a reflection of its use in superficial burns, as scalding is known to result in more superficial burns[14]. Next to that, our results showed that burns with other causes (e.g. electricity, chemical) had lower chances of being treated hydrosurgically, whereas these causes are known to result in deeper burns[14,15]. The exclusive use of hydrosurgery decreased in patients with extensive burns (TBSA >20%), which might be explained by the fact that patients with extensive burns have a higher chance of burn wounds with mixed depths.

In our study population, patients in the age category 0–4 years were more often treated with hydrosurgery. In these young children, scalds are the most common type of burn injury[16]. A younger age and scalds remained as independent predictors for the use of hydrosurgery in the multivariable analysis. Therefore, the high prevalence of scald injuries is not the only explanation for the more frequent use of hydrosurgery in young children. Conventional tangential knife excision is described to have a tendency to remove more viable tissue than is actually necessary for adequate debridement [6,12,17]. Our results that toddlers and infants had the highest chance of being treated with hydrosurgical excision may reflect the wish for a more precise debridement in the paediatric burns population to maximize preservation of viable dermis. Next to improvement of scar quality and scar contraction, this could potentially lead to a decrease in hypertrophic scarring, which is in fact a significant problem in children[18,19]. Another potential benefit of hydrosurgical debridement is that the small Versajet hand piece allows irregularly contoured and relatively inaccessible areas to be easily reached[1,7]. This is in line with our results that hydrosurgery was more often used for debridement of irregular contoured locations as the head and arms, and less in large flat body parts as the trunk. Because surgery characteristics are not linked to specific body locations in the R3 database we performed a subanalysis in patients with only one body part burned. This analysis also showed that the scalp, neck and genital area were more often treated with hydrosurgery alone.

Although we expected that hydrosurgery would be more often used in patients with smaller burns, we found that patients with a percentage TBSA beneath one percent were more often treated with conventional excision techniques only and that the median percentage TBSA excised was higher in the groups in which hydrosurgery was used. This might be explained by the higher costs of the Versajet™ compared with the costs of conventional excision techniques. The current

cost of the disposable Versajet™ II headpiece is €141,86 ($167.55) whilst the costs of a re-usable guard and handle of the Weck knife are respectively €0,91 ($1.08) and €20,91 ($24.70). The cost of one sterilized, single use Weck blade is €1,08 ($1.28). In our experience, burn specialist prefer to use a Weck knife in smaller burns to reduce treatment costs.

In current study, the mean volume of blood transfusion was lower in the group that was exclusively treated with hydrosurgery than in patients treated with conventional excision, even though the median TBSA excised was higher in the hydrosurgical treated group. Next to a possible more subtle debridement using the Versajet™ system, this might be the results of delayed debridement undertaken in the hydrosurgery group. To our knowledge, no other study compared the amount of blood loss in hydrosurgical and conventional treated burn patients. However, in wounds with delayed healing, maximum blood loss was found to be less in the hydrosurgical debridement group compared to the conventional debridement group in one clinical trial[9]. Our results also show that the prevalence of wound infection was significant lower in the group exclusively debrided with hydrosurgery compared to the conventional only debridement group. A few studies on chronic wounds have reported that hydrosurgery may decrease bacterial burden after debridement and therefore post-operative infection, but this was not confirmed by randomized trials in burn patients[7,10-12,20]. The differences in surgical management and clinical outcome might be explained by the possibility that the wounds that were treated with hydrosurgery alone were more likely to be superficial. This is supported by the lower use of dermal substitutes in this group. Unfortunately we were not able to adjust our results for burn depth.

Some shortcomings of our study have to be mentioned. As it is a retrospective study, data were not collected for the specific purpose of this study and was lacking in details on wound and surgery characteristics. As a result, we were not able to perform a multivariable analysis on the clinical outcomes and more prospective research is necessary to support the outcomes of our study. Nevertheless, the Dutch Burn Repository R3 database is closely linked to medical registers in three dedicated burn centres and study groups were large. Therefore, this database gives a unique picture of the use of hydrosurgery in burn care with comprehensive and generalizable data. The Dutch Burn Repository R3 registers burn depth estimated at admission and percentage TBSA excised during surgery. We were not able to conclusively conclude that hydrosurgery was the preferred debridement tool for deep dermal burns instead of full thickness burns during surgery, nor correct our outcomes for burn depth. Another shortcoming is the lack of long-term results. Although scar quality is considered to be one of the most important outcomes of burn surgery today, no clinical study compared the effect of hydrosurgical debridement on scarring in the long term. Hyland et al. performed a randomized trial in the paediatric burn population comparing hydrosurgery with conventional debridement. They did not observe significant differences in scarring at 3 or 6 months after injury measured with the Vancouver scar scale (VSS)[12]. Nevertheless, the follow-up duration of 6 months may not be adequate for scar quality

assessment and the VSS was formally not designed to assess burn scar severity, has a moderate reliability, and does not include the opinion of the patient[21-23]. Only one study showed a superior result after hydrosurgery was used for burn wound debridement[4]. Unfortunately, data of this retrospective study were not published. Hence, it remains unclear if hydrosurgical debridement results in better functional and cosmetic scar outcomes.

In conclusion, this study provides evidence that the use of the Versajet™ hydrosurgery system for burn wound debridement prior to skin grafting is substantial. In the three Dutch burn centres, it is often used in combination with sharp conventional tangential debridement with knives. Individual predictors for its use are a younger age, scalds, higher TBSA burned, and burns on convex locations.

Our study group currently performs a randomized trial to compare scar quality after hydrosurgical and conventional tangential excision, to optimize burn outcomes in the future and to provide new perspectives on the benefits of hydrosurgical debridement in burn surgery (Netherlands trial registry: NTR 6232)[24].

Conflict of interest statement

We declare that there is no conflict of interest including any financial, personal or other relationship.

Acknowledgements

We would like to thank the Dutch Burns Foundation Beverwijk, Red Cross Hospital Beverwijk; Martini Hospital Groningen, and Maasstad Hospital Rotterdam for their support.

Appendix A

The 'Dutch Burn Repository Group' consists of:
Burn Centre Beverwijk: E.C. Kuijper, A. Pijpe, D. Roodbergen, A.F.P.M. Vloemans, P.P.M. van Zuijlen.
Burn Centre Rotterdam: J. Dokter, A. van Es, C.H. van der Vlies.
Burn Centre Groningen: G.I.J.M. Beerthuizen, J. Eshuis, J. Hiddingh, S.M.H.J. Scholten-Jaegers.
Association of Dutch Burn centres: M.E. van Baar, E. Middelkoop, M.K. Nieuwenhuis, A. Novin.

REFERENCES

1. Klein MB, Hunter S, Heimbach DM, et al. The versajet water dissector: A new tool for tangential excision. *J Burn Care Rehabil*. 2005;26(6):483-487.
2. National institute for health and clinical excellence (2014) the versajet II hydrosurgery system for surgical debridement of acute and chronic wounds and burns NICE guideline (MIB1). .
3. Barret JP. Burn wound excision with hydrosurgery: Initial experience. *Injury Extra*. 2006;37(5):187-189.
4. Scholten SMHJ, Bosch van den M, Nieuwenhuis M, Niemeijer A, Beerthuizen GIJM. O27.3 scar quality after surgical treatment of deep dermal burns with hydrosurgery compared to quarded knife; short and long term outcome. *Burns: Supplement 1*. 2011;37(Supplement 1):S18-S18.
5. Schwartz, J.A. MD,Tissue healing and wound repair fellow, Goss SG, Facchin F, Avdagic E. Surgical debridement alone does not adequately reduce planktonic bioburden in chronic lower extremity wounds. *J Wound Care*. 2014;23(Sup9):S4-S13.
6. Esposito G, Anniboletti T, Palombo M, Palombo P. Versajet hydrosurgery: Our experience in adults and pediatric patients. *Burns*. 2009;35:23-Supl 1.
7. Gravante G, Delogu D, Esposito G, Montone A. Versajet hydrosurgery versus classic escharectomy for burn debridment: A prospective randomized trial. *J Burn Care Res*. 2007;28(5):720-724.
8. Caputo WJ, Beggs DJ, DeFede JL, Simm L, Dharma H. A prospective randomised controlled clinical trial comparing hydrosurgery debridement with conventional surgical debridement in lower extremity ulcers. *Int Wound J*. 2008;5(2):288-294.
9. Liu J, Ko JH, Secretov E, et al. Comparing the hydrosurgery system to conventional debridement techniques for the treatment of delayed healing wounds: A prospective, randomised clinical trial to investigate clinical efficacy and cost-effectiveness. *Int Wound J*. 2015;12(4):456-461.
10. Granick MS, Tenenhaus M, Knox KR, Ulm JP. Comparison of wound irrigation and tangential hydrodissection in bacterial clearance of contaminated wounds: Results of a randomized, controlled clinical study. *Ostomy Wound Manage*. 2007;53(4):64-6, 68-70, 72.
11. Mosti G, Iabichella ML, Picerni P, Magliaro A, Mattaliano V. The debridement of hard to heal leg ulcers by means of a new device based on fluidjet technology. *Int Wound J*. 2005;2(4):307-314.
12. Hyland EJ, D'Cruz R, Menon S, et al. Prospective, randomised controlled trial comparing versajet hydrosurgery and conventional debridement of partial thickness paediatric burns. *Burns*. 2015;41(4):700-707.
13. Cubison TC, Pape SA, Jeffery SL. Dermal preservation using the versajet hydrosurgery system for debridement of paediatric burns. *Burns*. 2006;32(6):714-720.
14. Brans TA, Dutrieux RP, Hoekstra MJ, Kreis RW, du Pont JS. Histopathological evaluation of scalds and contact burns in the pig model. *Burns*. 1994;20 Suppl 1:S48-51.

15. Hettiaratchy S, Dziewulski P. ABC of burns: Pathophysiology and types of burns. *BMJ*. 2004;328(7453):1427-1429.
16. Baptiste MS, Feck G. Preventing tap water burns. *Am J Public Health*. 1980;70(7):727-729.
17. Anniboletti T, Palombo M, Fasciani L, Delli Santi G, Palombo P. O27.6 the use of versajet hydrosurgery: 5 years experience. *Burns: Supplement 1*. 2011;37(Supplement 1):S19-S19.
18. Harrison CA, MacNeil S. The mechanism of skin graft contraction: An update on current research and potential future therapies. *Burns*. 2008;34(2):153-163.
19. Cubison TC, Pape SA, Parkhouse N. Evidence for the link between healing time and the development of hypertrophic scars (HTS) in paediatric burns due to scald injury. *Burns*. 2006;32(8):992-999.
20. Fraccalvieri M, Serra R, Ruka E, et al. Surgical debridement with VERSAJET: An analysis of bacteria load of the wound bed pre- and post-treatment and skin graft taken. A preliminary pilot study. *Int Wound J*. 2011;8(2):155-161.
21. van der Wal MB, Vloemans JF, Tuinebreijer WE, et al. Outcome after burns: An observational study on burn scar maturation and predictors for severe scarring. *Wound Repair Regen*. 2012;20(5):676-687.
22. Gangemi EN, Gregori D, Berchialla P, et al. Epidemiology and risk factors for pathologic scarring after burn wounds. *Arch Facial Plast Surg*. 2008;10(2):93-102.
23. Sullivan T, Smith J, Kermode J, McIver E, Courtemanche DJ. Rating the burn scar. *J Burn Care Rehabil*. 1990;11(3):256-260.
24. Legemate CM, Goei H, Middelkoop E, et al.Long-term scar quality after hydrosurgical versus conventional debridement of deep dermal burns (HyCon trial): study protocol for a randomized controlled trial. *Trials*. 2018 Apr 19;19(1):239.

8

Long-term scar quality after hydrosurgical versus conventional debridement of deep dermal burns (HyCon trial): study protocol for a randomized controlled trial

C.M. Legemate, H. Goei, E. Middelkoop, I.M.M.H. Oen, T.H.J. Nijhuis, K.A.A. Kwa, P.P.M. van Zuijlen, G.I.J.M. Beerthuizen, M.K. Nieuwenhuis, M.E. van Baar, C.H. van der Vlies

Trials. 2018 Apr 19;19(1):239

ABSTRACT

Background

Deep dermal burns require tangential excision of non-viable tissue and skin grafting to improve wound healing and burn-scar quality. Tangential excision is conventionally performed with a knife, but during the last decade hydrosurgery has become popular as a new tool for tangential excision. Hydrosurgery is generally thought to be a more precise and controlled manner of burn debridement leading to preservation of viable tissue and, therefore, better scar quality. Although scar quality is considered to be one of the most important outcomes in burn surgery today, no randomized controlled study has compared the effect of these two common treatment modalities with scar quality as a primary outcome. The aim of this study is, therefore, to compare long-term scar quality after hydrosurgical versus conventional tangential excision in deep dermal burns.

Methods/design

A multicenter, randomized, intra-patient, controlled trial will be conducted in the Dutch burn centers of Rotterdam, Beverwijk, and Groningen. All patients with deep dermal burns that require excision and grafting are eligible. Exclusion criteria are: a burn wound < 50 cm^2, total body surface area (TBSA) burned $> 30\%$, full-thickness burns, chemical or electrical burns, infected wounds (clinical symptoms in combination with positive wound swabs), insufficient knowledge of the Dutch or English language, patients that are unlikely to comply with requirements of the study protocol and follow-up, and patients who are (temporarily) incompetent because of sedation and/or intubation. A total of 137 patients will be included. Comparable wound areas A and B will be appointed, randomized and either excised conventionally with a knife or with the hydrosurgery system. The primary outcome is scar quality measured by the observer score of the Patient and Observer Scar Assessment Scale (POSAS); a subjective scar-assessment instrument, consisting of two separate six-item scales (observer and patient) that are both scored on a 10-point rating scale.

Discussion

This study will contribute to the optimal surgical treatment of patients with deep dermal burn wounds.

Trial registration

Dutch Trial Register, NTR6232. Registered on 23 January 2017.

BACKGROUND

Surgical debridement is an important step in the treatment of patients with deep dermal burns. The purpose is to remove necrotic and infectious materials and to prepare tissue for skin grafting and definitive wound closure[1]. Conventional surgical debridement of acute burn wounds consists of sharp tangential excision of non-viable tissue with hand-held knives such as the Goulian or Weck knife[2]. Adequate debridement with these knives is determined by the presence of punctuate bleeding and viable dermis. This procedure is not only associated with substantial blood loss, but also with the unnecessary removal of viable dermis[2,3]. Loss of dermis has been considered one of the main factors determining the quality of the scar and the degree of contraction of the healing wound[4-6]. Therefore, methods which maximally preserve dermis are essential. During the last decade, hydrosurgery has become popular in burn surgery as a new option for excision of non-viable tissue prior to skin grafting[7-9]. The Versajet™ hydrosurgery system (Smith and Nephew, St. Petersburg, FL, USA) was developed in 1997 for the purpose of debriding various types of wounds, including burn wounds, and is superseded by the Versajet II™ (Smith and Nephew) in 2011[8]. The Versajet II™ system works by producing a high-pressure jet of water across an aperture in an angled handpiece. The Venturi effect creates a vacuum that removes surface debris, which is sucked into the machine together with the irrigation fluid. The cutting and aspiration effects can be controlled by adjusting console power settings, handpiece orientation, and handpiece pressure. The vacuum that is created by the speed of the jet aims to lift only non-viable tissue and thus maximal dermal preservation. For this reason, hydrosurgical debridement of burns might lead to a better scar outcome compared to conventional sharp debridement.

Although burn specialists widely use hydosurgery as an alternative for conventional tangential debridement[6,7] only a limited number of studies is available on the effects of hydrosurgery in burn patients[10-12]. A guideline from the National Institute for Health and Care Excellence (NICE) recently reported that the Versajet™ is an efficient and safe wound debridement tool in both adults and children with acute and chronic wounds[8]. Up to now, two randomized controlled trials comparing hydrosurgical and conventional debridement in patients with burns have been published[13,14].

Gravante et al. described that adequate debridement of the wound bed was possible in all patients treated with the Versajet™ system[13]. The authors suggested that hydrosurgical excision was more precise in obtaining the correct dermal plane, but did not confirm this with objective measurements. Hyland et al. studied children under the age of 16 and histologically confirmed that significantly more viable dermis was preserved in the group of patients treated with hydrosurgery compared to the conventionally treated group of patients[14]. However, they did not observe significant differences in scar quality measured with the Vancouver scar scale (VSS) at

3 and 6 months post burn. Furthermore, they did not use any objective scar measurement tools and the study was limited by a relatively short follow-up period as scars mature over a period of at least one year[15,16]. Also, the VSS was formally not designed to indicate burn scar severity, has a moderate reliability and does not include the opinion of the patient[17]. Hence, it remains unclear whether hydrosurgery for the routine debridement of deep dermal burns prior to skin grafting leads to increased dermal preservation and better scar quality outcomes.

The aim of this study therefore, is to assess the effectiveness of hydrosurgical compared to conventional debridement in deep dermal burns. Long-term scar quality after hydrosurgical and conventional debridement of deep dermal burns in relation to histologically measured dermal preservation will be analyzed.

METHODS

Protocol and registration

The study was approved by the medical ethics committee (NL58875.101.16) and by the institutional review boards of each participating burn center. The methods applied were specified in advance, documented in a protocol, and registered (http://www.trialregister.nl, NTR6232). The protocol has been designed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines for interventional trials[18]. The SPIRIT checklist and figure are given in Additional file 1 and Figure 2, respectively.

Study design

A multicentre, randomized controlled trial with an intra-patient comparison of hydrosurgical and conventional debridement will be conducted in the three Dutch burn centers: Rotterdam, Beverwijk, and Groningen. As the healing process of burn wounds and scar formation differs between patients we chose an intra-patient design to provide representative outcomes and to limit inter-patient bias.

Participants

Patients of all ages with deep dermal burns and an indication for tangential excision and skin grafting are eligible for this trial, either hospitalized or under treatment in the outpatient clinic of one of the participating burns centers. Exclusion criteria are as follows: a burn wound $<50cm^2$, total body surface area (TBSA) burned $>$ 30%, full thickness burns, chemical or electrical burns, infected wounds (clinical symptoms in combination with positive wound swabs), patients with insufficient knowledge of the Dutch or English language, patients that are unlikely to comply with requirement of the study protocol and follow-up, and patients that are (temporary) incompetent because of sedation and/or intubation.

Patients are included after full, understandable and neutral explanation of the study by a member of the research team and after having given written informed consent.

Interventions

Conventional tangential excision

Tangential excision with guarded knives relies on the stepwise excision of a layer of tissue using a flat blade. The addition of a guard prevents the removal of excessive amounts of tissue, and most of these knives allow adjustment of the width of the gap between the blade and the guard. However, if the gap is too narrow the instrument will glide off the burn without any debridement taking place[2].

Hydrosurgical tangential excision

The Versajet™ II hydrosurgery system (Smith and Nephew. St. Petersburg, FL, USA) was CE marked in 2011 and was launched in 2012[8]. It uses a high-pressure jet of sterile saline to debride wounds. It is attached to a console, which is then operated by a foot pedal. The saline is forced out of a narrow nozzle and functions like a knife which allows debridement and aspiration of debris to occur simultaneously. Pressure can be adjusted (power setting 1-10) to facilitate the desired depth of debridement. As a result, the correct level might be reached more accurately, preserving as much dermis as possible. Hydrosurgery is preferentially suited to debride softer necrotic tissues, and cannot be used to debride full thickness burns as it doesn't cut through hard eschar. Versajet™ is reported to be used routinely in multiple centers around the world these days[8]. Nevertheless, a clear algorithm for its use is lacking, and burn specialists may choose individually whether hydrosurgery can be applied or not[9].

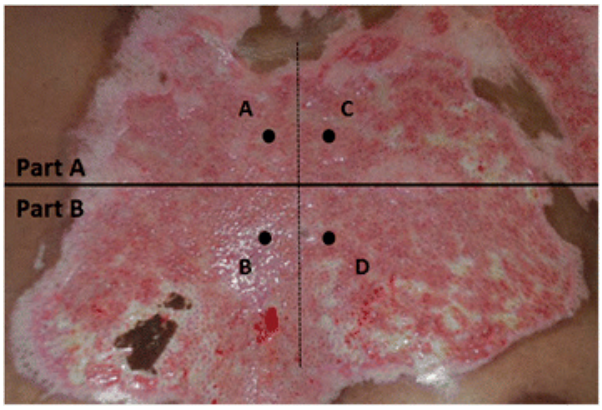


Figure 1. Location of punch biopsies.

A - Biopsy part A, before debridement. B - Biopsy part B, before debridement. C - Biopsy part A, after debridement. D - Biopsy part B, after debridement

Surgical procedure

Prior to surgery, the surgeon divides the study area into two adjacent parts of equal size and burn depth (part A and part B). These parts are randomly allocated to either conventional or hydrosurgical (Versajet™) tangential debridement. Two 3-mm punch biopsies of both intervention areas will be collected before and after debridement, according to a standardized method (Figure 1), to determine the amount of viable dermis before and after excision. Type of mesh graft and expansion will also be standardized per patient to ensure an equal mesh cover of the two intervention areas. Before and after surgery, standard wound care is given. After discharge, patients will be treated in an outpatient setting according to the local protocol.

Study outcomes

Primary outcome measure

Our primary outcome measure is scar quality assessed by the items of the observer scale of the Patient and Observer Scar Assessment Scale (POSAS) at 12 months post-surgery. The POSAS is recognized as a highly reliable scar rating scale, and consists of two numeric scales: the Patient Scar Assessment Scale (patient scale) and the Observer Scar Assessment Scale (observer scale)[19,20]. The observer scale contains the items vascularization, pliability, pigmentation, thickness, and relief. All items will be measured for part A and B of the study wound on a 10-point rating scale by two experienced, trained and independent observers to improve the reliability of the assessment. The average of the observers' scores will range from 1, which corresponds to the situation of normal skin, to 10, indicating the 'worst' imaginable scar.

Secondary outcome measures

Subjective scar assessment

Scar quality of part A and B of the study area will be measured at 3, 6 and 12 months post-surgery using the POSAS. Although the total score of the observer scale at 12 months is our primary outcome, the items of the patient scale of the POSAS will also be measured and analyzed separately. The patient scale of the POSAS contains the items color, pliability, thickness, relief, itching, and pain. These items will be scored on a 10-point rating scale and added to form the total patient score. In addition, patients and observers will score their overall opinion on the scar (1-10, numeric scale), and total scores of the patient and observers will be added to form a total score.

Scar elasticity

Scar elasticity will be measured with the Cutometer® (Courage-Khazaka electronic GmbH Cologne, Germany). The Cutometer® is a validated instrument to measures the viscoelasticity of the skin by analyzing its maximal extension (Uf in mm) in response to negative pressure[21].

Scar color and pigmentation

Scar color and pigmentation will be measured with the Dermaspectrometer® (Cortex Technology ApS Hadsund, Denmark), which is a reliable narrowband spectrometer that computes an erythema and melanin index[19].

Measurements with the Cutometer® and Dermaspectrometer® are performed at 3 and 12 months post burn on both parts (A and B) of the study area, and adjacent normal skin. For objective data collection, measurements will be performed following a fixed protocol.

Dermal preservation

During surgery, two punch biopsies will be taken out of both parts (A and B) of the study area, pre- and post-debridement, using a 3mm punch. The biopsies will be fixed in kryofix and processed into 3-5-μm histological slides. Sections will then be stained with hematoxylin and eosin (H&E) or picosirrius red. To determine the amount of dermal preservation and hence precision of debridement, the amount of viable tissue on pre- and post-debridement specimens will be recorded using light microscopy.

Baseline characteristics

Data registration will start on the day of randomization. Data regarding patients' baseline characteristics will be obtained from patients' medical records:

- Demographics: age, gender
- Burn characteristics: % total body surface area (TBSA) burned, affected anatomical site(s), wound etiology, time to surgery and burn depth of the study area. If possible, burn depth will accurately be determined on day 2–5 post burn by clinical evaluation and Laser Doppler Imaging (LDI) scan using the moorLDI2-Burn Imager™ (Moor Instruments, Axminster, UK) or similar[22].
- Clinical characteristics: skin type according to the Fitzpatrick skin type scale; wound healing time (measured in days till 95% re-epithelization); comorbidity; Weck knife, Versajet™ and dermatome settings; expansion of skin graft; adverse events (graft loss, wound infection) and need for reconstructive surgery.

	STUDY PERIOD					
	Enrolment	Allocation	Post-allocation			
TIMEPOINT	*Baseline (pre-surgery)*	Day of surgery	*Post-operative*	*3 months*	*6 months*	*12 months*
ENROLMENT:						
Eligibility screen	X					
Informed consent	X					
Allocation		X				
INTERVENTIONS:						
Conventional debridement		X				
Hydrosurgical debridement		X				
Biopsies		X				
Usual care	←	—	—	—	—	→
ASSESSMENTS:						
POSAS				X	X	X
Scar elasticity				X		X
Scar color and pigmentation				X		X

Figure 2. Schedule of enrolments, interventions and assessments

SAMPLE SIZE

Power calculation is based on the results obtained by an unpublished retrospective study on scar quality after hydrosurgery versus guarded knife excision in the Martini hospital in Groningen[9]. The primary outcome measure was scar quality assessed from photographs, and expressed in the total score of the observer part of the POSAS. Because scar quality was assessed from photographs, pliability was not taken into account[23]. Therefore, scar assessment contained the four items vascularization, pigmentation, thickness, and relief.

The lowest sum score, reflecting normal skin, was four and the highest score, reflecting the worst imaginable scar, was forty. In this study, the Observer Score of the POSAS questionnaire 12 months post-surgery was 14.7 in the hydrosurgery groups versus 16.7, with a pooled SD of 6.53. This results in an effect size of 0.3. Because of the within-subject design, a correction for correlated samples was included, assuming a correlation of 0.4 between POSAS Observer Score within one patient. Given a power of 0.8 and a level of significance of 0.05 a number of 105 patients is needed. Assuming a 30% drop out, 137 patients need to be recruited.

RANDOMIZATION

Randomization will occur in the operating theater after the wound is divided into two comparable study areas, defined as part A and part B. These areas are randomly assigned to receive either hydrosurgical or conventional debridement. Allocation of the treatment will be stratified per institute in blocks using the online randomisation program CASTOR, https://data.castoredc.com. The outcome will be displayed on the website, only visible for the person who performed the randomisation and the principal investigator. After randomization, the central trial coordinator will receive instructions with the inclusion number.

BLINDING

Patients are blinded as they are sedated during surgery and will not be aware which treatment they received on which part of the wound. Blinding surgical treatment is not possible, as the burn surgeon knows which part of the wound received which surgical treatment. Outcome assessment is blinded as the member of the research team who performs the follow-up measurements is unaware of the technique used for debridement of part A or part B. In case of randomisation related difficulties, the central trial coordinator can be contacted.

STATISTICAL ANALYSIS

Data analysis will be performed using SPSS PASW Statistics 23.0 (IBM, New York City).

Primary outcome

Differences in scar quality 12 months post-surgery assessed as the total score of the observer scale of the POSAS will be analyzed using the paired student-t test (in case of normal distribution) or the Wilcoxon signed-rank test (non-normal distribution).

Secondary outcomes

Differences in 12-month outcomes of the patient scar assessment, the observer scar assessment, scar elasticity (measured by the Cutometer®), scar color (measured by the DermaSpectrometer®) and differences in viable dermis after excision (measured by histopathology) will be analyzed using the Wilcoxon signed-rank test in case of non-normal distribution, or paired student t-test in case of normal distribution. Because of repeated measurements within patients and loss to follow-up, overall differences of scar quality measurements will be analyzed using generalized estimating equations, with unstructured working correlation matrix structure.

DISCUSSION

In this paper, we described the design of our study into long-term scar quality after hydrosurgical and conventional tangential excision of deep dermal burns. This will be the first study that assesses differences in scar quality between both treatment modalities at 12 months post-surgery in both adults and children with subjective and objective measurement tools.

Subjective scar quality will be assessed using the POSAS. The POSAS is unique as it takes the opinion of the patients into account which is mandatory for a clinically relevant scar evaluation[24].

Scar quality will not only be measured subjectively, but also with objective measurement tools concerning scar pigmentation, vascularity and pliability. Aside these evaluations, we want to support our results via quantitative analysis of the histological specimens. For a reliable follow-up, documentation of which area of the wound is part A, and which area is part B needs to be specific, to allow accurate assessment of the correct areas, as it is possible that there might be no differences visible at follow-up.

In this study, an accurate diagnosis of wound depth is essential to determine the indication for surgery. Therefore, all three burn centers are in possession of an LDI scan to assess burn wound depth, which has an accuracy of 95% in combination with a clinical evaluation of the wound[22,25,26]. Moreover, it can be used to make sure that part A and part B of the study area are of equal burn depth. To enhance the applicability and generalizability of this trial, we chose a multicentre trial design and will recruit patients treated in one of the three national Dutch burn centres. To increase generalizability, and because of the intra-patient design, we are forbearing regarding local clinical care; for example, timing of surgery and type of wound dressings. This study will contribute to the optimal surgical treatment of patients with deep dermal burn wounds and the results will be of high international value, as hydrosurgery is used worldwide.

TRIAL STATUS

This manuscript is a restructured and edited version of the REC approved protocol (version 3.2, 6 February 2017) to comply with the SPIRIT guidelines. Recruitment opened January 10, 2017, and is expected to be completed in January 2019. As of 16 April 2018, 56 patients had been recruited.

Additional file

Additional file 1: SPIRIT checklist (available online: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-018-2599-2#Sec21)

List of abbreviations

LDI: Laser Doppler Imaging, NICE: National Institute for Health and Care Excellence, POSAS: Patient and Observer Scar Assessment Scale, REC: Research
Ethics Committee, SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials, TBSA: Total Body Surface Area, VSS: Vancouver Scar Scale.

Ethics approval and consent to participate

This study has been approved by the Medical Research Ethics Committee of Rotterdam (NL58875.101.16).

Availability of data and material

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This project is funded by the Dutch Burn Foundation (project number 15.101).

Acknowledgements

The authors would like to thank N. Trommel, registered clinical research nurse, Burn Centre of the Maasstad Hospital Rotterdam, and J. Hiddingh, Msc, registered clinical research nurse, Burn Centre of the Martini Hospital Groningen, for their critical review of the research procedures

REFERENCES

1. Orgill DP. Excision and skin grafting of thermal burns. *N Engl J Med*. 2009;360(9):893-901.
2. Jeffery SL. Device related tangential excision in burns. *Injury*. 2007;38 Suppl 5:S35-8.
3. Gurfinkel R, Rosenberg L, Cohen S, et al. Histological assessment of tangentially excised burn eschars. *Can J Plast Surg*. 2010;18(3):e33-6.
4. Matsumura H, Nozaki M, Watanabe K, et al. The estimation of tissue loss during tangential hydrosurgical debridement. *Ann Plast Surg*. 2012;69(5):521-525.
5. Mensik I, Lamme EN, Brychta P. Depth of the graft bed influences split-skin graft contraction. *Acta Chir Plast*. 2003;45(3):105-108.
6. Cubison TC, Pape SA, Jeffery SL. Dermal preservation using the versajet hydrosurgery system for debridement of paediatric burns. *Burns*. 2006;32(6):714-720.
7. Klein MB, Hunter S, Heimbach DM, et al. The versajet water dissector: A new tool for tangential excision. *J Burn Care Rehabil*. 2005;26(6):483-487.
8. NICE. the versajet II hydrosurgery system for surgical debridement of acute and chronic wounds and burns. NICE medtech innovation briefing: National institute for health care excellence; . 2014: 1-23.
9. Scholten-Jaeger S, Bosch van den M, Nieuwenhuis M. Scar quality after surgical treatment of deep dermal burns with hydro-surgery compared to quarded knife, short and long-term outcome. . ;Poster Burns: 37S:S18. 2011.
10. Kimble RM, Mott J, Joethy J. Versajet hydrosurgery system for the debridement of paediatric burns. *Burns*. 2008;34(2):297-8; author reply 299.
11. Rennekampff HO, Schaller HE, Wisser D, Tenenhaus M. Debridement of burn wounds with a water jet surgical tool. *Burns*. 2006;32(1):64-69.
12. Tenenhaus M, Bhavsar D, Rennekampff HO. Treatment of deep partial thickness and indeterminate depth facial burn wounds with water-jet debridement and a biosynthetic dressing. *Injury*. 2007;38 Suppl 5:S39-45.
13. Gravante G, Delogu D, Esposito G, Montone A. Versajet hydrosurgery versus classic escharectomy for burn debridment: A prospective randomized trial. *J Burn Care Res*. 2007;28(5):720-724.
14. Hyland EJ, D'Cruz R, Menon S, et al. Prospective, randomised controlled trial comparing versajet hydrosurgery and conventional debridement of partial thickness paediatric burns. *Burns*. 2015;41(4):700-707.
15. van der Wal MB, Vloemans JF, Tuinebreijer WE, et al. Outcome after burns: An observational study on burn scar maturation and predictors for severe scarring. *Wound Repair Regen*. 2012;20(5):676-687.
16. Gangemi EN, Gregori D, Berchialla P, et al. Epidemiology and risk factors for pathologic scarring after burn wounds. *Arch Facial Plast Surg*. 2008;10(2):93-102.
17. Sullivan T, Smith J, Kermode J, McIver E, Courtemanche DJ. Rating the burn scar. *J Burn Care Rehabil*. 1990;11(3):256-260.
18. Chan AW, Tetzlaff JM, Gotzsche PC, et al. SPIRIT 2013 explanation and elaboration:

Guidance for protocols of clinical trials. *BMJ*. 2013;346:e7586.

19. Draaijers LJ, Tempelman FR, Botman YA, Kreis RW, Middelkoop E, van Zuijlen PP. Colour evaluation in scars: Tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? *Burns*. 2004;30(2):103-107.

20. van der Wal MB, Tuinebreijer WE, Bloemen MC, Verhaegen PD, Middelkoop E, van Zuijlen PP. Rasch analysis of the patient and observer scar assessment scale (POSAS) in burn scars. *Qual Life Res*. 2012;21(1):13-23.

21. Draaijers LJ, Botman YA, Tempelman FR, Kreis RW, Middelkoop E, van Zuijlen PP. Skin elasticity meter or subjective evaluation in scars: A reliability assessment. *Burns*. 2004;30(2):109-114.

22. Hoeksema H, Van de Sijpe K, Tondu T, et al. Accuracy of early burn depth assessment by laser doppler imaging on different days post burn. *Burns*. 2009;35(1):36-45.

23. 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(7):1960-5; discussion 1966-7.

24. Pape SA, Baker RD, Wilson D, et al. Burn wound healing time assessed by laser doppler imaging (LDI). part 1: Derivation of a dedicated colour code for image interpretation. *Burns*. 2012;38(2):187-194.

25. Jeng JC, Bridgeman A, Shivnan L, et al. Laser doppler imaging determines need for excision and grafting in advance of clinical judgment: A prospective blinded trial. *Burns*. 2003;29(7):665-670.

26. S.M. Monstrey, H. Hoeksema, R.D. Baker, J. Jeng, R.S. Spence, D. Wilson, S.A. Pape, (2011) Burn wound healing time assessed by laser Doppler imaging. Part 2: Validation of a dedicated colour code for image interpretation. Burns 37 (2):249-256

9

Scar quality after hydrosurgical versus conventional debridement for burns: a within-patient randomised double-blind trial

C.M. Legemate, K.A.A. Kwa, H. Goei, A. Pijpe, E. Middelkoop, P.P.M. van Zuijlen, G.I.J.M. Beerthuizen, M.K. Nieuwenhuis, M.E. van Baar, C.H. van der Vlies, the HyCon Study Group.

Accepted for publication in the British Journal of Surgery

ABSTRACT

Background

Tangential excision of burned tissue followed by skin grafting is the cornerstone of burn surgery. Hydrosurgery has become popular for tangential excision with the hypothesis that enhanced preservation of vital dermal tissue reduces scarring. The objective of this trial was to compare scar quality after hydrosurgical versus conventional debridement prior to split skin grafting.

Methods

This double-blind randomised within-patient multicentre RCT was conducted in patients with burns that required split skin grafting. One wound area was randomized to hydrosurgical debridement, the other to Weck knife debridement. The primary outcome was scar quality at 12 months assessed with the observer part of the Patient and Observer Scar Assessment Scale (POSAS). Secondary outcomes included complications, scar quality, colour, pliability and histological dermal preservation.

Results

A total of 137 patients were randomised. At 12 months, scars of the hydrosurgical debrided wounds were statistically significantly better in terms of the POSAS observer total item score (mean 2.42 (95 per cent c.i. 2.26 – 2.59) vs. 2.54 (95 per cent c.i. 2.36 – 2.72); $p=0.023$) and overall opinion score (mean 3.08 (95 per cent c.i. 2.88 – 3.28) vs. 3.30 (95 per cent c.i. 3.09 – 3.51); $p=0.006$). Patient reported scar quality and pliability measurements were significantly better for the hydrosurgical debrided wounds. Complication rates did not differ between both treatments. Histologically, significant more dermal was preserved with hydrosurgery ($p<0.001$).

Conclusion

One year post-surgery scar quality and pliability seemed better for hydrosurgical debrided burns. This is probably the result of enhanced histological preservation of dermis.

Registration number

Trial NL6085 (NTR6232 (http://www.trialregister.nl)).

Funding

The Dutch Burns Foundation, grant number 15.101.

INTRODUCTION

Early debridement and split skin grafting is the standard of care for deep dermal burn wounds to maximize recovery of the affected area and minimize pathological scarring.[1]

Conventional surgical debridement consists of sharp tangential excision of non-viable tissue with hand-held knives until bleeding tissue is encountered as a marker for vital tissue.[2] Commonly used instruments for conventional debridement include the Watson knife, the Humby knife, the Goulian or the Weck knife. Hydrosurgical debridement is an alternative to conventional knife debridement. The principle of hydrosurgery is the emission of a jet of water across an aperture that causes a localised vacuum to simultaneously cut, irrigate and suction tissue. The speed of the jet of water can be adjusted by the surgeon and is claimed to lift only non-viable tissue, thereby achieving an accurate wound debridement with maximal preservation of viable dermis.[3]

Loss of dermis has been considered one of the main factors determining the quality of a scar.[4] Burn specialists widely use hydrosurgery as an alternative for conventional tangential debridement. The underlying hypothesis is that scar quality would be better after hydrosurgical debridement as it enables surgeons to accurately debride burned tissue with maximal preservation of viable dermis in contrast to conventional surgical debridement, which is associated with the unnecessary tissue loss. A recent Cochrane review showed uncertainty whether or not hydrosurgical debridement and skin grafting is better than conventional surgical debridement and skin grafting for the treatment of acute partial-thickness burns and concluded that more trials are needed.[5]

The aim of this within patient randomized clinical trial (RCT) was to compare and evaluate long term scar quality of patients whose burns were debrided with hydrosurgical or conventional techniques prior to split skin grafting.

METHODS

The HyCon study (long-term scar quality after Hydrosurgical versus Conventional debridement for deep dermal burns) is a multi-centre, within-patient randomised, double-blind trial. The Medical Ethics Committee and the institutional review boards of each participating hospital approved the study protocol, which has been published elsewhere.[6] The study was registered at the Netherlands Trial Register before the start of recruitment (Trial NL6085 (NTR6232)). The study was conducted according to the principles of the Declaration of Helsinki (Ethics manual World Association revision 2013) and in accordance with the Medical Research Involving Human Subjects Act (WMO) and the CONSORT statement for reporting within-person randomized trials.[7]

Setting and recruitment

Participants were recruited in the three specialized burn centres in the Netherlands: the Maasstad Hospital in Rotterdam, the Martini Hospital in Groningen and the Red Cross Hospital in Beverwijk. National guidelines advise to refer a patient to one of these specialized burn centers if they fulfill one of the Emergency Managements of Severe Burns (EMSB) referral criteria.[8, 9]

Eligible patients had burns with a surface area larger than 50cm^2 that required debridement and split-skin grafting. There was no age restriction. Patients with full thickness burns were excluded as the hydrosurgery system cannot cut through hard leather-like eschar. Other exclusion criteria were wound infection, insufficient knowledge of the Dutch or English language and patients who were unlikely to comply with the requirements of the follow-up. Patients or their legal representatives gave written informed consent before inclusion in the study. The inclusion criteria were adapted to overcome low eligibility rates in the first months of the inclusion period. Contrary to the published protocol, both study areas did not have to be adjacent if they were both suitable for hydrosurgical and conventional debridement and of equal depth determined by an experienced burn physician preferably in combination with a Laser Doppler Imaging scan. Also, patients with a total body surface area (TBSA) burned of more than 30% were included.[6]

Procedures and interventions

Every participant acted as their own control. Two similar wound areas of at least 25cm^2 were selected by the surgeon in each patient and assigned A or B. If possible, the study areas were adjacent. Otherwise, a similar burn wound at the contralateral body part or the nearest comparable burn wound was chosen. Photographs were taken to facilitate identification of both areas during follow-up. After assignment by the surgeon and before surgery started, the study areas were randomly allocated to either hydrosurgical or conventional debridement with a Weck knife using a web-based automated randomisation system (https://data.castoredc.com) in the operating room by a member of the research team. During the operation, the study areas were debrided with the VERSAJET™ Hydrosurgery System (Smith+Nephew, London UK) or conventional surgery using a hand-held knife. Both study areas were debrided during the same procedure and covered with the same size meshed split skin graft or Meek wall grafts and identical non-adhesive wound dressings. Graft harvesting, meshing and fixation were done following local treatment protocols. This design allowed comparison of hydrosurgical versus conventional debridement within the same participant, while controlling for variations in healing and scarring that could occur between patient groups. Dressings were left in situ for 5-7 days. Both study areas were followed until complete wound closure (at least 95 per cent re-epithelialization) was achieved, assessed by a member of the research team and documented with photographs and notes from a clinician in the patient's file as instructed by a standard operating procedure. Standard of care involved visits to the outpatient clinic at 3, 6 and 12 months post-surgery.

Clinicians/researchers that assessed scar quality during follow-up and patients were blinded to the modality used to debride the study areas.

Baseline characteristics

Investigators not involved in the clinical care of participants were responsible for trial recruitment, allocation and data collection. They recorded the following baseline characteristics for all included patients: age, gender, Fitzpatrick skin type, comorbidities, percentage TBSA burned, wound aetiology, burn depth, and time to surgery. During surgery, Weck knife, Versajet settings and skin graft expansion were registered.

Primary outcome measure

The primary outcome was scar quality at 12 months assessed by the clinician/researcher with the Patient and Observer Scar Assessment Scale (POSAS) version 2.0.[10] The POSAS questionnaire consists of two six-item scales; one for the observer (clinician/researcher) and one for the patient. The observer total item score was chosen as the primary outcome because it has been demonstrated to produce valid and reproducible results by trained evaluators.[11, 12] The observer part includes the items vascularity, pigmentation, thickness, relief, pliability and surface area. These items were separately scored on a 10-point rating scale with 1 corresponding to the situation of normal skin and 10 indicating the worst imaginable scar. Two independent observers scored the scar quality to improve the reliability of the assessment.[11] The mean score of the items of both observers formed the observer total item score.[11, 12]

Secondary outcome measures

Secondary outcome measures included complications and wound healing, scar colour, scar pliability and dermal preservation measured by histology.

Complications and wound healing

The presence of complications, such as infection (clinical signs in combination with a positive swab), percentage graft loss and hematoma were registered per study area. Prolonged wound healing was defined as more than 2 weeks to achieve ≥95% reepithelialisation.

Scar quality measures

Patient reported scar quality was measured with the total item score of the patient part of the POSAS. The patient part of the POSAS includes the parameters pain, itch, color, thickness, relief, and pliability. Pain and itch were scored between 1 (no pain/itch) and 10 (extreme pain/itch). Each of the other items was scored between 1 (no difference with normal skin) and 10 (very different from normal skin). The mean score of these items formed the Patient total item score. In addition to the item scores, both observers and patients gave a score for their overall opinion on the scar on a 10-point rating scale, where 1 resembles normal skin and 10 resembles

the worst imaginable scar. Because scar quality changes over time, the POSAS was used to assess scar quality during standard follow-up visits to the outpatient clinic at 3, 6 and 12 months postoperatively. At the 12 month visit, patients were asked to indicate the degree of clinical difference they noticed between study area A and B on a 5-point Likert scale (much worse, worse, the same, better, a lot better). The purpose of this question was to gain insight in patients' perspectives regarding what is clinically important.

Scar colour

Scar colour was evaluated measuring erythema and melanin with the DSM II ColorMeter (Cortex Technology, Hadsund, Denmark).[19] Colour results were expressed as the absolute difference between healthy skin and scar to eliminate season-related influence of sun exposure on the erythema and melanin scores.

Scar pliability

Scar pliability was measured using the Cutometer Skin Elasticity Meter 575 (Courage and Khazaka GmbH, Cologne, Germany). The Cutometer measures the vertical deformation of the skin in millimetres into the circular aperture of the probe after a controlled vacuum. Two Cutometer parameters that were previously shown to be the most reliable were used: elasticity (Ue) and maximal extension (Uf).[13] To eliminate influence of different anatomic locations, elasticity was analysed using the ratio of the scar and normal skin.

Colour and pliability measurement procedures

To prevent measurement bias within the scar, the colour and pliability measurements were performed on five locations following a standardized method including 5 scar measurements.[14] The average score of these five scar measurements was used. The first option for the control measurement was the patient's unaffected contralateral site. In cases where the contralateral site was also affected, the most comparable and unaffected spot near the scar was used. Measurements were performed at 3 and 12 months post-surgery.

Dermal preservation

The amount of remaining dermis after debridement was evaluated with histopathology. During surgery, punch biopsies (diameter 3mm) were taken from both study areas after debridement. In addition to the hematoxylin and eosin (H&E) staining described in our published protocol, we used a Herovici polychrome staining to analyse the biopsies to differentiate between mature collagen and granulation tissue.[15, 16] All resection specimens were processed and sampled using a standard protocol (Appendix).

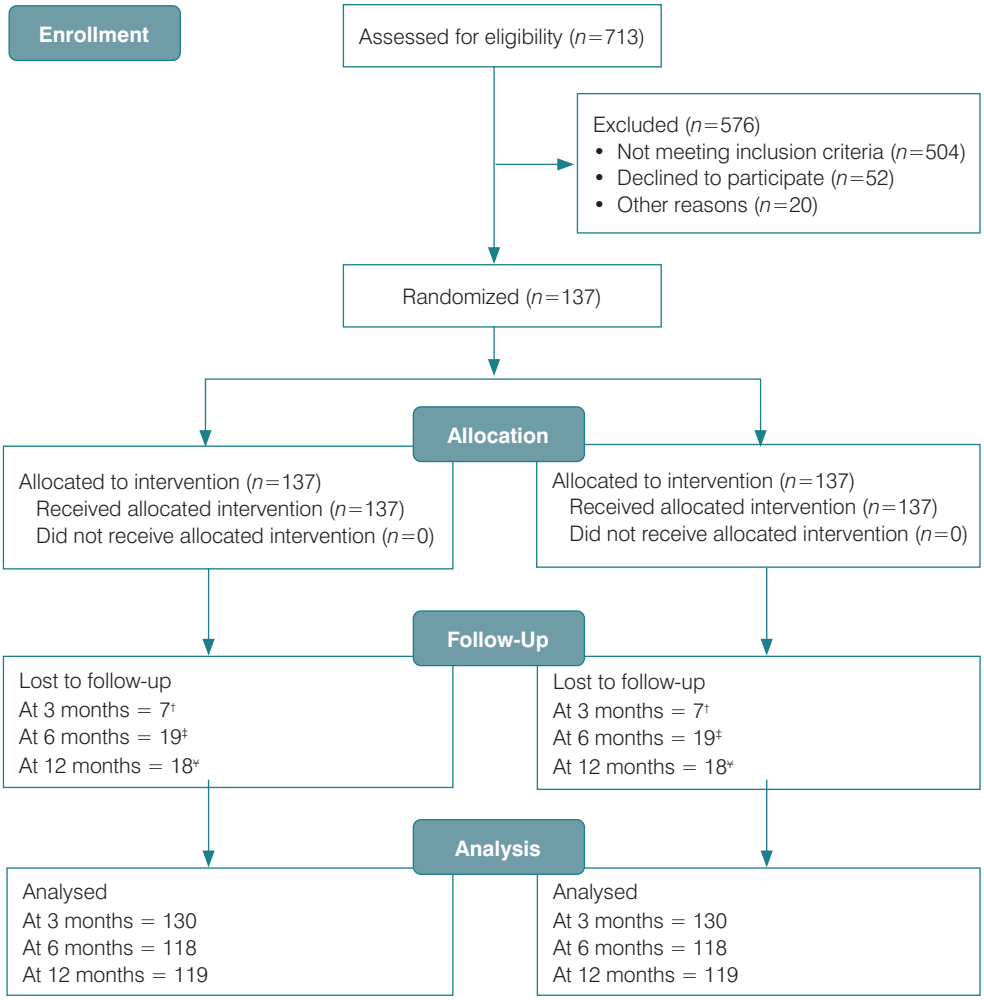


Figure 1. CONSORT diagram for the trial

*'other' reasons for exclusion were *missed by the study team* or *participation in another intervention study.* †6 loss to follow-up, 1 drop out because of new self inflicted burn wounds in study areas. ‡ 17 loss to follow-up, 1 drop out because of new wounds in study areas, 1 deceased. ¥16 loss to follow-up, 1 drop out because of new wounds on study areas, 1 deceased.

Statistical analysis

A sample size calculation was performed based on the results of an unpublished retrospective study on scar quality after hydrosurgery versus guarded knife excision assessed from photographs by carers.[17] In this study, the total item score of the observer scale of the POSAS questionnaire 12 months post-surgery was 14.7 in the hydrosurgery group versus 16.7 in the conventional debridement group with a pooled standard deviation (SD) of 6.53, resulting in an effect size of 0.3. Given a power of 0.90, a significance level of 0.05 and including a correction

for correlated samples and lost-to-follow-up, a sample size of 137 was calculated. Continuous data were first tested for normality. Normally distributed data are presented as mean (95 per cent c.i.) and testing was performed with paired t-tests. Non-normally distributed data are presented as median (interquartile range (i.q.r.)) values and analysed with the Wilcoxon signed ranks test for paired data. Effect sizes for the paired t-test were represented using cohen's d, for Wilcoxon signed rank test *r* was used.[18] The McNemar's test was used for paired dichotomous values and odds ratios were used to represent the effect size. Because of repeated measurements within patients, overall differences between both treatments in subjective scar quality outcomes was analysed using generalized estimating equation (GEE) with an exchangeable correlation matrix structure. Significance was set at $p < 0.050$. Analyses were conducted using SPSS® version 25 (IBM, Armonk, New York, USA) and Stata® version 16 (StataCorp, College Station, Texas, USA).

RESULTS

From January 2017 to July 2019, 713 patients were screened for inclusion, of whom 137 were eligible to be randomized (Fig. 1). Patient characteristics and surgical details are shown in Table 1. An example of assignment of the study area and a scar at 12 months post-surgery are shown in Fig. 2.

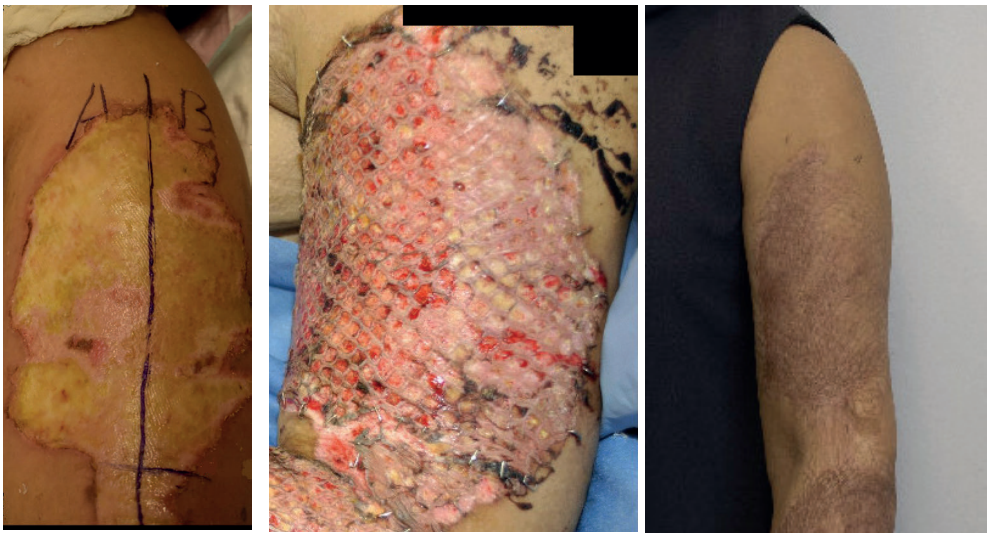


Figure 2. Allocation of wound areas on the left upper arm before randomization (left), wound inspection 5 days post-surgery (middle) and scarring during 12 months follow-up (right) in a 61 year old female. Part A was randomized to debridement with a Weck knife, Part B was randomized to debridement with hydrosurgery.

Table 1. Baseline characteristics

Characteristic	Total (n = 137)
Age (years)†	45 (25 – 59)
Sex ratio (F:M)	54:83
Skin type	
Fitzpatrick 1-2	79 (57)
Fitzpatrick 3-4	49 (36)
Fitzpatrick 5-6	9 (7)
Diabetes	11 (8)
% TBSA burned†	7 (4 – 14)
Aetiology	
Flame	85 (62)
Scald	26 (20)
Fat	20 (15)
Other	6 (4)
Surgical characteristics	
Time from injury to surgery (days)†	15 (10 – 19)
Versajet setting*	5.25 (4.82 – 5.69)
Blade weck knife	
0.008 inch	17 (12)
0.010 inch	19 (14)
0.012	81 (59)
Unknown	20 (15)
Skin graft expansion	
1:1	16 (12)
1:1.5	66 (49)
1:2	10 (7)
1:3	30 (22)
1:6 (Meek technique)	15 (11)

Values in parentheses are percentages unless indicated otherwise; values are *mean(SD) †Median (i.q.r.); TBSA = Total Body Surface Area

Primary outcome

Scar quality measured with the observer total item score of the POSAS at 12 months was statistically significant better for hydrosurgical debrided burns than for the conventional debrided burns (mean difference -0.12 (95 per cent c.i. -0.22 to -0.02), p = 0.023) (Table 2)).

Table 2. POSAS scores 12 months post-surgery

	Hydrosurgical debridement	Conventional debridement	Effect-size§	p- value
Observer scar score				
Total item score†	2.42 (2.26 – 2.59)	2.54 (2.36 – 2.72)	-0.21	0.023*
Vascularity†	2.47 (2.26 – 2.69)	2.59 (2.37 – 2.81)		
Pigmentation†	2.81 (2.62 – 3.00)	2.89 (2.70 – 3.08)		
Thickness†	2.30 (2.08 – 2.51)	2.81 (2.57 – 3.19)		
Relief†	2.66 (2.44 – 2.87)	2.81 (2.58 – 3.05)		
Pliability‡	2.50 (1.50 – 3.00)	2.50 (1.50 – 3.50)		
Surface area‡	1.50 (1.00 – 2.00)	1.50 (1.00 – 2.00)		
Overall opinion score†	3.08 (2.88 – 3.28)	3.30 (3.09 – 3.51)	-0.25	0.006*
	Hydrosurgical debridement	**Conventional debridement**	**Effect-size§**	**p- value**
Patient scar score				
Total item score‡	2.68 (1.67 – 4.33)	3.00 (1.83 – 4.83)	-0.14	0.019¥
Pain‡	1 (1 – 1)	1 (1 – 1)		
Pruritus‡	1 (1 – 3)	1 (1 – 3)		
Colour‡	4 (3 – 6)	4 (3 – 6)		
Stiffness‡	3 (1 – 5)	3 (1 – 6)		
Thickness‡	2 (1 – 5)	3 (1 – 6)		
Relief‡	3 (1 – 5)	4 (2 – 6)		
Overall opinion score‡	4 (2– 6)	4 (3 – 6)	-0.15	0.024¥

POSAS scores range from 1 – 10. A lower score correlates with a better scar. Observer scores are the mean scores of the six items scored by 2 clinicians/researchers.

†mean (95 per cent c.i.), ‡ median (i.q.r),

§ Effect size for paired t-test represented using cohen's d. Effect size for Wilcoxon signed rank test represented using r (Cohen, 1988)

*paired t-test, ¥ Wilcoxon signed rank test

Secondary outcomes

Wound healing and complications

Time to re-epithelialization did not differ between both intervention groups. No significant differences in wound infection, percentage graft loss and other complications were found between treatment groups (Table 3).

Scar quality

One year after surgery, scar quality scores were significantly lower (i.e. reflecting a better scar) for hydrosurgical debrided burns in terms of the observer overall opinion score, the patient total item score and patient overall opinion score (Table 2). Observer and patient reported POSAS scores at 3 and 6 months post-surgery are presented in Table 3. On average there was a significant

better outcome for the hydrosurgical debrided wounds over time in terms of the observer total item score (-0.16 (95 per cent c.i. -0.25 – -0.06, p=0.001)), observer overall opinion score (-0.22 (95 per c.i. -0.34 – -0.09, p=0.001)), and patient total item score (-0.29 (95 per cent c.i. -0.49 – -0.09, p=0.0024)), but not for the patient overall opinion score (-0.18 (95 per cent c.i. -0.75 –0.40, p=0.547) using GEE analyses. At 12 months post-surgery, 56 patients (48%) rated the hydrosurgical debrided study area as a better scar on the Likert scale, 30 patients (26%) rated the conventional debrided study area as better, and 30 patients (26%) said they noticed no difference between both study areas.

Scar colour

The erythema index of the hydrosurgical debridement and conventional debrided study areas did not differ significantly at 3 and 12 months post-surgery (Table 3). The melanin index for scars of wounds that were treated with hydrosurgery were significantly more comparable to normal skin at 3 months, but did not differ at 12 months (Table 3).

Table 3. Secondary outcomes

	Hydrosurgical debridement	Conventional debridement	Hydrosurgical versus Conventional	P value*
			Odds ratio	
Complications				
Wound infection	19 (13.9)	20 (14.6)	0.93 (0.47 – 1.86)	1.000[£]
Graft loss (partial or total)	3 (2.2)	8 (5.8)	0.36 (0.09 – 1.39)	0.227[£]
Prolonged wound healing	29 (21.1)	22 (16.1)	1.40 (0.68 – 2.87)	0.065[£]
Other	1 (0.7)	3 (2.2)	0.33 (0.03 – 3.22)	0.625[£]
			Effect size[a]	
Wound healing				
Re-epithelialization 5-7 days post-surgery (%)[†]	80.0 (68.96 – 79.64)	81.2 (73.20 – 82.32)	-0.13	0.144
Time to re-epithelialization (days)[‡]	7 (5 – 13)	7 (5 – 12)	-0.10	0.353[¥]
Observer POSAS sores				
Total item score				
3 months[†]	3.04 (2.91 – 3.22)	3.18 (3.03 – 3.36)	-0.21	0.021
6 months[†]	2.72 (2.57 – 2.91)	2.93 (2.75 – 3.10)	-0.29	0.002
Overall opinion score				
3 months[†]	3.87 (3.69 – 4.11)	4.07 (3.86 – 4.27)	-0.21	0.031
6 months[†]	3.51 (3.33 – 3.74)	3.73 (3.54 – 3.94)	-0.26	0.006

Table 3. Continued

Patient POSAS sores				
Total item score				
3 months†	4.14 (3.79 – 4.50)	4.55 (4.22 – 4.88)	-0.28	0.002
6 months‡	3.33 (2.21 – 4.83)	3.67 (2.54 – 5.17)	-0.11	0.093ᵛ
Overall opinion score				
3 months‡	5 (3 – 7)	6 (4 – 8)	-0.14	0.011ᵛ
6 months‡	5 (2 – 6)	5 (3 – 7)	-0.21	0.001ᵛ
ColourΩ				
Erythema				
3 months‡	6.26 (3.38 – 10.67)	6.68 (3.11 – 10.35)	-0.02	0.560ᵛ
12 months‡	3.61 (2.08 – 8.82)	4.08(1.60 – 9.51)	-0.04	0.596ᵛ
Melanin				
3 months‡	13.24 (6.97 – 21.72)	13.96 (8.13 – 22.54)	-0.16	0.041ᵛ
12 months‡	7.19 (4.30 – 13.87)	6.24 (3.77 – 15.02)	0.24	0.607ᵛ
Pliability§				
Elasticity *(Ue)*				
3 months‡	0.55 (0.35 – 0.74)	0.49 (0.32 – 0.71)	0.08	0.225ᵛ
12 months‡	0.73 (0.58 – 0.91)	0.70 (0.53 – 0.89)	0.15	0.029ᵛ
Maximal extension *(Uf)*				
3 months†	0.58 (0.50 – 0.72)	0.55 (0.47 – 0.62)	0.18	0.089
12 months‡	0.75 (0.62 – 0.91)	0.72 (0.56 – 0.90)	0.14	0.039ᵛ
Histopathological findings				
Dermal preservation (μm)‡	1748 (1213 – 2175)	1265 (689 – 1989)	0.23	<0.001ᵛ

Values in parentheses are percentages unless indicated otherwise; †mean (95 per cent c.i.), ‡median (i.q.r)

*Paired t-test unless indicated otherwise; ᵛWilcoxon signed rank test and ᶜMcNemar's test

ᵒ Effect size for paired t-test represented using cohen's d. Effect size for Wilcoxon signed rank test represented using *r* (Cohen, 1988)[18]

ΩMeans are calculated as absolute difference between scar tissue and the non-affected skin

§Values represent the ratio between scar tissue and non-affected skin

Scar pliability

At 12 months post-surgery, scars of hydrosurgical debrided wounds were more comparable to normal skin in terms of the pliability parameters elasticity and maximal extension ($p = 0.029$ and $p = 0.039$, Table 3).

Histopathological findings

Of the hydrosurgical debrided study areas, 104 biopsies were included for analyses. Of the conventional debrided study areas, 101 biopsies were qualified for analyses. More dermis was left in the punch biopsies of wounds that were debrided with hydrosurgery ($p < 0.001$, Table 3).

DISCUSSION

The use of hydrosurgery led to better scar-quality outcomes, as reported by clinicians and patients, up to 1 year postsurgery. Objective scar pliability measures were also significantly better which was probably the result of better preservation of dermis after hydrosurgical debridement.

An important topic to discuss is whether statistical differences in POSAS outcomes present a clinically significant difference. Effect sizes in observer outcomes were small (ranging from −0.21 to −0.29) but not trivial.[19] However, the effect sizes of patient outcomes were smaller (ranging from −0.11 to −0.28).[20] Of the trial population, 48% considered the hydrosurgically debrided study area as better or much better at 12 months post-surgery versus 26% of the conventionally debrided study area. In addition, unpublished data from our institute suggest that patients consider differences between −0.08 and −0.39 in patient POSAS item scores as important, which may indicate that the differences identified are at least of some importance to patients. However, in the absence of a minimal clinically important difference in POSAS score, uncertainty remains over what difference in outcome should be considered clinically important.

The goal of debridement of a burn wound is to remove injured and non-viable tissue to create the optimal wound bed for autologous split-thickness skin grafting.[21-23] An essential asset of an effective debridement tool is to remove as much necrotic tissue as possible while preserving as much vital tissue as possible, to improve clinical, functional, and cosmetic outcomes.[4, 24, 25] Although specialists in burns have long recognized the association between the depth of dermal injury and the degree of scarring, the cellular and molecular basis of the relationship remains poorly understood. Dunkin et al. hypothesized that different depths of dermal injury may result in different inflammatory responses and cytokine profiles, which, in turn, provide an environment for a different proliferative response.[25] The current study is the first to report a relationship between dermal preservation and better clinical scar outcomes. POSAS item scores related to elasticity (thickness, relief, pliability, and stiffness) differed most between both study groups. These results, in combination with better measurement of pliability for scars after hydrosurgical debridement, suggest that the preservation of dermis leads to better scar quality in terms of how scars 'feel' rather than how scars 'look' (vascularity, pigmentation, and colour). Further studies are necessary to better understand the relationship between dermal preservation (i.e. selective debridement) and scar quality. Time to wound healing and complication rates did not

differ between both treatment groups and can therefore be excluded as causes for superior scar quality after hydrosurgical debridement. This also implies that both techniques provided sufficient debridement.

Although several studies have reported that hydrosurgery can be used maximally to preserve dermis, only one has reported histological evidence to support this in burns.[26-31] Hyland et al. performed a randomized controlled trial to study scar quality after hydrosurgical and conventional debridement in children with partial-thickness burns.[30] They also confirmed greater loss of dermis in conventionally debrided burns via an analysis of histological specimens. They also found better scar scores in favour of the hydrosurgery group at 3 and 6 months postburn, but this difference was not statistically significant. However, they did not use a within-patient design, did not report the distribution of patient characteristics that may have influenced scar quality, and ended their follow-up at 6 months.

A recent Cochrane review reported low-quality evidence for the potential benefits of hydrosurgical debridement over conventional debridement that are desirable to clinicians, such as faster operating time, improved usability, fewer procedures, less blood loss, and a shorter hospital stay.[5] To reduce treatment costs, burn specialists often prefer to use a Weck knife in smaller burns. The current cost of the disposable Versajet headpiece is €141.86 ($167.55), while the costs of a reusable Weck knife guard and handle are €0.91 ($1.08) and €20.91 ($24.70), respectively. The cost of one sterilized, single-use Weck blade is €1.08 ($1.28). However, evidence for the cost-effectiveness of hydrosurgery in burns is limited and further research on its long-term benefits, such as fewer reconstructive surgery procedures, is necessary.[32]

The strengths of this trial include the comprehensive inclusion criteria (patients of all ages and most burn aetiologies), which allowed application to a broad patient population. The within-patient design reduced factors that could lead to confounding of scar outcomes. Much effort was put into minimizing the risk of bias due to incomplete outcome data, which resulted in a remarkably low drop-out rate after 1 year follow-up (13%). The multicentre approach and pragmatic character (mimicking routine clinical practice) improve the generalizability of the results. Another strength is the use of the POSAS instrument; it is validated, incudes most relevant scar characteristics, and is the most frequently used scale.[33-36]

The key assets of the trial also create the main limitations, including the within-patient design and outcome measure. For patients, the POSAS may have been difficult to rate for two study areas. In particular, after adjustment of the protocol that allowed study areas to be on different parts of the body, the patient's opinion on scar quality might have been biased based on the location of the scar, especially if scars are further apart on the body or if one is in the sight and the other is not. To increase reliability, the observer part of the POSAS was used as the primary outcome,

and was scored by two independent trained observers. To improve the feasibility of the trial, there were no restrictions or standards for delivery of the intervention, depth of graft harvesting, skin graft fixation, or wound treatment post-surgery. However, both wound areas were treated the same within patients and therefore this may lead to improvement of generalizability of the results rather than bias of scar-quality outcomes. Previous research has shown that burn surgeons tend to use hydrosurgery more often in children and more superficial wounds (like scalds).[6] Subgroup analyses were not part of the initial research plan and further research is necessary to gain insight into the benefits of hydrosurgery for different patient and burn categories. Scar outcomes present data from specialized burn care, which might not be comparable to outcomes in non-specialized centres. Nevertheless, hydrosurgical debridement is easy to learn and the device is not difficult to use. Therefore, its use may lead to better scar outcomes in centres where surgeons are not frequently practising burn debridement with guarded knives.

Funding

This research was funded by the Dutch Burns Foundation (grant number 15.101).

Disclosure. The authors declare no conflict of interest.

The HyCon Study Group

Collaborators: J. Dokter, K.L.M. Gardien, J. Hiddingh, H.W.C. Hofland, Y. Lucas, A. Meij-de Vries, T.H.J. Nijhuis, I.M.M.H. Oen, D.T. Roodbergen, S.M.H.J. Scholten-Jaegers, M. Stoop, M. Vlig, N. Trommel

APPENDIX

The biopsies were fixed in kryofix, processed into 5μm histological slides and stained with a Herovici polychrome staining.[37] A digital camera (Nikon DS-Ri2, Nikon, Amsterdam, the Netherlands) mounted on an Axioskop40FL microscope (Zeiss, Badhoevedorp, the Netherlands) was used to take images of the slides. Digital image analysis software (NIS-Elements 4.4, Nikon) was used to determine the amount of remaining dermis, measured as the length in μm between two parallel lines; one at the subcutis and one at the dermal surface (Figure 1). If the border between dermis and subcutis was not visible on the slides, or only subcutis was present, slides were excluded for analyses.

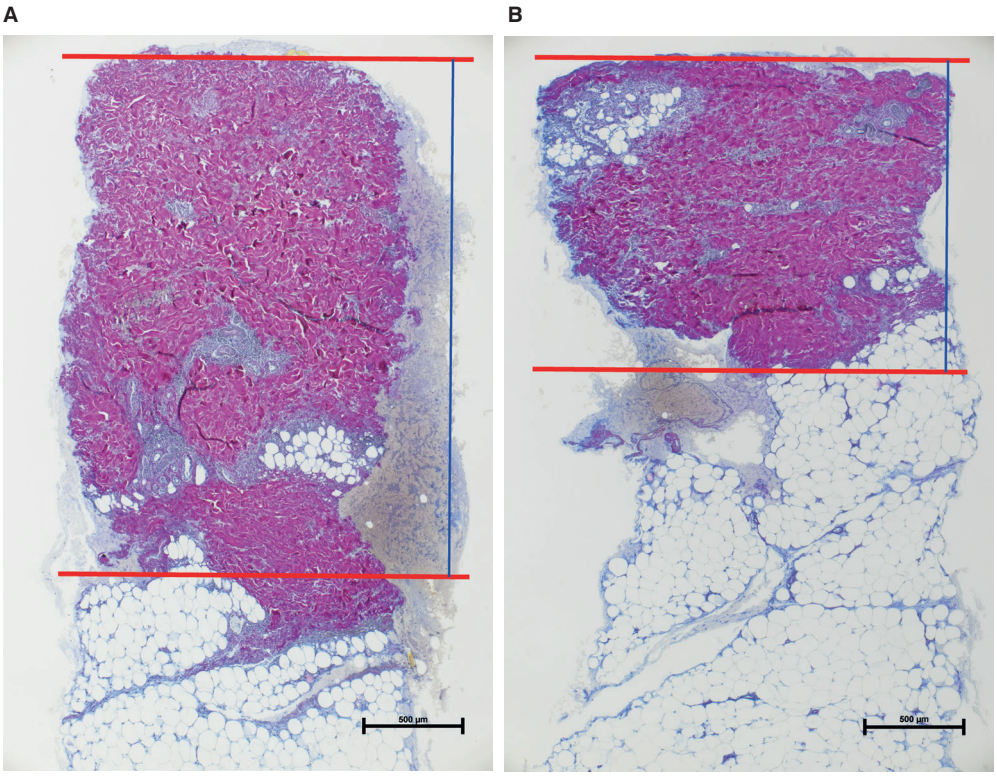


Figure 1a and 1b. Histological specimen demonstrating an example of Herovici polychrome staining and measurement procedure after hydrosurgical (Figure 1a, 2362 μm dermis left) and conventional (Figure 1b, 1265 μm dermis left) debridement.

REFERENCES

1. Muller M GD, Herndon DN. Operative Wound Management. In: *Total Burn Care*. Elsevier inc., 2007; 177-195.
2. Orgill DP. Excision and skin grafting of thermal burns. *N Engl J Med* 2009;360(9): 893-901.
3. The Versajet II hydrosurgery system for surgical debridement of acute and chronic wounds and burns. Medtech innovation briefing [MIB1]. www.nice.org.uk/advice/mib1 [18 august 2020.
4. Mensik I, Lamme EN, Brychta P. Depth of the graft bed influences split-skin graft contraction. *Acta Chir Plast* 2003;45(3): 105-108.
5. Wormald JC, Wade RG, Dunne JA, Collins DP, Jain A. Hydrosurgical debridement versus conventional surgical debridement for acute partial-thickness burns. *Cochrane Database Syst Rev* 2020;9: CD012826.
6. Legemate CM, Goei H, Gostelie OFE, Nijhuis THJ, van Baar ME, van der Vlies CH, Dutch Burn Repository G. Application of hydrosurgery for burn wound debridement: An 8-year cohort analysis. *Burns* 2019;45(1): 88-96.
7. Pandis N, Chung B, Scherer RW, Elbourne D, Altman DG. CONSORT 2010 statement: extension checklist for reporting within person randomised trials. *BMJ* 2017;357: j2835.
8. Van Yperen DT, Van Lieshout EMM, Nugteren LHT, Plaisier AC, Verhofstad MHJ, Van der Vlies CH, Burns study g. Adherence to the emergency management of severe burns referral criteria in burn patients admitted to a hospital with or without a specialized burn center. *Burns* 2021.
9. Australian and New Zealand Burn Association Referral Criteria. In; 2012.
10. http://www.posas.org.
11. Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, van Zuijlen PP. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. *Plast Reconstr Surg* 2004;113(7): 1960-1965; discussion 1966-1967.
12. van der Wal MB, Tuinebreijer WE, Bloemen MC, Verhaegen PD, Middelkoop E, van Zuijlen PP. Rasch analysis of the Patient and Observer Scar Assessment Scale (POSAS) in burn scars. *Qual Life Res* 2012;21(1): 13-23.
13. Draaijers LJ, Botman YA, Tempelman FR, Kreis RW, Middelkoop E, van Zuijlen PP. Skin elasticity meter or subjective evaluation in scars: a reliability assessment. *Burns* 2004;30(2): 109-114.
14. Verhaegen PD, van der Wal MB, Bloemen MC, Dokter J, Melis P, Middelkoop E, van Zuijlen PP. Sustainable effect of skin stretching for burn scar excision: long-term results of a multicenter randomized controlled trial. *Burns* 2011;37(7): 1222-1228.
15. Herovici C. [Picropolychrome: histological staining technic intended for the study of normal and pathological connective tissue]. *Rev Fr Etud Clin Biol* 1963;8: 88-89.
16. Jaspers MEH, Feroldi F, Vlig M, de Boer JF, van Zuijlen PPM. In vivo polarization-sensitive optical coherence tomography

of human burn scars: birefringence quantification and correspondence with histologically determined collagen density. *J Biomed Opt* 2017;22(12): 1-8.

17. Scholten SMHJ, den M, Nieuwenhuis M, Niemeijer A, Beerthuizen GIJM. O27.3 Scar quality after surgical treatment of deep dermal burns with hydro-surgery compared to quarded knife; short and long term outcome. *Burns* 2011;37: S18.
18. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 2012;141(1): 2-18.
19. Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. *J Grad Med Educ* 2012;4(3): 279-282.
20. Ferguson CJ. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice* 2009;40(5): 532-538.
21. Eldad A, Weinberg A, Breiterman S, Chaouat M, Palanker D, Ben-Bassat H. Early nonsurgical removal of chemically injured tissue enhances wound healing in partial thickness burns. *Burns* 1998;24(2): 166-172.
22. Greenhalgh DG. Management of Burns. *N Engl J Med* 2019;380(24): 2349-2359.
23. Robson MC, Krizek TJ. Predicting skin graft survival. *J Trauma* 1973;13(3): 213-217.
24. Jones T, McDonald S, Deitch EA. Effect of graft bed on long-term functional results of extremity skin grafts. *J Burn Care Rehabil* 1988;9(1): 72-74.
25. Dunkin CSJ, Pleat JM, Gillespie PH, Tyler MPH, Roberts AHN, McGrouther DA. Scarring occurs at a critical depth of skin injury: precise measurement in a graduated dermal scratch in human volunteers. *Plast Reconstr Surg* 2007;119(6): 1722-1732.
26. Kimble RM, Mott J, Joethy J. Versajet hydrosurgery system for the debridement of paediatric burns. *Burns* 2008;34(2): 297-298; author reply 299.
27. Matsumura H, Nozaki M, Watanabe K, Sakurai H, Kawakami S, Nakazawa H, Matsumura I, Katahira J, Inokuchi S, Ichioka S, Ikeda H, Mole T, Smith J, Martin R, Aikawa N. The estimation of tissue loss during tangential hydrosurgical debridement. *Ann Plast Surg* 2012;69(5): 521-525.
28. Cubison TC, Pape SA, Jeffery SL. Dermal preservation using the Versajet hydrosurgery system for debridement of paediatric burns. *Burns* 2006;32(6): 714-720.
29. Tenenhaus M, Bhavsar D, Rennekampff HO. Treatment of deep partial thickness and indeterminate depth facial burn wounds with water-jet debridement and a biosynthetic dressing. *Injury* 2007;38 Suppl 5: S39-45.
30. Hyland EJ, D'Cruz R, Menon S, Chan Q, Harvey JG, Lawrence T, La Hei E, Holland AJ. Prospective, randomised controlled trial comparing Versajet hydrosurgery and conventional debridement of partial thickness paediatric burns. *Burns* 2015;41(4): 700-707.
31. Gravante G, Delogu D, Esposito G, Montone A. Versajet hydrosurgery versus classic escharectomy for burn debridment: a prospective randomized trial. *J Burn Care Res* 2007;28(5): 720-724.
32. Kakagia DD, Karadimas EJ. The Efficacy of Versajet Hydrosurgery System in Burn Surgery. A Systematic Review. *J Burn Care Res* 2018;39(2): 188-200.

33. Jones LL, Calvert M, Moiemen N, Deeks JJ, Bishop J, Kinghorn P, Mathers J, team P. 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(8): 1682-1692.

34. Forbes-Duchart L, Cooper J, Nedelec B, Ross L, Quanbury A. Burn therapists' opinion on the application and essential characteristics of a burn scar outcome measure. *J Burn Care Res* 2009;30(5): 792-800.

35. 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): 275-285.

36. Carriere ME, Kwa KAA, de Haas LEM, Pijpe A, Tyack Z, Ket JCF, van Zuijlen PPM, de Vet HCW, Mokkink LB. Systematic Review on the Content of Outcome Measurement Instruments on Scar Quality. *Plast Reconstr Surg Glob Open* 2019;7(9): e2424.

37. Herovici C. [Picropolychrome: histological staining technic intended for the study of normal and pathological connective tissue]. *Rev Fr Etud Clin Biol* 1963;8:88-9. [published Online First: 1963/01/01]

10

Course of scar quality of donor sites following split skin graft harvesting: Comparison between patients and observers

C.M. Legemate, P.J. Ooms, N. Trommel, H. Goei, Y. Lucas, E. Middelkoop, M.E. van Baar, C.H. van der Vlies

Wound Repair Regen. 2020 Sep;28(5):696-703

ABSTRACT

There exists little to no data on the development of donor-site scars that remain after split skin graft harvesting. The objectives of this study were to (a) examine changes in characteristics of donor-site scar quality over time and (b) assess the agreement between patient-reported and observer-reported donor-site scar quality in a burn population. A prospective cohort study was conducted including patients who underwent split skin grafting for their burn injury. Patients and observers completed the Patient and Observer Scar Assessment Scale (POSAS) for the first harvested donor site at 3 and 12 months post-surgery. This study included 80 patients with a median age of 34 years. At 3 months post-surgery, the patients scored the POSAS items itch and color as most deviant from normal skin, both improved between 3 and 12 months (3.1 vs 1.5 and 5.0 vs 3.5, respectively [$P < .001$]). Other scar characteristics did not show significant change over time. The patients' overall opinion score improved from 3.9 to 3.2 ($P < .001$). Observers rated the items vascularization and pigmentation most severe, only vascularization improved significantly between both time points. Their overall opinion score decreased from 2.7 to 2.3 ($P < .001$). The inter-observer agreement between patients and observers was considered poor (ICC < 0.4) at both time points. Results of current study indicate that observers underestimate the impact of donor-site scars. This has to be kept in mind while guiding therapy and expectations.

BACKGROUND

Split skin grafting remains a widely used reconstructive technique for chronic wounds, burns and other traumatic wounds. The procedure involves harvesting of the full epidermis and part of the dermis, creating a secondary wound at the donor-site. Because skin grafting is necessary to cover the wound, scars of these donor-sites might be considered as subservient. Unlike for other wounds, there is little to no data on the development of donor-site wounds and their final appearance, even though the donor-site can be a considerable burden to patients during and after the healing process[1,2].

Integrating scar evaluations of patients in clinical assessments is promoted based on findings that patient-rated scar severity is directly related to psychological distress whereas observer-rated scar severity is unrelated to psychological distress[3]. Assessment of scars by both patients' and professionals' provides more useful information regarding the patients' well-being compared to focusing on the separate assessments only[4].

Eskes et al. investigated patients' and observers' judgements and satisfaction with respect to donor-site scars at 3 months after wound healing[5]. They found discrepancies between patients' and observers' opinions on different characteristics of donor-site scars[5]. However, they included mainly male patients who underwent surgery in a non-acute setting and only included adults with a mean age of 59.6 years[6]. Furthermore, scars were assessed at twelve weeks post-surgery, which limits the insight into the final situation of the scar as the active transformation processes and maturation of scars takes at least one year[7-9].

The objectives of this study were to 1) examine changes in characteristics of donor-site scar quality over time and get insight into final scar appearance at 12 months post-surgery; and 2) to assess the extent of agreement between patients' ratings and observers' rating of donor-site scar characteristics in a burn population. The ultimate aim of current study was to improve information given to patients and eventually ensure high-quality patient-centered care.

PATIENTS AND METHODS

Study design and patients

An observational prospective cohort study was performed. From February 2016 – February 2017 patients were included in the burn center of the Maasstad Hospital in Rotterdam, the Netherlands. Patients of all ages who underwent split skin grafting for an acute burn were eligible to participate. Patients were excluded if they had cognitive impairments or were unable to understand or answer questionnaires in Dutch or English. Written informed consent was obtained from each

patient and patients received standard treatment. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were adhered to in this study and manuscript. The study was approved by the regional Medical Ethics Committee (reference number L2016119) and conducted according to the principles of the Declaration of Helsinki. Scar quality of the first harvested donor-site was evaluated at 3 and 12 months post-surgery during routine outpatient visits. Other study parameters were documented during admission, surgery and outpatient visits. These were patient characteristics: age at surgery, gender, skin type. Registered clinical characteristics were burn-related: % total burned body surface area (TBSA), % TBSA excised, length of stay, and donor site-related: location on the body, >2 weeks to re-epithelization, and wound infection.

Treatment

Patients received standard treatment peri-operatively. Split-thickness skin grafts were harvested at a depth of 0.2mm (0.007 inch) with an electric dermatome. Adrenaline soaked alginate dressings were place on the wounds immediately after grafting to reduce blood loss. Afterwards, donor-site wounds were covered with an alginate dressing, cotton wool and elastic bandages which were removed 2 weeks post-surgery.

Scar outcome

The Patient and Observer Scar Assessment Scale (POSAS) version 2.0 was used to assess scar quality[10]. The POSAS is a validated measurement scale for scar quality of burn and linear scars and therefore seemed most suitable to assess scar quality of donor-sites[10-13]. The POSAS consists of a patient and observer (i.e. caregiver) part. Both patient and observers rated the same scar on six different scar characteristics, with roughly an overlap of four characteristics (Table 2). This enabled the identification of differences between patients and observers and specific scar characteristics that may be more troublesome than others from the patients' point of view.

The patient part involves the scar characteristics pain, itch, color (a combination of vascularization and pigmentation), thickness, relief (surface roughness) and pliability (stiffness). The parameters pain and itch measure the extent to which the donor-site scars have been painful or itching over the past few weeks. The parameters color, thickness and pliability describe the patients' judgement of whether the color, thickness, and stiffness of the donor-site scar differed from the normal skin. The parameter relief, which includes the surface roughness of the donor-site scar area, described the presence of surface irregularities. The observer scale includes vascularization, pigmentation, thickness, surface, relief and pliability. Each of the six scar characteristics was rated on a 10-point scale, which ranges from 1 (comparable to normal skin) to 10 (very different from normal skin).

In addition, patients and observers complete one item to measure their overall opinion on the donor-site scar, ranging from 1 (best scar imaginable) to 10 (worst scar imaginable).

STATISTICAL ANALYSIS

Characteristics

Descriptive statistics were used to present data on patient characteristics (age, sex, and Fitzpatrick skin type), clinical characteristics (cause of the burn, percentage total body surface area (TBSA) burned, percentage TBSA excised, and length of stay) and donor-site characteristics (anatomical location of the donor-site, >2 weeks to re-epithelization, infection).

Changes of scar quality over time

The Wilcoxon signed-rank test was used to analyze differences per POSAS item between 3 and 12 months. A p -value <0.05 was considered significant. Additionally, the effect size was calculated by dividing the Z-value by the square root of number of cases[14]. An effect size above 0.5 was considered as a 'large effect', between 0.3 and 0.5 as a 'moderate effect' and between 0.1 and 0.3 as a 'small effect' and beneath 0.1 'trivial'. In other words, the larger the effect size, the greater the change in scar quality[14].

Agreement between patients and observers

The agreement between patients and observers on the POSAS items was assessed in three ways. First, the inter-observer reliability (IOR) was used to determine the agreement between the patient-ratings and observer- ratings regarding the different items on quality of the donor-site scars. The IOR was expressed as intra-class correlation coefficient (ICC) for the POSAS items, including their 95% confidence intervals (CI), and calculated using a two-way mixed random effect model for single measures consistency[15,16]. The ICC was calculated for the corresponding items (color, vascularity, pigmentation, relief and thickness) and overall opinion item. To be able to compare the color assessments of the patients with the color assessments of the professionals, an average score for the items vascularity and pigmentation was calculated. ICC-values range from no agreement (0) to perfect agreement (1)[17,18]. ICC-values beneath 0.4 are considered as 'poor agreement', between 0.4 and 0.6 as 'moderate agreement', between 0.6 and 0.8 as 'good agreement' and above 0.8 as 'very good agreement'[17,18]. Second, the 95% limits of agreement approach (Bland & Altman plots) was used to assess the score agreement between the patients' and observers' judgement of the overall opinion[19,20]. Third, a sub analysis on the corresponding POSAS items was performed to analyze whether the patients judged their scars more severe, identical or less severe compared to the observers. All data analyses were performed using statistical software (IBM, SPSS, V.24.0).

RESULTS

Out of 113 eligible patients, 80 patients were included in the study. Participants had a median age of 34 (range 0-84) and most were male (65.0%)). Median percentage TBSA burned was 6.0 (range 0.5-55) and most donor-sites were placed on the thigh (73.8%) (Table 1). During the three-month follow-up, 73 patients (91%) completed the POSAS. During the twelve-month follow-up, 72 patients (90%) completed the POSAS (Figure 1). A non-response analysis showed no differences in age, gender and %TBSA burned.

Table 1. Baseline characteristics

Demographics	**Total population (n=80)**
Age, median (range)	34 (0-84)
Male, n(%)	52 (65.0)
Skin type [FP >3], n (%)	21 (26.4)
Clinical characteristics	
Cause of the burn	
Scald (%)	19 (23.8)
Flame/fire (%)	39 (48.8)
Hot fat (%)	4 (5.0)
Other (%)	18 (22.5)
%TBSA burned, median (range)	6 (0.5-55)
% TBSA excised, median (range)	2 (0.5-50)
LOS in days, median (range)	16 (0-94)
Donor-site characteristics	
Location of the donor-site	
Thigh (%)	59 (73.8)
Other (%)	21 (26.4)
Time to re-epithelization (>2 weeks), n (%)	58 (74.4)
Wound infection, n (%)	7 (8.8)

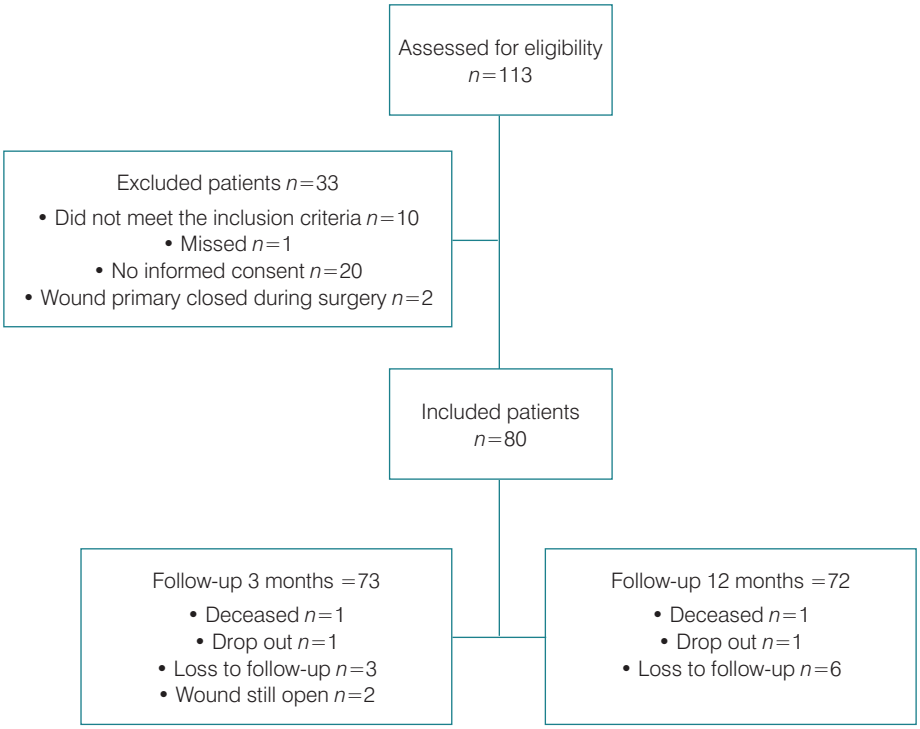


Figure 1. Inclusion flowchart

Scar quality and changes over time

At 3 months post-surgery, the patients rated the scar characteristics color and itch most severe (Figure 2). The mean scores for color and itch significantly decreased (e.g. improved) between 3 and 12 months (3.1 *vs*. 1.5 and 5.0 *vs*. 3.5, respectively both $p<0.001$) with an effect size of 0.53 and 0.48, indicating a moderate effect. No significant change was seen for the other items (Figure 2, Supplementary Digital Content 1), but the mean overall opinion of the patients significantly decreased from 3.9 to 3.2 (effect size 0.49, $p<0.001$).

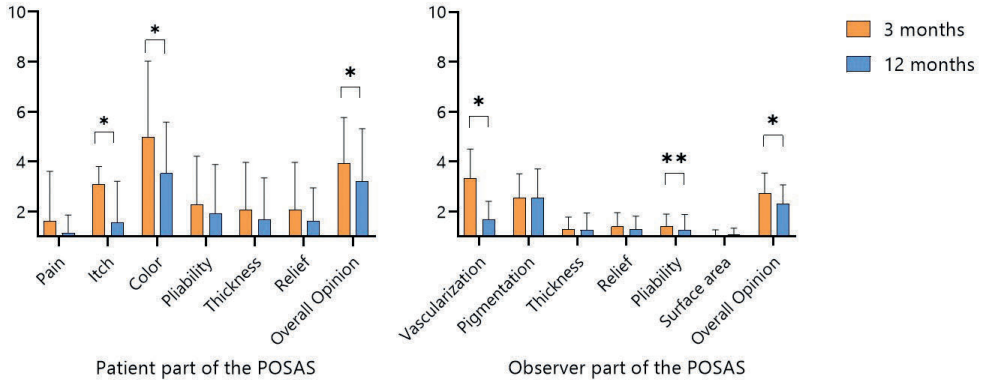


Figure 2. Patients and Observers (i.e. caregivers) scar quality scores at 3 and 12 months post-surgery.

Bars represent the mean item score and standard deviation. A lower POSAS score correlates with a better scar; a score of 10 reflects the worst imaginable scar. A score of 1 means no difference from normal skin. *p<0.001, ** p<0.05 (Wilcoxon signed-rank test).

Agreement between patients and observers

Table 2 shows the agreement in terms of inter-observer agreement between patients and observers on the corresponding POSAS items. The agreement on the items pliability, thickness and relief increased between 3 and 12 months. Agreement on POSAS items were all poor, however at 12 months at best for the item color/pigmentation (0.38, 95% CI 0.16-0.56).

Table 2. Inter-observer reliability between patients and caregivers

POSAS items	3 months ICC (95% CI)	12 months ICC (95% CI)
Color/vascularity	0.29 (0.07 – 0.49)	0.14 (-0.09 – 0.36)
Color/pigmentation	0.05 (0.28 – 1.11)	0.38 (0.16 – 0.56)
Color/combination*	0.08 (0.30 – 1.17)	0.20 (-0.03 – 0.41)
Pliability	0.30 (0.08 – 0.50)	0.31 (0.09 – 0.54)
Thickness	0.19 (-0.05 – 0.40)	0.36 (0.15 – 0.55)
Relief	0.23 (-0.01 – 0.44)	0.36 (0.14 – 0.54)
Overall opinion	0.20 (-0.02 – 0.41)	0.24 (0.02 – 0.45)

POSAS: Patient and Observer Scar Assessment Scale;

*Combination: Average score of vascularity and pigmentation

ICC-values range from no agreement (0) to perfect agreement (1) [17 18]. ICC-values beneath 0.4 are considered as 'poor agreement', between 0.4 and 0.6 as 'moderate agreement', between 0.6 and 0.8 as 'good agreement' and above 0.8 as 'very good agreement' [17 18]

The limits of agreement approach showed that 95% of the patient overall opinion item differed up to 4.74 points from the observers with a systematic difference of 1.05 at 3 months and 0.89 at 12 months, indicating a slightly better agreement between the assessors at twelve months. Both plots visualize that the difference between the patients and observers tends to get larger as the average overall opinion score increases (i.e. if the overall opinion gets worse) (Figure 3).

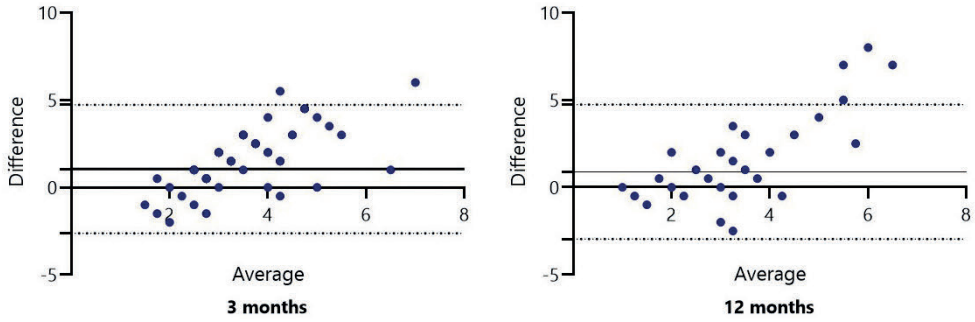


Figure 3. Bland and Altman plots demonstrating the agreement of POSAS overall opinion scores between patients and caregivers at three (left) and twelve (right) months post-surgery.

Each dot in the figure represents a donor site scar judged by the patient and observers. The difference between the two ratings is on the Y-axis and the average of both ratings on the X-axis. The dotted lines represent the 95% confidence intervals (limits of agreement -2.63 to 4.74 and -2.95 to 4.74) and the black line the mean difference between the raters (mean difference 1.055 and 0.892).

Figure 4 presents the agreement in terms of the proportion of patients that score their scar more severe, identical or less severe compared to the observer. Patient and observer scar assessment scores of the items relief and pliability were identical in more than 50% of the cases, whereas the agreement on color (vascularity and pigmentation) was below 20% at both time points. Overall, patients rated the scar characteristic color (pigmentation and vascularization) and their overall opinion on the scar more severe than observers. If patients rated their scar less severe than observers, the mean difference was less than one point for all items (Supplementary Content 2). If patients rated the scar more severe than observers, the mean difference was more than 2 points for all items (Supplementary Content 2).

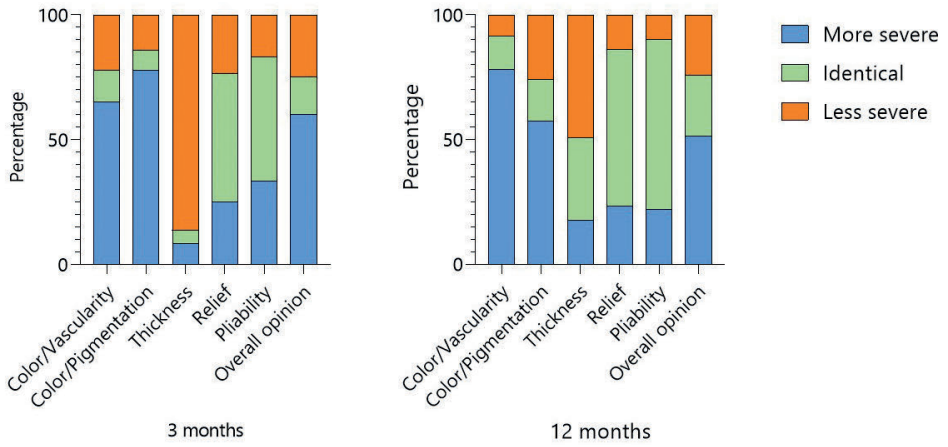


Figure 4. The agreement in terms of the proportion of patients that score their scar quality more severe, identical or less severe compared to the observer

DISCUSSION

To our knowledge, this is the first study that investigated patient- and observer-reported scar quality of donor-sites up to one year post-surgery. Patients' and observers' perceptions of scar quality only slightly improved during scar maturation. The agreement between patients and observers generally increased between 3 and 12 months, but remained 'poor' for all items of the POSAS. Results of this study indicate that caregivers seem to underestimate the impact of scars on patients. Especially the items on color were rated more severe by patients.

The magnitude of the observed improvement in scar quality over time was limited in our study population. Patient satisfaction regarding donor-sites may improve over time as a result of scar maturation but may on the other hand deteriorate as a result of psychological sequelae, especially when the recipient-site is completely healed and the patient might have expected that the donor-site scar would fade. The small changes over time that we observed in the perception of scar quality and satisfaction are consistent with previous studies on other scar types[7 9]. Our results show that items on color (color on the patients scale, vascularization and pigmentation on the observer scale) are rated the worst out of all items by both patients and observers at both time-points. Patients rated color less severe at 12 months post-surgery while observers only rated the item vascularization less severe at 12 months post-surgery. This might indicate that the appreciation of the item 'color' grows as a result of the reduction of erythema and that pigment changes contribute less to the improvement. This is in line with one study that measured erythema and melanin indexes of donor site scars with an objective measurement instrument

in patients with chronic leg ulcers[2]. That study found that the erythema index decreased 109% from 3 to 12 months post-surgery, whilst pigmentation only decreased 24% compared to normal skin. Studies that investigated patient- and observer-reported scar quality of burn wounds also found that items on color were rated worse than other items[7 21]. The severe scores and reduction over time that we found for patient-reported itching are also in accordance with previous studies regarding burn- and linear scars[7 12 13].

The items pliability, thickness and relief seem of less importance in our study population than to populations with burn scars or linear scars[7 12 13]. Nevertheless, it is of note that none of the POSAS items had a mean score of 1 (i.e. the same as normal skin) at 12 months post-surgery.

Patients and observers showed only poor to moderate agreement on scar quality at both time points. For the overall opinion on the scar, agreement was poor as well. Patients especially scored color worse than observers, whilst observers seemed to underline the importance of thickness. This might be due to the fact that hypertrophy of donor-site scars is not expected by the observers and may be seen as pathological scarring, whilst patients might compare the donor-site scar to their burn scar and consequently downgrade hypertrophy. The limited agreement on specific items that we found between patients and observers is consistent with studies that reported differences in scar appreciation between patients and observers on linear scars[5 22-24]. Eskes et al. only found a 'moderate' agreement on the overall opinion item of the POSAS on donor-site scars in a general trauma population and also a 'poor' agreement for all other items. However, they did not study the magnitude or direction of the differences between patients' and observers'[5]. Hoogewerf et al. studied the magnitude and direction of differences in assessment of facial burn scars for patients and observers in a Dutch burn population[4]. They found that 53% of the patients' and observers' assessments were identical. In our study population this was only 24% and 37%, at respectively 3 and 12 months. In other words; there seems to be a worse agreement in the judgment of donor-site scars compared to the recipient site.

Surgeons should be aware of the fact that patients might have different views on scar outcomes after split skin grafting. Patients are often well informed on the (development of) scars at the recipient site, but just have to deal with the emergence of a donor-site scar. Our results can be used to manage patients' expectations regarding donor-site scar quality after split skin grafting. Effective communication can improve patient satisfaction and outcome. However, it is difficult to predict psychological distress based on the severity of disfigurement. Therefore, future studies should investigate the relationship between donor-site scar quality and psychological distress or quality of life[25]. Results of our study can also be used as a starting point for scar quality improvement. Concerns about donor-site scarring may be more significant than surgeons might expect; the mean overall opinion of the patient on the donor-site scar was still 3.2 after one year (in comparison: two large cohort studies found that patients scored the overall opinion on

their burn scar 4.1, at 1 - ≥5 years post burn[7,21]). Many studies have been performed on donor-site management ranging from different types of wound dressings to more innovative (surgical) techniques, including harvesting of the skin from a different location (i.e. buttock or skull) or the use of other harvesting methods, like dermal grafting[26-29]. Caregivers should be aware of these options, which may increase scar quality. However, most studies on the outcomes of these treatment options focus on early and rapid re-epithelialization but lack data on patient reported (long term) scar appearance[26]. Further research is, therefore, necessary to investigate which patients benefit from these techniques.

This study has limitations that should be noted. The POSAS was used to assess scar quality. Although this is the only scar assessment scale that takes the opinion of the patient into account and has been validated for judgment of burn and linear scars, there has never been a reference or golden standard with regard to the quality of scars. In our study, this is of less importance because the patients' perception is the ultimate outcome to come to the best patient-centered care. Nevertheless, a minimal important change analysis is never done for the POSAS and it is therefore unknown if patients or observers judged the observed changes between 3 and 12 months important or meaningful. Another limitation is that this study is part of an explorative cohort study wherefore it was decided a priori to include patients over a one-year period. The accuracy of the agreement might be limited as illustrated by the wide confidence intervals and might be due to an insufficient sample size. However, the agreement between items was 'poor' at best. So even if the number of included patients would decrease confidence limits, this will probably not lead to a 'good' agreement.

CONCLUSION

The agreement on donor-site scar quality between patients and observers is limited. Surgeons should be aware that patients might have a different view on donor-site scars. This realization is important to manage patient expectations regarding scar quality after split skin grafting and pre-surgical counseling of patients with regard to anticipated anxiety about scar appearance and quality improvement. Effective communication may improve patient satisfaction.

Source of funding

C.M. Legemate received a grant from The Dutch Burns Foundation to support this study (grant number: 15.101). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

None

SUPPLEMENTARY CONTENT 1

POSAS item scores at 3 and 12 months post-surgery

Patients				
Scar characteristic	**Mean 3 months**	**SD**	**Mean 12 months**	**SD**
Pain	1,61	2	1,14	0,723
Itch	3,1	0,7071068	1,55	1,663
Color	4,97	3,05505	3,53	2,055
Pliability	2,28	1,944	1,92	1,97
Thickness	2,08	1,897	1,68	1,668
Relief	2,08	1,897	1,63	1,326
Overall Opinion	3,92	1,847	3,22	2,103

Observers				
Scar characteristic	**Mean 3 months**	**SD**	**Mean 12 months**	**SD**
Pain	1,61	2	1,14	0,723
Pigmentation	2,5479	0,96884	2,56	1,14774
Thickness	1,2945	0,49185	1,26	0,6895
Relief	1,3904	0,5728	1,28	0,54052
Pliability	1,3973	0,50662	1,2533	0,63869
Surface area	1,0548	0,21346	1,0667	0,26423
Overall Opinion	2,726	0,82091	2,32	0,75624

SUPPLEMENTARY CONTENT 2

Comparison of patients' scores and caregivers' scores on corresponding POSAS items

	3 months			12 months		
POSAS item	More severe n (%); D	Identical n (%)	Less severe n (%); D	More severe n (%); D	Identical n (%)	Less severe n (%); D
Color/vascularity[ab]	47 (65.3); 2.7	9 (12.5)	32 (22.3); 0.7	57 (78.0); 2.7	10 (13.7)	6 (8.2); 0.1
Color/pigmentation[ab]	56 (77.9); 3.3	6 (8.3)	10 (14.0); 0.8	42 (57.6); 2.5	12 (16.4)	19 (26.0); 0.5
Thickness[ab]	6 (8.4); 4.5	4 (5.6)	62 (86.2); 0.1	13 (17.7); 2.8	24 (32.9)	36 (49.4); 0.1
Relief[ab]	18 (25.2); 2.7	37 (51.4)	17 (23.6); 0.5	17 (23.4); 2.3	46 (63.0)	10 (13.7); 0.4
Pliability[ab]	24 (33.5); 2.2	36 (50.0)	12 (16.7); 0.5	16 (21.9); 3.5	50 (68.5)	7 (9.6); 0.2
Overall opinion[ab]	44 (60.3); 2.5	11 (15.1)	18 (24.7); 0.8	38 (51.4); 2.6	18 (24.3)	18 (24.4); 0.6

[*] D: mean difference between patients' and observers' scores (patient – observer)

[a] missing at 3 three months: color/vascularity (7), color/pigmentation (7), thickness (7), relief (7), pliability (7) overall opinion (7);

[b] missing at twelve months: color/vascularity (8), color/pigmentation (8), thickness (8), relief (8), pliability (8), overall opinion (8)

REFERENCES

1. Garcia E, Stone E, Chan LS, et al. Donor-site preferences in women during autologous skin grafting. *Plast Reconstr Surg* 2014;133(3):378e-82e. doi: 10.1097/PRS.0000000000000062
2. Danielsen PL, Jorgensen LN, Jorgensen B, et al. Erythema persists longer than one year in split-thickness skin graft donor sites. *Acta Derm Venereol* 2013;93(3):281-5. doi: 10.2340/00015555-1455 [published Online First: 2012/09/19]
3. Brown BC, Moss TP, McGrouther DA, et al. Skin scar preconceptions must be challenged: importance of self-perception in skin scarring. *J Plast Reconstr Aesthet Surg* 2010;63(6):1022-9. doi: 10.1016/j.bjps.2009.03.019 [published Online First: 2009/06/09]
4. Hoogewerf CJ, van Baar ME, Middelkoop E, et al. Patient reported facial scar assessment: directions for the professional. *Burns* 2014;40(2):347-53. doi: 10.1016/j.burns.2013.07.015
5. Eskes AM, Brolmann FE, van de Kar AL, et al. Values of patients and caregivers for donor site scars: an inter-observer analysis between patients and caregivers and prediction of cosmetic satisfaction. *Burns* 2012;38(6):796-801. doi: 10.1016/j.burns.2012.04.011
6. Dokter J, Vloemans AF, Beerthuizen GI, et al. Epidemiology and trends in severe burns in the Netherlands. *Burns* 2014;40(7):1406-14. doi: 10.1016/j.burns.2014.03.003
7. van der Wal MB, Vloemans JF, Tuinebreijer WE, et al. Outcome after burns: an observational study on burn scar maturation and predictors for severe scarring. *Wound Repair Regen* 2012;20(5):676-87. doi: 10.1111/j.1524-475X.2012.00820.x
8. Bond JS, Duncan JA, Mason T, et al. Scar redness in humans: how long does it persist after incisional and excisional wounding? *Plast Reconstr Surg* 2008;121(2):487-96. doi: 10.1097/01.prs.0000299183.88334.37 [published Online First: 2008/02/28]
9. Bond JS, Duncan JA, Sattar A, et al. Maturation of the human scar: an observational study. *Plast Reconstr Surg* 2008;121(5):1650-8. doi: 10.1097/PRS.0b013e31816a9f6f [published Online First: 2008/05/06]
10. Draaijers LJ, Tempelman FRH, Botman YAM, et al. The patient and observer scar assessment scale: A reliable and feasible tool for scar evaluation. *Plastic and Reconstructive Surgery* 2004;113(7):1960-65. doi: 10.1097/01.prs.00012207.28773.56
11. 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(1):6-18. doi: 10.1016/j.burns.2011.09.021
12. 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(2):514-22.
13. Truong PT, Lee JC, Soer B, et al. Reliability and validity testing of the Patient and Observer Scar Assessment Scale in evaluating linear scars after breast cancer surgery. *Plast Reconstr Surg* 2007;119(2):487-94. doi: 10.1097/01.prs.0000252949.77525.bc

14. Cohen J. Statistical power analysis for the behavioral science: Elsevier 1988.
15. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86(2):420-8.
16. McGraw KOW, S.P. . Forming Inferences about Some Intraclass Correlation Coefficients *Psychological Methods* 1996;1(1):30-46.
17. Durani P, McGrouther DA, Ferguson MW. Current scales for assessing human scarring: a review. *J Plast Reconstr Aesthet Surg* 2009;62(6):713-20. doi: 10.1016/j.bjps.2009.01.080
18. Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. *Stat Med* 1994;13(23-24):2465-76.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307-10.
20. de Vet HC, Terwee CB, Knol DL, et al. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59(10):1033-9. doi: 10.1016/j.jclinepi.2005.10.015
21. Spronk I, Polinder S, Haagsma JA, et al. Patient-reported scar quality of adults after burn injuries: A five-year multicenter follow-up study. *Wound Repair Regen* 2019 doi: 10.1111/wrr.12709
22. Kallidonis P, Kyriazis J, Kamal W, et al. Hybrid laparoendoscopic single-site surgery of upper urinary tract with the use of mini-laparoscopic instruments: cosmetic outcome and midterm oncological outcome. *World J Urol* 2016;34(9):1221-8. doi: 10.1007/s00345-016-1761-2 [published Online First: 2016/01/20]
23. Gencdal S, Aydogmus H, Aydogmus S, et al. Mini-Laparoscopic Versus Conventional Laparoscopic Surgery for Benign Adnexal Masses. *J Clin Med Res* 2017;9(7):613-17. doi: 10.14740/jocmr3060w [published Online First: 2017/06/15]
24. Pini G, Goezen AS, Schulze M, et al. Small-incision access retroperitoneoscopic technique (SMART) pyeloplasty in adult patients: comparison of cosmetic and post-operative pain outcomes in a matched-pair analysis with standard retroperitoneoscopy: preliminary report. *World J Urol* 2012;30(5):605-11. doi: 10.1007/s00345-011-0740-x [published Online First: 2011/08/24]
25. Rumsey N, Clarke A, White P. Exploring the psychosocial concerns of outpatients with disfiguring conditions. *J Wound Care* 2003;12(7):247-52. doi: 10.12968/jowc.2003.12.7.26515 [published Online First: 2003/08/05]
26. Serebrakian AT, Pickrell BB, Varon DE, et al. Meta-analysis and Systematic Review of Skin Graft Donor-site Dressings with Future Guidelines. *Plast Reconstr Surg Glob Open* 2018;6(9):e1928. doi: 10.1097/GOX.0000000000001928 [published Online First: 2018/10/24]
27. Radharaman, Kumar P, K SA, et al. The role of recruited minced skin grafting in improving the quality of healing at the donor site of split-thickness skin graft-A comparative study. *Burns* 2019;45(4):923-28. doi: 10.1016/j.burns.2018.11.018 [published Online First: 2019/01/22]
28. Lindford AJ, Kaartinen IS, Virolainen S, et al. The dermis graft: another autologous option for acute burn wound coverage.

Burns 2012;38(2):274-82. doi: 10.1016/j.burns.2011.08.009 [published Online First: 2011/09/10]

29. Legemate CM, Lucas Y, Oen I, et al. Regrafting of the Split-Thickness Skin Graft Donor-Site: Is It Beneficial? *J Burn Care Res* 2020;41(1):211-14. doi: 10.1093/jbcr/irz166 [published Online First: 2019/10/21]

11

Patient-reported scar quality of donor-sites following split-skin grafting in burn patients: Long-term results of a prospective cohort study

C.M. Legemate, P.J. Ooms, N. Trommel, E. Middelkoop, M.E. van Baar, H. Goei, C.H. van der Vlies

Burns. 2021 Mar;47(2):315-321

ABSTRACT

Background

Skin grafting is the current gold standard for treatment of deeper burns. How patients appraise the donor-site scar is poorly investigated. The aim of this study was to evaluate long-term patient-reported quality of donor-site scars after split skin grafting and identify possible predictors.

Methods

A prospective cohort study was conducted. Patients were included in a Dutch burn centre during one year. Patient-reported quality of donor-site scars and their worst burn scar was assessed at 12 months using the Patient and Observer Scar Assessment Scale (POSAS). Mixed model analyses were used to identify predictors of scar quality.

Results

This study included 115 donor-site scars of 72 patients with a mean TBSA burned of 11.2%. The vast majority of the donor-site scars (84.4%) were rated as having at least minor differences with normal skin (POSAS item score ≥2) on one or more scar characteristics and the overall opinion on 80.9% of the donor-site scars was that they deviated from normal skin 12 months after surgery. The overall opinion on the donor-site scar was 3.2 ± 2.1 vs. 5.1 ± 2.4 on the burn scar. A younger age, female gender, a darker skin type, and location on the lower leg were predictors of reduced donor-site scar quality. In addition, time to re-epithelization was associated with scar quality.

Conclusion

This study provided new insights in long-term scar quality of donor-sites. Donor-site scars differed from normal skin in a large part of the population 12 months after surgery. Results of this study can be used to inform patients on the long-term outcomes of their scars and to tailor preventive or therapeutic treatment options.

BACKGROUND

In present day burn care, excision and skin grafting is the cornerstone in the treatment of deeper burns to facilitate wound healing and provide a good functional and aesthetic scar outcome.[1-3]

On the one hand, skin grafting offers an important therapeutic option in the treatment of burn wounds. On the other hand, donor-sites that remain after skin grafting form scars, which may heal aesthetically displeasing with noticeable depigmentation and hypertrophy.[4-6] Scars of the donor-sites are rectangular, linear-shaped and commonly placed on the patients' thigh, arms or back. Patients just have to accept this extra scar whilst it may have an impact on their quality of life.[4 7]

The incorporation of patients' values and opinions is endorsed to ensure high-quality patient-centred care.[8-10] Although scar quality is one of the most important outcomes in burn surgery today, there is no evidence to support therapeutic decision-making regarding skin grafting and expected donor-site morbidity. In massive burn injuries, donor-site scarring might be of limited importance. However, when treating smaller injuries, other treatment options might be considered if significant distress for the patient is expected after surgery.

Clinical observations at our institution have shown that caregivers seem to underestimate the impact of donor-site scarring on patients.[11] Therefore, the main aim of this study was to evaluate long-term patient-reported scar quality of donor-sites one-year post-surgery. Our secondary aim was to identify factors related to patient-reported scar quality of donor-sites in burn patients.

METHODS

Design and Participants

The present study is part of an observational prospective cohort study. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were adhered to in this study and manuscript. Patients of all ages who underwent excision and split-skin grafting for a burn wound between February 2017 and February 2018 in the burn centre of the Maasstad Hospital in Rotterdam were asked to participate. Patients were included if they were able to comply with the study protocol and signed informed consent. A maximum of 3 donor sites per patient were included. The study was conducted according to the principles of the Declaration of Helsinki and Dutch laws and approved by the regional Ethics Committee (reference number L2016119).

Treatment

Skin grafts were harvested at a depth of 0.2mm (0.007 inch) with an electric Aesculap® dermatome. Adrenaline soaked gauzes were placed on the wounds immediately after grafting to reduce blood loss. Afterwards, donor site wounds were covered with an alginate dressing, cotton wool and elastic bandages, which were removed 2 weeks post-surgery.

Scar quality assessment

Scar quality was assessed at 12 months post burn in the outpatient clinic. The patient part of the Patient and Observer Scar Assessment Scale (POSAS) version 2.0 was used to assess the scar quality of their donor sites and of the burn scar that they indicated as most severe. The patient scored the items pain, itch, color, pliability, thickness, and relief. All items were scored on a 10-point rating scale. A lower score correlates with a better scar, where 1 resembles 'normal skin' and 10 resembles 'very different from normal skin'. The mean POSAS score was calculated by summing up the six item scores and dividing this by 6. Furthermore, patients were asked to give their overall opinion of the scar on a scale from 1 (best scar imaginable) to 10 (worst scar imaginable). The outcomes of the POSAS were divided into 3 categories: (1) low score, no differences with normal skin: POSAS item score 1; (2) intermediate scores, minor differences with normal skin: POSAS item score 2 or 3; (3) high scores, major differences with normal skin: POSAS item score ≥ 4. These cut-off points are arbitrary in the absence of commonly used cut-off points and in the absence of a minimal important change analysis of the POSAS [12].

Other study parameters

Other study parameters were documented during admission, surgery and outpatient visits. These were patient characteristics: age at surgery, gender, skin type, diabetes yes/no and smoking yes/no. Registered clinical characteristics were burn-related: % total burned body surface area (TBSA), % TBSA excised, length of stay, POSAS of the burn scar, and donor site-related: location on the body, location in relation to the burn wound, surface area, >2 weeks to re-epithelization, application of pressure garment and application of silicone gel.

Statistical analysis

We compared the main baseline characteristics of participants and nonparticipants to determine if there were any relevant differences between the groups using the independent t-test or Mann Whitney U tests (for continuous variables) and chi^2 test (for categorical variables). Descriptive statistics were used to assess long-term scar quality and characterize patients with low and high POSAS scores. Pearson statistics were used to identify the correlation between patient rated POSAS scores of the donor-site scar and burn scar (i.e. recipient site scars).

Univariable and multivariable mixed model analyses were performed to determine the predictive value of patient-, clinical- and donor-site-related factors for the mean POSAS score and

mean overall opinion of the POSAS. Mixed model analysis was used to take into account the dependency of the multiple observations within the participants if more than one donor site per patient was included. Factors with univariable $p<0.20$ were selected for multivariable analyses. A backward selection procedure was used to obtain the final models for the outcomes, in which only variables with $p<0.10$ were selected. IBM SPSS Statistics 23 and STATA version 14 were used for the analysis.

RESULTS

A total of 114 patients were screened for eligibility during the study period. Of these, 106 patients were eligible to participate and 80 patients signed informed consent. At 12 months post-surgery, 7 patients were lost to follow-up and 1 patient deceased, resulting in a total study population of 72 patients with 115 donor site scars. Patients included in the analysis had a mean age of 37.4 ± 23.0 years, 23.8% were aged ≤16years, and most were male (65.3%) (Table 1). Most burns were caused by flames (51.4%). Mean %TBSA burned was 11.2 ± 11.4, mean length of hospital stay was 24.8 ± 23.2 days, and most participants had only 1 donor site (62.6%). Most donor-sites were placed on the patients' thigh (76.5%).

Donor site scar quality

The mean POSAS score (based on the six POSAS items) was 1.9 ± 1.2 (range 1.0 - 7.2) at one-year post-surgery. Eighteen patients (25.0%) scored all six items as 1, indicating that their donor site scar did not deviate from normal skin (all had 1 donor site scar). These patients had a mean age of 43.1 ± 24.6 years and most (64.3%) were male.

Thus, for the other donor-site scars (n=97, 84.3%), patients reported at least minor differences (i.e. POSAS item score ≥2) on one or more scar characteristics. Six patients (8.0%) with a total of 8 donor sites (6.1%) reported a relatively high POSAS score (i.e. POSAS item score ≥4) for all POSAS items). These patients had a mean age of 29.7 ± 23.9 years and most (87.5%) were female.

The item 'color' was appreciated worst; for 41% of the scars, major differences compared to normal skin were reported and for 43% of the scars minor differences were reported (Fig. 1). For the scar characteristics itch, pliability, thickness and relief 8-12% of the donor site scars were rated with high scores (POSAS item score ≥4), while 73-88% were rated with no differences compared to normal skin (POSAS item score =1). The lowest ratings were for the item pain; 97% of the scars were rated as 'no difference to normal skin', resulting in a mean score of 1.1 ± 0.6 (Fig. 1).

Table 1. Patient demographics and clinical data

Patient characteristics	**No. of patients (n = 72)**
Age, mean (SD, range)	37.43 (23.0, 0-84)
Gender: Male, n(%)	47 (65.3%)
Fitzpatrick skin type	
I	12 (10.4%)
II	65 (56.5%)
III	12 (10.4%)
IV	18 (15.7%)
V	7 (6.1%)
VI	1 (0.9%)
Diabetes, n(%)	6 (5.2%)
Smoking, n(%)	35 (30.4%)
Clinical characteristics	
Burn aetiology	
Flame	37 (51.4%)
Scald	18 (25%)
Other	17 (23.6%)
%TBSA burned, mean (SD, range)	11.2 (11.4, 0.1-55)
%TBSA excised, mean (SD, range)	6.2 (7.1, 0.1-50)
Length of stay (days), mean (SD)	24.8 (23.2)
Donor site characteristics	**No. of Donor sites (n = 115)**
Location, n(%)	
Upper back	1 (0.9%)
Upper arm	12 (10.4%)
Lower arm	1 (0.9%)
Thigh	88 (76.5%)
Lower leg	1 (0.9%)
Same limb as burn wound, n(%)	56 (48.7%)
Adjacent to burn wound, n(%)	39 (39%)
Surface (cm2), mean(SD)	167.5 (173.4)
Time to re-epithelization (>2 weeks), n (%)	28 (24.3%)
Wound infection, n(%)	8 (7.0%)
>1 time harvested, n(%)	3 (2.6%)
Application of pressure garment, n(%)	2 (1.7%)
Application of silicone gel, n(%)	19 (17.3%)

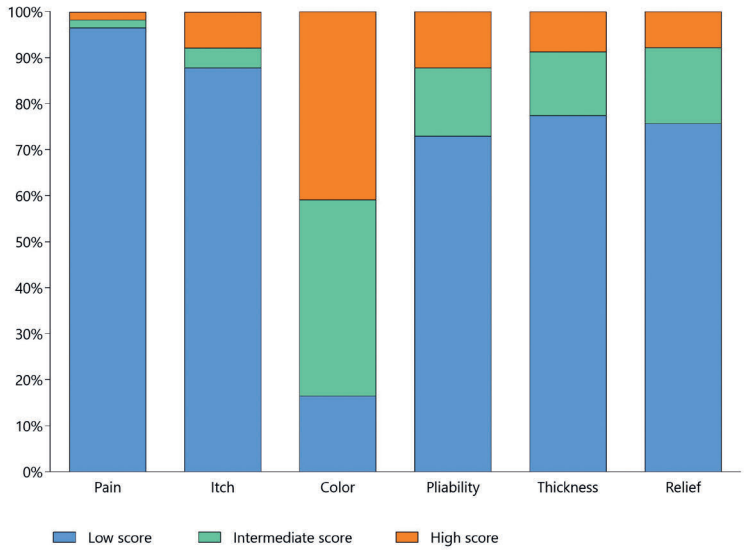


Figure 1. Proportion of donor sites for which patients scored low, intermediate, and high scores for scar-related problems on items of the patient part of the POSAS at 12 months post-surgery.

Low scores, no differences with normal skin; POSAS item score 1; intermediate scores, minor differences with normal skin: POSAS item score 2 or 3; high scores, major differences with normal skin: POSAS item score ≥4.

Patients' mean overall opinion of their donor site scars was 3.2 ± 2.2 (range 1-10) (Fig. 2). Twenty-two scars (19%, in 16 patients) were rated as 1 (i.e. 'best scar imaginable'). These patients had a mean age of 38.6 ± 24.6 years and most of these patients were male (81.3%). Thus, for all other scars (80.9%) at least minor dissatisfaction with the scar was reported. For 40 scars, 27 patients reported a relatively poor overall opinion (i.e. POSAS score ≥ 4). These patients had a mean age of 31.3 ± 21.3 years and 47.5% were male. In total, two patients rated 4 scars as 10 (i.e. 'worst scar imaginable). These patients were both female and had a mean age of 35.5 ± 13.4 years. Figure 2 shows the mean and standard deviation of the POSAS item scores of the donor-site scar and most severe burn scar (as indicated by the patient). The items 'pain' (1.1 ± 0.7 vs 1.9 ± 1.8), 'itch' (1.6 ± 1.7 vs 2.7 ± 2.3), 'color' (3.5 ± 2.1 vs. 5.2 ± 2.4), and 'overall opinion' (3.2 ± 2.1 vs. 5.1 ± 2.4) items differed least. The items 'pliability' (1.9 ± 2.0 vs. 4.2 ± 2.6), 'thickness' (1.7 ± 1.7 vs. 4.4 ± 2.8), and 'relief' (1.6 ± 1.3 vs 4.8 ± 2.6) differed most. All items had a very low or low ICC (Pearson's r <0.30).

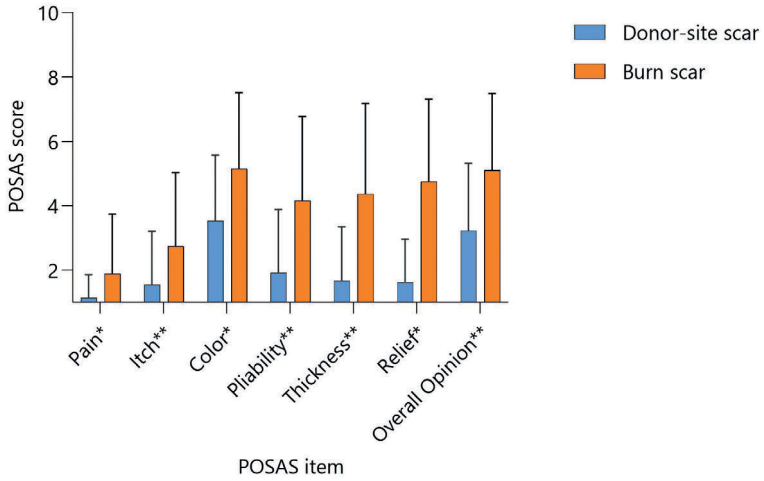


Figure 2. Patient reported POSAS scores of their donor-site and burn scar 12 months post-surgery.

A lower POSAS score correlates with a better scar; a score of 10 reflects the worst imaginable scar. *ICC<0.3 (very low), **ICC 0.3-0.5 (low).

Predictors of long-term donor-site scar quality

The results of univariable and multivariable mixed model analysis are shown in tables 2 and 3 respectively. In the final model, a higher age was associated with a better donor-site scar quality (i.e. a lower mean POSAS score (r=-0.01, SE=0.01; p=0.046)). Female gender (r=0.76, SE=0.27; p=0.004), a higher Fitzpatrick skin type (r=0.27, SE=0.13,; p=0.12) and time to re-epithelization exceeding 2 weeks (r=0.66, SE=0.26; p=0.016) were associated with a poorer scar quality (i.e. higher mean POSAS score).

For the overall opinion, a higher age was associated with a better score (i.e. lower POSAS score (r=0.02, SE=0.01; p=0.045). Female gender (r=1.40, SE=0.48; p=0.045), location on the lower leg (r=0.77, SE=0.43; p=0.077) and time to re-epithelization exceeding 2 weeks (r=0.79, SE=0.39; p=0.044) were associated with a poorer overall opinion on the donor-site scar. None of the clinical characteristics were associated with patient-reported donor-site scar quality at 12 months (Table 2).

Table 2. Univariable mixed model analysis of predictors of long-term donor-site scar quality

	Mean 6 item POSAS score			Overall opinion score		
Patient characteristics	**R coefficient**	**SE**	**p-value**	**R coefficient**	**SE**	**p-value**
Age (years)	-0.01	0.01	0.114	-0.02	0.01	0.158
Female gender	0.78	0.29	**0.008**	1.28	0.50	**0.010**
Fitzpatrick skin type	0.27	0.13	**0.042**	0.35	0.23	0.119
Diabetes	-0.38	0.63	0.554	0.43	1.09	0.694
Smoking	-0.58	0.31	0.061	-0.83	0.53	0.117
Clinical characteristics						
%TBSA burned	0.00	0.01	0.824	0.01	0.03	0.908
%TBSA excised	-0.01	0.02	0.526	-0.02	0.04	0.537
Total no. donor-sites	-0.18	0.35	0.604	0.34	0.60	0.569
Length of stay	0.01	0.01	0.695	0.01	0.01	0.496
Donor site characteristics						
Location - Body part						
Trunk	-0.07	0.65	0.911	-0.66	1.08	0.541
Upper arm	0.37	0.37	0.320	0.07	0.63	0.914
Lower arm	-0.28	0.65	0.665	-1.47	1.06	0.166
Upper leg	-0.08	0.21	0.689	-0.23	0.35	0.519
Lower leg	0.02	0.27	0.955	0.84	0.44	0.060
Location on same limb as burn wound (yes)	-0.20	0.24	0.401	-0.27	0.41	0.513
Location adjacent to burn wound (yes)	-0.08	0.24	0.750	-0.01	0.41	0.987
Surface	0.01	0.00	0.789	-0.03	0.01	0.722
Time to re-epithelization (>2 weeks)	0.64	0.24	**0.008**	0.75	0.41	0.066
Wound infection	0.66	0.40	0.105	0.93	0.68	0.172
>1 time harvested	0.01	0.01	0.391	0.01	0.01	0.441
Use of pressure garment	-0.06	1.22	0.959	1.31	2.07	0.527
Use of silicone gel	0.74	0.38	0.054	1.04	0.67	0.118

Table 3. Multivariable mixed model analysis of predictors of long-term donor-site scar quality

	Mean 6 item POSAS score*			Overall opinion score¥		
Patient characteristics	**R coefficient**	**SE**	**p-value**	**R coefficient**	**SE**	**p-value**
Age (years)	-0.01	0.01	0.046	-0.02	0.01	0.045
Female gender	0.76	0.27	0.004	1.40	0.48	0.004
Fitzpatrick skin type	0.21	0.12	0.067			
Donor site characteristics						
Location				0.77	0.43	0.077
Lower leg						
Time to re-epithelization (>2 weeks)	0.66	0.26	0.017	0.79	0.39	0.044

*Explained variance: 32.3%

¥Explained variance: 17.3%

DISCUSSION

This prospective cohort study assessed patient-reported quality of donor-site scars in a burn population one year after surgery. The majority of the scars (84.4%) were rated as having at least minor differences with normal skin (POSAS item score ≥2) on one or more scar characteristics. The overall opinion on the majority of the donor-site scars (80.9%) was that they deviated from normal skin.

The overall opinion of patients on their donor-site scar differed less than 2 points (POSAS 1-10 point scale) and patient-reported quality of burn scars and donor-site scars were not correlated, which might indicate that the individual opinion of the patient is of more importance than biological or genetic factors. A younger age, female gender and time to re-epithelization were associated with reduced scar quality (both mean POSAS item score and overall opinion on the scar). In addition, a darker skin was associated with a poorer scar quality (POSAS item score) and location on the lower leg was associated with a poorer overall opinion of the patient.

A former study from our research group found that the agreement on donor-site scar quality between patients and caregivers is poor and that caregivers seem to underestimate the impact of donor-site scars in – a subgroup of- patients. Many studies have been performed on donor-site management, ranging from different types of wound dressings to more innovative (surgical) techniques. However, patient-reported outcomes were hardly reported.[13] Our results show that location on the lower leg was a predictor of reduced patient satisfaction, which might be due to the fact that this area is more often visible than the upper leg. Harvesting of the skin from a different location (i.e. buttocks or skull) may lead to a less visible donor-site and might therefore be a relatively simple option to improve overall satisfaction of patients. The use of other harvesting methods, like dermal and minced skin grafting, have been described to

reduce donor-site morbidity.[13-16] Also, methods that aim to improve selective debridement (e.g. enzymatic or hydrosurgical debridement) of burn tissue may reduce the need for skin grafting and consequently, donor-site scarring.[17 18] If poor patient satisfaction regarding scar quality of a donor-site is expected, this might be an argument to support the decision to refrain from skin grafting. Local, pedicle and free flaps or the use of a skin stretching device for primary closure have been described as successful in the treatment of acute burn wounds and eliminate the need for donor-sites.[19-21] Another option, although costly and time consuming, is the use of allogenic skin substitutes or dermal regeneration products to support the wound environment and autologous regeneration in such way that skin grafting (and therefore donor-site scarring) may be reduced.[17 18 22] Conversely, if no problems regarding donors-site scar quality are expected, early debridement and skin grafting may lead to a decrease of the length of hospital stay.[23]

Articles that report donor-site scar quality are scarce. Most investigate difference in cosmetic outcome after the use of different types of wound dressings and only a few used patient-reported outcome measurement instruments.[13] Schulz et al. evaluated donor-site scar quality ≥2 years after application of Biobrane or Dressilk in 11 patients and found that patients reported all POSAS items ≤2 for their donor-site scar. These lower POSAS scores might indicate that donor-site scar quality improves after one year. On the other hand, the patients that they included in their study were older, no children were included and more males were included compared to our study population. Similar to our results, color was appreciated worst.[24]

To our knowledge, only two studies investigated the relationship between patient- and other clinical factors and patient-reported scar quality of donor-sites.[25 26] Karlsson et al. reported POSAS results 8 years post-surgery that were similar to our study results, but did not find a significant relationship between age, sex, healing time and patient-reported scar quality. However, they invited patients retrospectively, resulting in a study population of only 27 patients. McBride et al. studied patient reported donor-site scarring in children, but did not find a relationship with age or sex.[26] Studies that assessed predictors of patient-reported quality of scars after general surgical procedures and burn injuries have, in line with our study, reported female gender as a predictor for a worse scar outcome.[12 27 28] Wallace et al. hypothesized that immune and hormone responses might result in hypertrophic scarring in females.[29] Nevertheless, other studies on hypertrophic scars did not find female gender as an independent predictor.[30-32] Garcia et al. state that their clinical observations showed that female burn patients frequently have greater difficulty choosing a donor-site location and therefore conclude that scar outcome in females is more important than in men.[5] This finding is comparable with a previous study that described that women express greater concern with their appearance than men.[33] Moreover, many studies on health related quality of life after burn injury report female gender as a predictor of a reduced health related quality of life.[23] This supports the gender differences in the patients' opinion found in our study and suggest that this outcome might be based on culture rather than biological differences between males and

females. One study that used the patient scale of the POSAS to assess the quality of burn scars also found differences in age categories on the items pain, color, pliability and thickness.[28] It is important to realize that in children under the age of 5, parents complete the patient part of the POSAS. In literature, it has been stated that this may lead to underestimation of the true magnitude of the problem because pain and pruritus are difficult to assess through the parents.[30] On the other hand, parents may be very concerned about the appearance of the angular donor-site scars and how they evolve if their child grows and what they might think when they go into puberty.

An important strength of this study is that the study was conducted in a dedicated burn center, and thus reflects donor-site outcome after specialized (scar) treatment. Another asset of the study was the prospective design which is preferred for the development of association and prediction models.[34] Because of the strict study protocol and study conduct there were no missing values in the patient-, clinical and donor site characteristics. Although patients signed informed consent, they were not aware of the predictors that we aimed to investigate and could therefore not influence the outcome. This study also has some limitations. We used the POSAS to assess scar quality and used arbitrary cut-off points in the absence of a commonly used cut-off point or a minimal important change analysis of the POSAS. Nevertheless, the POSAS is the only validated scar outcome measure that takes the opinion of the patient into account.

CONCLUSION

This study provides important new insights in long-term scar quality of donor-sites as stated by burn patients. Even one year after surgery the mean overall opinion of patients on donor-site scars was remarkably high (POSAS score 3.2 (scale 1-10)). Moreover, 37% of the patients reported a poor overall opinion on the donor-site scar (i.e. POSAS score ≥ 4). Especially color of the donor site-scars was judged to remain deviant from normal skin. A younger age, female gender, a darker skin type, location on the lower leg and prolonged time to re-epithelization predict patient-reported reduced donor-site scar quality. Our study provides data that can help to better inform patients on the long-term outcome of their injury. Furthermore, preventive and therapeutic measures can be tailored to further improve long-term donor-site scar quality.

Acknowledgments

None

Conflict of interest

None

Funding

This research was financially supported by the Dutch Burns Foundation (grant number 15.101).

REFERENCES

1. Cubison TCS, Pape SA, Parkhouse N. Evidence for the link between healing time and the development of hypertrophic scars (HTS) in paediatric burns due to scald. *Burns* 2006;32(8):992-99. doi: 10.1016/j.burns.2006.02.007
2. Engrav LH, Heimbach DM, Reus JL, et al. Early excision and grafting vs. nonoperative treatment of burns of indeterminant depth: a randomized prospective study. *J Trauma* 1983;23(11):1001-4. doi: 10.1097/00005373-198311000-00007 [published Online First: 1983/11/01]
3. Ong YS, Samuel M, Song C. Meta-analysis of early excision of burns. *Burns* 2006;32(2):145-50. doi: 10.1016/j.burns.2005.09.005 [published Online First: 2006/01/18]
4. Brown BC, McKenna SP, Siddhi K, et al. The hidden cost of skin scars: quality of life after skin scarring. *Journal of Plastic Reconstructive and Aesthetic Surgery* 2008;61(9):1049-58. doi: 10.1016/j.bjps.2008.03.020
5. Garcia E, Stone E, Chan LS, et al. Donor-site preferences in women during autologous skin grafting. *Plast Reconstr Surg* 2014;133(3):378e-82e. doi: 10.1097/PRS.0000000000000062
6. Rotatori RM, Starr B, Peake M, et al. Prevalence and Risk Factors for Hypertrophic Scarring of Split Thickness Autograft Donor Sites in a Pediatric Burn Population. *Burns* 2019;45(5):1066-74. doi: 10.1016/j.burns.2019.02.003 [published Online First: 2019/03/03]
7. Coull A. Making sense of ... split skin graft donor sites. *Nursing times* 1991;87(40):52-4.
8. Bensing J. Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine. *Patient Education and Counseling* 2000;39(1):17-25. doi: 10.1016/s0738-3991(99)00087-7
9. 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). *Social Science & Medicine* 1997;44(5):681-92. doi: 10.1016/s0277-9536(96)00221-3
10. Draaijers LJ, Tempelman FRH, Botman YAM, et al. The patient and observer scar assessment scale: A reliable and feasible tool for scar evaluation. *Plastic and Reconstructive Surgery* 2004;113(7):1960-65. doi: 10.1097/01.prs.00012207.28773.56
11. Legemate CM, Ooms PJ, Trommel N, et al. Course of scar quality of donor-sites following split skin graft harvesting: comparison between patients and observers. *Wound Repair Regen* 2020 doi: 10.1111/wrr.12840 [published Online First: 2020/06/11]
12. Spronk I, Polinder S, Haagsma JA, et al. Patient-reported scar quality of adults after burn injuries: A five-year multicenter follow-up study. *Wound Repair Regen* 2019 doi: 10.1111/wrr.12709
13. Serebrakian AT, Pickrell BB, Varon DE, et al. Meta-analysis and Systematic Review of Skin Graft Donor-site Dressings with Future Guidelines. *Plast Reconstr Surg Glob Open* 2018;6(9):e1928. doi: 10.1097/GOX.0000000000001928 [published Online First: 2018/10/24]

14. Radharaman, Kumar P, K SA, et al. The role of recruited minced skin grafting in improving the quality of healing at the donor site of split-thickness skin graft-A comparative study. *Burns* 2019;45(4):923-28. doi: 10.1016/j.burns.2018.11.018 [published Online First: 2019/01/22]
15. Lindford AJ, Kaartinen IS, Virolainen S, et al. The dermis graft: another autologous option for acute burn wound coverage. *Burns* 2012;38(2):274-82. doi: 10.1016/j.burns.2011.08.009 [published Online First: 2011/09/10]
16. Legemate CM, Lucas Y, Oen I, et al. Regrafting of the Split-Thickness Skin Graft Donor-Site: Is It Beneficial? *J Burn Care Res* 2020;41(1):211-14. doi: 10.1093/jbcr/irz166 [published Online First: 2019/10/21]
17. Rosenberg L, Krieger Y, Bogdanov-Berezovski A, et al. A novel rapid and selective enzymatic debridement agent for burn wound management: a multi-center RCT. *Burns* 2014;40(3):466-74. doi: 10.1016/j.burns.2013.08.013
18. Hirche C, Citterio A, Hoeksema H, et al. Eschar removal by bromelain based enzymatic debridement (Nexobrid((R))) in burns: An European consensus. *Burns* 2017;43(8):1640-53. doi: 10.1016/j.burns.2017.07.025
19. Bloemen MC, van der Wal MB, Verhaegen PD, et al. Clinical effectiveness of dermal substitution in burns by topical negative pressure: a multicenter randomized controlled trial. *Wound Repair Regen* 2012;20(6):797-805. doi: 10.1111/j.1524-475X.2012.00845.x
20. Herndon DN, Spies M. Modern burn care. *Semin Pediatr Surg* 2001;10(1):28-31.
21. Verhaegen PD, Bloemen MC, van der Wal MB, et al. Skin stretching for primary closure of acute burn wounds. *Burns* 2014;40(8):1727-37. doi: 10.1016/j.burns.2014.03.014
22. Boyce ST, Kagan RJ, Yakuboff KP, et al. Cultured skin substitutes reduce donor skin harvesting for closure of excised, full-thickness burns. *Ann Surg* 2002;235(2):269-79. doi: 10.1097/00000658-200202000-00016
23. Spronk I, Legemate CM, Dokter J, et al. Predictors of health-related quality of life after burn injuries: a systematic review. *Crit Care* 2018;22(1):160. doi: 10.1186/s13054-018-2071-4
24. Schulz A, Rothermund I, Lefering R, et al. Long-term Scar Quality after Treatment of Standardized Partial-Thickness Skin Graft Donor Sites. *Adv Skin Wound Care* 2018;31(3):109-17. doi: 10.1097/01.ASW.0000527287.28216.65 [published Online First: 2018/02/14]
25. Karlsson M, Elmasry M, Steinvall I, et al. Scarring At Donor Sites After Split-Thickness Skin Graft: A Prospective, Longitudinal, Randomized Trial. *Adv Skin Wound Care* 2018;31(4):183-88. doi: 10.1097/01.ASW.0000530684.31491.5f [published Online First: 2018/03/22]
26. McBride CA, Kimble RM, Stockton KA. Prospective randomised controlled trial of Algisite M, Cuticerin, and Sorbact(R) as donor site dressings in paediatric split-thickness skin grafts. *Burns Trauma* 2018;6:33. doi: 10.1186/s41038-018-0135-y [published Online First: 2018/12/07]
27. de Vries CEE, Dekker AC, van Veen RN, et al. Trocar port scar quality in morbidly obese patients after bariatric surgery.

Surgery for Obesity and Related Diseases 2018;14(5):616-22. doi: 10.1016/j.soard.2018.01.035

28. Rashaan ZM, Kwa KAA, van der Wal MBA, et al. Patterns and predictors of burn scar outcome in the first 12 months after burn: The patient's perspective. *Burns* 2019 doi: 10.1016/j.burns.2019.03.025

29. Wallace HJ, Fear MW, Crowe MM, et al. Identification of factors predicting scar outcome after burn in adults: A prospective case-control study. *Burns* 2017;43(6):1271-83. doi: 10.1016/j.burns.2017.03.017

30. van der Wal MBA, Vloemans JFPM, Tuinebreijer WE, et al. Outcome after burns: An observational study on burn scar maturation and predictors for severe scarring. *Wound Repair and Regeneration* 2012;20(5):676-87. doi: 10.1111/j.1524-475X.2012.00820.x

31. Gangemi EN, Gregori D, Berchialla P, et al. Epidemiology and risk factors for Pathologic scarring after burn wounds. *Archives of Facial Plastic Surgery* 2008;10(2):93-102. doi: 10.1001/archfaci.10.2.93

32. Sood RF, Hocking AM, Muffley LA, et al. Race and Melanocortin 1 Receptor Polymorphism R163Q Are Associated with Post-Burn Hypertrophic Scarring: A Prospective Cohort Study. *Journal of Investigative Dermatology* 2015;135(10):2394-401. doi: 10.1038/jid.2015.197

33. Harris DL, Carr AT. Prevalence of concern about physical appearance in the general population. *British Journal of Plastic Surgery* 2001;54(3):223-26. doi: 10.1054/bjps.2001.3550

34. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162(1):W1-73. doi: 10.7326/M14-0698

12

General discussion & Future perspectives

Research in this thesis focused on patient reported outcome measures (PROMs) and clinical outcomes in burn care. An overview of available evidence on patient reported Health Related Quality of life (HRQL) after burn injury was provided (Part I), clinimetric properties of PROMs used in burn care were assessed (Part II) and clinical studies into outcomes of burn surgery were performed (Part III). The aim was to critically appraise available evidence of often used measurement instruments and to contribute to improved surgical burn care through the evaluation of surgical techniques.

PART I HEALTH-RELATED QUALITY OF LIFE AFTER BURN INJURY

Part I of this thesis is dedicated to patient reported HRQL after burn injury, which is essential to qualify the burden of burns in survivors. As a first step, we performed a systematic review to gain insight which instruments have been used to measure HRQL in burn patients. The review in Chapter 2 revealed 94 articles that studied outcomes of HRQL after burn injury in adults since 1975. Figure 1 shows how important and often used HRQL became as an outcome in burn research during the last years.

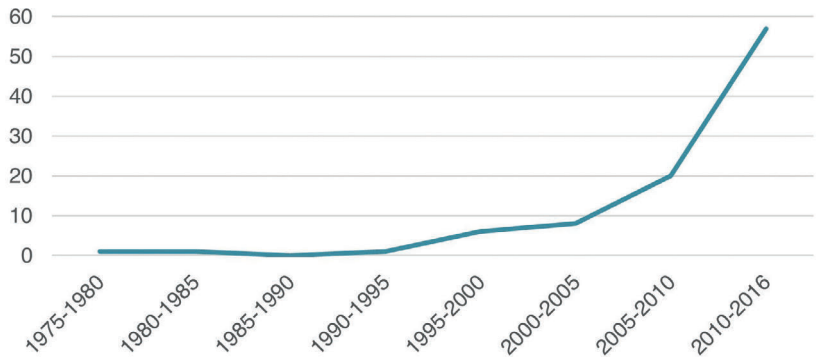


Figure 1. Number of studies (Y-axis) assessing HRQL of burn patients per year (X-axis).
(data derived from the search in Chapter 2 of this thesis (October 2016))

Results of the review showed that 20 different instruments were used to assess HRQL. In addition, we found a large variation in the use of disease-specific and/or generic instruments (26% of the studies used both) and time-assessment points (varying from pre-injury to >10 years). Also, studies had different ways of reporting of the results, including mean or median scores, domain scores versus total scores and 0-100 scores or standardized norm scores. This made it difficult to combine study results in order to study recovery patterns. Therefore, we choose to analyze recovery patterns of the three most applied instruments; the BSHS-B (46%), SF-36 (42%) and EQ-5D (9%). Results of the studies showed that most domains of HRQL were affected shortly

after injury but recovered over time, except for the domains work and heat sensitivity (BSHS-B), emotional role limitations and physical role limitations (SF-36), and pain/discomfort (EQ-5D). In the long-term, limitations of regular daily activities, including work, seemed to develop. We hypothesized that this is probably due to emotional problems as patients accomplish less and perform not as accurately as they were used to due to the injury.

The following step in optimizing treatment and rehabilitation of burn patients was to identify patients at risk of a decreased HRQL after burn injury. The systematic review in Chapter 3 analyzed 32 articles that studied and identified factors predicting HRQL after burn injury. Also in this review, we refrained from pooling of results due to the heterogeneity in used HRQL instruments, time points and data presentation. Furthermore, reported predictors and statistical reporting varied. Therefore, we were not able to take into account the sample size nor the reported strengths of predictors in the individual articles. Not entirely unexpected, indicators of burn severity and post-burn psychological factors were strong predictors of HRQL.

During the reviewing process we concluded that children with burn injuries form a distinct group in HRQL research. One important aspect of HRQL research in children is that, next to the child's perception, it often includes the parent's perception on a child's health status (a parent-proxy score). Chapter 4, therefore, provides a separate overview of study designs, instruments, methodological quality, outcomes and predictors of HRQL research in children after burn injury. As 27 studies were included, the relatively low number of studies compared to studies for adults found in Chapter 2 and Chapter 3 was striking. Again, we were unable to pool the data due to differences in instruments, data presentation, and the use of different norm scores. Similar to adults, most often used instruments were burn-specific: the Burns Outcome Questionnaire (BOQ) 0 to 4 years old (30%) and the BOQ 5 to 18 years old (41%). For the latter questionnaire, a child and parent version are available. An older age seemed related to a reduced HRQL. A possible explanation is that younger children are less affected by the trauma and adjust easier to their scars, disabilities and appearance as they grow older. Furthermore, indicators of burn severity and health status were found to predict HRQL in children. Assessment of individual data showed that HRQL improved over time and that parental proxy scores were in general comparable to children's self-ratings except for the domains 'appearance' and '(parental) concern'. It might be that parents are constantly reminded on the trauma by seeing the scar, which results in having more concerns. Also, parent views are influenced by general parental concerns, parental post-traumatic stress, and their perspectives on pre-burn functioning of their child. Presumably, patients do not see their scar constantly and might focus more on their progression during recovery. Compared to the results of Chapter 3, there was very low attention for psychological predictors in the studies found in children and no conclusion can be drawn on this aspect.

Clinical implications

Clinicians should be aware that both mental and physical support during the rehabilitation of burn patients is important. Also, a follow-up period of at least one year is necessary as some problems develop over time. Next to the severity of burns, HRQL seemed mostly affected by the psychological response of an individual to the trauma. In the Netherlands, the Burn centers Outcome Registry the Netherlands (BORN) has become part of standard burn care in 2018.[1] In 2020, this registration also started for pediatric burn patients. All admitted patients and/or caregivers are requested to participate, and a range of outcomes (including HRQL) are assessed during admission (pre-injury health status), one week, three months and twelve months after discharge from the burn center. Hereafter, annual follow-up will follow. Clinicians can use these data to identify patients at risk for a decreased HRQL and use this information to tailor their treatment.

In children, clinicians should be aware that special attention is needed for a child's experience on their appearance. Special attention is needed on parental concerns as many parents keep expressing their worries about their child's recovery years after injury, regardless of the child's age. Former studies have reported that family functioning, social support, and personality characteristics play an important role on child outcomes.[2, 3] As it seemed that children that were burned at an older age experience more problems, age specific treatment or prevention seems to be implicated. Hence, parental counseling and the inclusion of parents in the aftercare is important to improve HRQL measures in children, regardless of the child's age.

Future research

Future studies should focus on the development on a guideline on how to measure HRQL after burn injury world-wide, in both children and adults. This guideline should not only inform researchers on which measurement instrument should be used, but also on which time points the assessment should take place and how data should be presented. The first step to come to a worldwide consensus and a universally applied HRQL instrument for assessment in adults has recently been taken by an international group of clinicians and researchers.[4] They recommended a standardized assessment schedule of 4-6 weeks, 3 months, 6 months, 12 months and 24 months after burn injury, which is in accordance with the most frequently used time assessment points in research. Consensus on which instrument should be used could be achieved by an international Deplhi study: a method that is based on structured communication that relies on a panel of experts and patients, who are receiving questionnaires in several rounds.[5, 6] Standardization of reporting of results will make it easier to compare and combine study results into the development and predictors of HRQL, which is an important advantage as the study population of burn patients is often relatively small. Also, this will lead to a decreased risk of outcome reporting bias and would thereby contribute to the quality of future research and clinical burn care. Of course, this initiative should be rolled out for adults, adolescents and children.

Recently, a screening tool has been developed with the aim that it can be used in clinical practice to inform adult patients on expected HRQL outcomes and provide clinicians insights into the expected recovery of HRQL.[7] The aim is that, after further validation of the tool, clinicians may use patients characteristics and patient-reported data to identify patients who might benefit from additional rehabilitation care and timely interventions can be started to improve patient-centered care. The current tool is designed for adult burn patients, but should also be developed for children and adolescents to enhance future improvement of pediatric burn care.

PART II CLINIMETRIC STUDIES ON OUTCOMES

In Part II of this thesis, results of clinimetric studies into HRQL and scar quality, two often used PROMs in burn research, were performed. The reviews in Part I of this thesis show that there are many different PROMs used to evaluate HRQL in burn patients and that standardization of outcome measurement would be a step forward for future research on HRQL after burn injury. An important first step is to evaluate which instrument(s) should be recommended to assess HRQL. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative recommends that the selection should be based on 1) the quality of measurement properties of an instrument, as these are required to draw reliable and valid conclusions and 2) feasibility aspects, like access to the instrument and ease of use.[8, 9]

The quality of measurement properties of HRQL instruments validated for burn patients were evaluated using the COSMIN checklist in Chapter 5. None of the instruments provided enough evidence on their measurement properties to be highly recommended for routine use based on the COSMIN guidelines.[9] This is possibly due to the fact that only recently standards for the development of PROMs became available whilst the COSMIN guideline requires high standards of development and analyses and calls for distinct reporting of results. Therefore, we believe that some of the studies might have been of higher quality than rated in this review as a result of incomplete reporting, even though researchers may have performed extensive studies. According to the COSMIN guidelines, the first step in evaluating the development of a PROM is the question whether there is a clear description provided of the construct to be measured (i.e. a description of the different health aspects they purport to measure). Although HRQL is a widely used outcome, there is no definition or consensus available on which domains the outcome HRQL should cover.[4] Our review focused on instruments for which authors stated that they aimed to measure HRQL. Other PROMs that may assess only specific aspects of HRQL were therefore not included. For example, the Life Impact Burn Recovery Evaluation (LIBRE) profile that aims to measure social participation includes items on social role and personal relationships, which may also be important aspects of HRQL.[10] The COSMIN guidelines indicate that if no instrument can be recommended based on superior measurement properties, the

most important property of a measurement instrument is content validity, followed by structural validity and internal consistency.[11] The recently developed Brisbane Burns Scar Impact Profile (BBSIP) was the only instrument with good quality evidence for sufficient content validity as both professionals and patients were included in the development and state-of-the art qualitative data methods were used in line with the strict COSMIN guidelines.[12] However, it is important to note that this questionnaire focusses on the impact of burn scarring on HRQL while domains like work and daily activities or emotional reactions may also be influenced by other trauma-related factors and not all patients may (only) suffer from scarring. The BSHS-B was the only instrument with high-quality evidence for internal consistency, which is (together with structural validity) the second important measurement property according to the COSMIN guideline. Moreover, it seemed the most feasible instrument as it is relatively short and freely available in 14 languages whilst the BBSIP is only available in English. The review found good quality evidence for sufficient hypotheses testing for construct validity of the EQ-5D and SF-36, which suggests that these instruments can adequately determine differences between groups that differ in burn severity. Nevertheless, they miss important content that is relevant for patients after burn injuries as items specific for burns and related to scarring (self-esteem, stigmatization, physical appearance) are missing.

Another important aspect of a PROM is interpretability. Interpretability is not a measurement property, like validity and reliability, because it does not refer to the quality of an instrument.[13] It refers to what the score of an instrument means. In Chapter 6, the minimal important change (MIC) and minimal clinically important difference (MCID), two important interpretability aspects, of the patient part of the POSAS were explored. We were not able to determine the MIC, as the best cut-off values that we found for the differences that patients gave in POSAS scores between 3-6 and 6-12 months were not able to adequately differentiate between unchanged and improved scars. We hypothesize that patients may have found it difficult to remember what their scar was like 3 or 6 months earlier, even though we showed them pictures of the scar. Also, acceptance or other psychological factors may have influenced patients' POSAS ratings overtime. However, one important finding seemed to be that patients already appreciated a small change in POSAS as important. Of more importance in clinical research is what a difference in POSAS score means when results of a treatment are evaluated, which is expressed by the MCID. Using two different methods we found MCID values of -0.08 and -0.39. No guidelines for aggregation of these results exist.[14] Nevertheless, a MCID value of -0.08 seemed adequate to differentiate between patients that scored a scar as 'better' and patients that scored a scar as 'the same' or 'worse' (sensitivity 0.69, specificity 0.81). This difference in score is very small on the 1-10 POSAS scale and therefore small differences seem already important to patients.

Clinical implications

For now, it is recommended to use the BSHS-B to evaluate disease-specific HRQL in clinical burn research in adults based on the quality of measurement properties. The generic instruments EQ-5D and SF-36 can be best used to make comparisons with population norms or other patient groups. In children, none of the burn-specific instruments could be recommended based on the quality of the measurement properties or feasibility aspects and no evidence on measurement properties of generic instruments after burn injury exists. Therefore, it seems that the decision should be based on the research goal (evaluation of recovery/treatment or comparison with other patient groups) and the availability of the instrument (in the right language and/or if a permission fee is required).

The MCID of the POSAS is important to interpret results of clinical trials. However, it cannot directly be used to interpret differences between intervention groups in randomized clinical trial arms when baseline values are compared with improvement after a certain treatment to inform patients about what they might expect from a certain treatment in clinical practice.[15] The reason is that the results in each intervention group are presented as a mean value of all patients in that group, some may have indicated a larger difference than the MCID value, others may have smaller values. Therefore, former researchers have proposed a responder analysis.[16, 17] In a responder analysis, the percentage of patients in each arm of the trial that exceed the MCID value is calculated and these values are compared. In clinical practice, for example, one can tell a patient that about 75% of the patients experience a clinically relevant improvement after a given treatment. In this way, the MCID can be used to facilitate and improve shared decision making.

Future research

The first step to evaluate which instrument(s) should be recommended for use is to come to an overall consensus on the content that an instrument should include to comprehensively measure HRQL after burn injury. As mentioned in the discussion of part I, it is important that all relevant stakeholders (clinicians, researchers, but also patients and their caregivers) are included in this decision.[8, 11] If consensus is achieved, the overview provided in Chapter 5 could help to decide if a new instrument should be developed, or that more research into the measurement properties of already available instrument is needed in order to improve them before they can be recommended.

Currently, a new version of the POSAS, the POSAS 3.0, is being developed. The main reason for initiating this project was the lack of patient input in the development of the Patient Scale of the POSAS 2.0.[18] Also, Rasch analyses on the Patient Scale of the POSAS 2.0 suggested that the number of answering categories (10-point scale) might be reduced without losing information.[19] An important challenge for further research will be to discover the best way to measure different outcome domains of scar quality to guide treatment and rehabilitation strategies after burn injury.

PART III OUTCOMES OF BURN SURGERY

Part III focused on clinical outcomes of surgical debridement and split skin grafting, which is the standard of care for dermal burn wounds that are not expected to heal in 14-21 days. Over the last decade, hydrosurgery has become available in burn surgery as an alternative technique for tangential excision alongside the gold standard of conventional tangential excision by guarded knives. The retrospective analyses in Chapter 7 showed that the use of the VERSAJET™ hydrosurgery system (Smith+Nephew, London UK) in specialized Dutch burn care is substantial and that it is more often used in younger patients, scalds, burns on irregularly contoured body areas and in patients with a larger percentage total body surface area (%TBSA) burned. This is in line with our clinical experiences that 1) hydrosurgery seems more precise in debridement of burned tissue and therefore is used preferably in children and for burns on irregular surface areas like the face and hands 2) hydrosurgery is more often used in more superficial burns since full thickness burns are not easily debrided with hydrosurgery (scalds are more likely to be superficial than, for example, fire or chemical burns) and 3) hydrosurgery is regularly used next to conventional knives if the operating team consist of more than 1 surgeon, which is often the case in patients with a larger %TBSA burned. Hydrosurgery was used in 52% of the surgically treated patients despite little or no evidence of beneficial effects on related outcomes such as scar quality and substantial higher costs.[20, 21] One assumption of hydrosurgery is that it contributes to a more selective method of debridement with more preservation of viable dermis and therefore a better scar quality.[22, 23] Scar quality is the most relevant outcome for burn patients in the long run, especially in young patients, who will suffer from scar sequelae the rest of their lives. To compare the effect of hydrosurgery with conventional knife debridement in relation to dermal preservation and scar quality, we designed a study protocol for a randomized controlled trial (RCT) which is presented in Chapter 8. We choose an intra-patient design to overcome the effect of inter-patient differences on scarring (i.e. the presence of genetic risk factors for pathological scarring). The downside of this decision was that we were unable to conduct a cost-effectiveness study along the RCT. In addition to consecutive and comprehensive scar assessments, biopsies of the wound bed were taken before and after debridement to objectively measure dermal preservation by histology. To approach the full spectrum of scar quality evaluation we followed patients up to 1-year post-surgery. Results of the RCT are shown in Chapter 9. At 12 months, scars of the hydrosurgical debrided wounds were statistically significantly better in terms of the POSAS observer scores, POSAS patient scores and pliability measurements, whilst there was no difference in complications in wound healing after both treatments. The differences in POSAS scores in our study population were small, nevertheless, we determined that 48% of the patients rated the hydrosurgical debrided study area as better or much better 12 months post-surgery compared to 26% for conventional debrided study areas. Also, results of Chapter 6 support that the difference in patient POSAS scores can be interpreted as clinically important to patients. In line with the hypotheses of an increased scar quality when more dermis is preserved,

there was significantly more dermal tissue left after hydrosurgical debridement measured with histology. Although specialists in burns have long recognized the association between the depth of dermal injury and the degree of scarring this is, to our knowledge, the first study that proved a relationship between preservation of dermal tissue measured with histology and clinical better scar quality outcomes after burn surgery in humans. One of the strengths of this trial was the comprehensive inclusion criteria (patients of all ages and aetiologies) which allows application to a broad patient population with burns. Nevertheless, eligibility rates were lower than expected, as the amount of burns that would be suitable for hydrosurgical debridement was overestimated during the formation of the study protocol. Besides, one should keep in mind that dedicated burn surgeons determined which wound areas were randomized and therefore may have chosen wounds that seemed more suitable for hydrosurgical debridement than others. On the other hand, this mimics routine clinical practice and clinical decision-making.

Another important matter during split skin grafting is the side issue of a donor-site wound, and later scar, which is created when autologous healthy skin is harvested. A recent systematic review showed that studies that report donor-site scar morbidity are scarce and only a few studies used PROMs.[24] In 2021, a Delphi panel aimed to achieve consensus on the reduction in donor-skin (in percentages) that would be clinically meaningful to validate expensive treatment options like cellular and/or tissue products.[25] Unfortunately, the panel only included experts and patients were not asked for their opinion. In Chapter 10 and Chapter 11 results of a prospective cohort study in which patients were asked about their opinion on the scar quality of their donor-site are presented. We found that patients' perceptions on scar quality of donor-sites only slightly improved during scar maturation. One of the hypotheses for the small improvement was that scar quality improved over time as a result of scar maturation. On the other hand, the opinion of the patient regarding the scar may have deteriorated as a result of psychological sequelae, especially when the recipient site is completely healed and the patient might have expected that the donor-site scar would fade. Especially items on colour had a bad rating. Another important finding was that observers (i.e. caregivers/clinicians) seemed to underestimate the impact of donor-site scarring on patients. To improve management of patient expectations and pre-surgical counselling we explored patients' opinions on donor-site scars in relation to recipient-site scars (i.e. burn scars) and identified patients at risk for a low opinion on donor-site scars in Chapter 11. The mean overall opinion on scars of donor-sites measured with the POSAS was 3.2 (SD 2.1) compared to a mean overall opinion of 5.1 (SD 2.4) on the burn scar. This difference of less than 2 on a 1 – 10 point scale seems to indicate that the impact of donor-site scarring comes rather close to the impact of burn scars. From our clinical experience and the amount of attention that exists for burn scars in combination with the results of Chapter 10, clinicians need to be more aware of the morbidity of donor-site scarring in patients. A younger age, female gender, a darker skin type, and location on the lower leg were predictors of reduced donor-site scar quality. In addition, time to re-epithelization was associated with scar quality.

Clinical implications

Burn specialists seem to have a preference for specific patients and burn wounds on which they apply hydrosurgery. Although this is mostly based on the ease in use for particular burns, they also seem to keep the costs in mind. Hydrosurgery appeared to improve dermal preservation during debridement and resulted in a better scar quality after split skin grafting. Improved scar quality aspects were mostly related to pliability of the scar. During rehabilitation, scar therapy (like micro-needing, silicone cream, sheets and garment), is often necessary to improve pliability. This calls for repeated hospital visits and, in some cases, reconstructive surgery might even be necessary. If at least one of these treatments can be reduced in a patient, the costs of the VERSAJET™ hydrosurgery system (Smith+Nephew, London UK) of ±€140 per procedure may seem less important.[26, 27] Although surgical debridement and split skin grafting is still the cornerstone in treatment of deeper burns, surgeons should be aware that patients might have a different view on donor-site scars. This realization is important to manage patient expectations regarding scar quality after split skin grafting. A better management of expectations before surgery may improve patient satisfaction. If a poor patient satisfaction regarding a donor-site is expected, this might be an argument to support the decision to refrain from skin grafting, use more costly and/or time-consuming cellular or tissue products, or harvest skin from the scalp.[28] On the contrary, if no problems regarding donors-site scarring are expected, early excision and skin grafting leads to an early wound closure and a shorter stay in hospital which may lead to better quality of life outcomes and lower costs. Furthermore, identification of people at risk of reduced satisfaction regarding their donor-site may lead to tailored preventive and therapeutic measures and improve long-term satisfaction regarding donor-site scar quality.

Future research

To create robust guidelines for the approach of deeper burns there is a need for more high quality research on debridement techniques, in particular on alternatives to surgical knife debridement. Although RCTs are considered the highest in the order of levels of evidence, two recent reviews showed that only three RCTs with overall limited quality could be found to evaluate the advantages of the use of other techniques over tangential excision with a knife.[29, 30] Moreover, none of the studies addressed long term follow-up in terms of PROMs or scar outcomes. We experienced the performance of a RCT with a long follow-up period as cumbersome and time consuming for both patients and clinicians, but not impossible. We would like to encourage others to contribute to the evidence of debridement techniques by performing RCTs or large prospective cohort studies including PROMs with a follow-up period of at least one year to facilitate evidence-based decision making and patient-centered care in burn treatment. As mentioned before, burn surgeons tend to use hydrosurgery more often in more superficial wounds (like scalds) and children. To gain insight in the benefits of hydrosurgery for different patient categories, further analysis of our data could demonstrate if different patient or burn categories benefit more from hydrosurgical debridement compared to others.

In recent years, a growing trend has emerged towards more patient involvement in treatment decision making (shared decision making (SDM)). In the Netherlands, formal institutions such as the Dutch Healthcare Institute and Netherlands Federation of University Medical Centers (NFU) mention SDM as a key element of value based health care which aims to improve outcomes that matter to patients/populations while optimizing resource utilization.[31] One of the most important steps in the process of SDM is that the physician carefully explains the benefits and harms of all reasonable treatment options, including necessity of (early) surgery and donor-site morbidity. Future studies should examine 1) if clinicians currently apply SDM in clinical practice, 2) the preferences burn patients and parents have regarding involvement in treatment decision making (time of debridement, placement of the donor-site), and 3) if SDM leads to a higher degree of satisfaction with the decision-making process and final treatment.

FUTURE PERSPECTIVES

Standardization of outcome measurement

As pointed out in Part I, standardization of outcome measurement would be a big step forward for future research on burns. When all research groups measure the same outcomes, using the same validated instruments and thereby producing uniform results, researchers will be able to compare and combine their data. The development of a Core Outcome Set (COS), a scientifically agreed minimum set of the most important outcomes to be reported in all studies of a medical condition, is likely to provide the answer. Patients should have a genuine say in the development process otherwise there is a likelihood that the COS will omit important outcomes and that, ultimately, research will fail to give definitive information about whether treatments benefit patients or not. After the decision is made on 'what' to measure, consensus should be reached on 'how' to measure the outcomes, i.e. which outcome measurement instrument should be selected. This selection must be based on quality of measurement properties of the instruments (to gain valid and reliable results) and feasibility aspects (like availability of the instrument and ease of use).[8]

Recently, Young et al. initiated the development of a COS for burns, however, they restricted their focus to outcomes to be used in RCTs, which are not necessarily the outcomes used in daily burn care.[32] Especially to aid justification of more modern or various debridement tools, the inclusion of long-term validated scar outcome assessments and PROMs in future research is necessary. This is supported by several research initiatives that used modern state-of-the art methods, including patient participation, to develop PROMs on scarring.[12, 18] Furthermore, the International Consortium for Health Outcomes Measurement (ICHOM) is an organization which promotes the use of core outcome sets for various diseases. Burns is one of the 16 conditions and diseases that are considered as a topic for new standard sets.[33, 34]

Treatment strategies

The evidence presented in Part III of this thesis highlights the necessity for the burn community to consider a paradigm shift away from always reaching for the knife towards modern debridement approaches. Besides hydrosurgery, enzymatic debridement has been shown to be a new debridement technique comparable to knife debridement in terms of safety, efficacy and speed.[20, 29, 30, 35] More importantly is that both have the potential to improve dermal preservation and, therefore, long term scar outcomes. Nevertheless, use of novel tools is still not common practice for many burns facilities.[36] Challenges include feasibility of use, learning curves, available expertise and costs. Another topic of debate in burn treatment remains the timing of debridement.[37-40] Postponing excision of deep partial-thickness burns could allow time to assess burn depth, as superficial partial-thickness burns do not require skin grafting, but it may then result in increased risk of infection, increased length of stay and more extensive scarring. Novel debridement techniques contribute to novel perspectives in this debate. When the eschar of a deep partial thickness burn is removed with hydrosurgery, it is hypothesized that bacterial load is reduced and spontaneous healing with the use of biological dressings is optimized, potentially leading to a better outcome.[22, 41] Enzymatic debridement claims to only selectively debride burn eschar while viable dermal tissue remains unaffected during the debridement process and can therefore be used almost immediately after presentation to the burn center.[35] In this additional way, both techniques may minimize the risk of overzealous early excision, and subsequently poor scar quality outcomes, while allowing early burn management and a potential shorter hospital stay.

Patient participation

In 2018, the Burn centers Outcome Registry the Netherlands (BORN) started for adults, followed by the introduction of BORN-kids in the autumn of 2020.[1] Patients register several outcomes, like HRQL and burn scar quality at several moments during the first year after their injury, followed by an annual follow-up. This facilitates a comprehensive registration of the evolvement of short- and long-term consequences of burns. Patients are currently able to enter an online personal health environment platform that contains information and tools that can help them during their recovery including a decision aid for scar therapy to choose the treatment that best suits the patient's health status and perspectives. This tool uses data from research and the outcome registration. So, especially in aftercare patient involvement is getting more integrated. In acute burn care, however, many decisions are primarily made by the healthcare provider with limited consultation with the patient about their preferences, goals and values. An example of this is the decision between early skin grafting versus conservative treatment for intermediate depth burns. A recent systematic review provides no clear best treatment strategy.[40] Early skin grafting has the possibility of a shortened hospital stay and earlier rehabilitation, but the potential for worse scars (including an additional donor-site scar). Conservative treatment leads to a longer in-hospital stay and more painful bandage changes with the possibility of delayed surgery, but the

potential of less morbidity (no need for surgery) and less severe scars. In the future, integration of patient perspectives and clinical data should be used to uncover insights into patient-relevant outcomes of acute burn treatment. Especially for treatment options like early skin grafting versus conservative treatment for intermediate depth wounds and the use of novel techniques like tissue engineered skin constructs. The three Dutch Burn centers (Beverwijk, Groningen and Rotterdam) recently received a ZonMw grant to explore these topics. The aim is to create a decision aid to serve as a practical guide for patients and healthcare providers to understand and prioritize the preferences and goals of patients and to make sure that patients are involved in the decision about which treatment strategy to pursue (shared decision making). Ultimately, this should lead to improved patient-relevant outcomes and optimal value based acute burn care.

REFERENCES

1. Haanstra TM, van Baar ME, van Loey NEE, Middelkoop E, Nieuwenhuis MK, Pijpe A, van Schie C, Scholten-Jaegers SM, van der Vlies CH, Van Zuijlen PPM. THE BURN CENTRES OUTCOME REGISTRY THE NETHERLANDS (BORN): DEVELOPMENT AND FIRST EXPERIENCES. *Wound Repair and Regeneration* 2018;26(2): A22-A22.
2. Bakker A, Maertens KJ, Van Son MJ, Van Loey NE. Psychological consequences of pediatric burns from a child and family perspective: a review of the empirical literature. *Clin Psychol Rev* 2013;33(3): 361-371.
3. Egberts MR, van de Schoot R, Boekelaar A, Hendrickx H, Geenen R, Van Loey NE. Child and adolescent internalizing and externalizing problems 12 months postburn: the potential role of preburn functioning, parental posttraumatic stress, and informant bias. *Eur Child Adolesc Psychiatry* 2016;25(7): 791-803.
4. Edgar DW, Van Daele U, Spronk I, van Baar M, van Loey N, Wood FM, Kazis LB, Meirte J. Seeding the value based health care and standardised measurement of quality of life after burn debate. *Burns* 2020.
5. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med* 2011;8(1): e1000393.
6. Prinsen CA, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, Adams D, Terwee CB. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. *Trials* 2014;15: 247.
7. Spronk I, Van Loey NEE, Sewalt C, Nieboer D, Renneberg B, Moi AL, Oster C, Orwelius L, van Baar ME, Polinder S, Quality of life study g. Recovery of health-related quality of life after burn injuries: An individual participant data meta-analysis. *PLoS One* 2020;15(1): e0226653.
8. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials* 2016;17(1): 449.
9. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018;27(5): 1171-1179.
10. Ryan CM, Shapiro GD, Rencken CA, Griggs C, Jeng JC, Hickerson WL, Marino M, Goverman J, Kazis LE, Schneider JC, Board LA. The Impact of Burn Size on Community Participation: A Life Impact Burn Recovery Evaluation (LIBRE) Study. *Ann Surg* 2020;Publish Ahead of Print.
11. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, Bouter LM, de Vet HCW, Mokkink LB. COSMIN methodology for evaluating the content validity of patient-reported outcome

measures: a Delphi study. *Qual Life Res* 2018;**2**7(5): 1159-1170.

12. Tyack Z, Ziviani J, Kimble R, Plaza A, Jones A, Cuttle L, Simons M. 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(7): 1405-1419.

13. de Vet HCW TC, Mokknk LB, Knol DL. *Measurement in medicine. A practical guide to biostatistics and epidemiology* (1st edn). Cambridge Univ Press: Cambridge, 2011.

14. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, Croft P, de Vet HC. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010;63(5): 524-534.

15. Terwee CB, Peipert JD, Chapman R, Lai JS, Terluin B, Cella D, Griffith P, Mokkink LB. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Qual Life Res* 2021;30(10): 2729-2754.

16. Brozek JL, Guyatt GH, Schunemann HJ. How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. *Health Qual Life Outcomes* 2006;4: 69.

17. Schunemann HJ, Akl EA, Guyatt GH. Interpreting the results of patient reported outcome measures in clinical trials: the clinician's perspective. *Health Qual Life Outcomes* 2006;4: 62.

18. van Zuijlen PPM, Mokkink LB, Hoogewerf CJ, de Vet HCW. The official update of the POSAS: An invitation to share experiences to improve the POSAS in 'Project POSAS 3.0'. *Burns* 2017;43(4): 893-894.

19. van der Wal MB, Tuinebreijer WE, Lundgren-Nilsson A, Middelkoop E, van Zuijlen PP. Differential item functioning in the Observer Scale of the POSAS for different scar types. *Qual Life Res* 2014;23(7): 2037-2045.

20. Wormald JCR, Wade RG, Dunne JA, Collins DP, Jain A. Hydrosurgical debridement versus conventional surgical debridement for acute partial-thickness burns. *Cochrane Database of Systematic Reviews* 2020(9).

21. Kakagia DD, Karadimas EJ. The Efficacy of Versajet Hydrosurgery System in Burn Surgery. A Systematic Review. *J Burn Care Res* 2018;39(2): 188-200.

22. Cubison TC, Pape SA, Jeffery SL. Dermal preservation using the Versajet hydrosurgery system for debridement of paediatric burns. *Burns* 2006;32(6): 714-720.

23. Matsumura H, Nozaki M, Watanabe K, Sakurai H, Kawakami S, Nakazawa H, Matsumura I, Katahira J, Inokuchi S, Ichioka S, Ikeda H, Mole T, Smith J, Martin R, Aikawa N. The estimation of tissue loss during tangential hydrosurgical debridement. *Ann Plast Surg* 2012;69(5): 521-525.

24. Asuku M, Yu TC, Yan Q, Boing E, Hahn H, Hovland S, Donelan MB. Split-thickness skin graft donor-site morbidity: A systematic literature review. *Burns* 2021.

25. Gibson ALF, Smiell J, Yu TC, Boing EA, McClure EB, Merikle E, Holmes JHt. Determining clinically meaningful thresholds for innovative burn care products to reduce autograft: A US burn surgeon Delphi panel. *Burns* 2021;47(5): 1066-1073.

26. Hop MJ, Langenberg LC, Hiddingh J, Stekelenburg CM, van der Wal MB,

Hoogewerf CJ, van Koppen ML, Polinder S, van Zuijlen PP, van Baar ME, Middelkoop E. Reconstructive surgery after burns: a 10-year follow-up study. *Burns* 2014;40(8): 1544-1551.

27. Hop MJ, Polinder S, van der Vlies CH, Middelkoop E, van Baar ME. Costs of burn care: a systematic review. *Wound Repair Regen* 2014;22(4): 436-450.
28. Roodbergen DT, Vloemans AF, Rashaan ZM, Broertjes JC, Breederveld RS. The scalp as a donor site for skin grafting in burns: retrospective study on complications. *Burns Trauma* 2016;4: 20.
29. Edmondson SJ, Ali Jumabhoy I, Murray A. Time to start putting down the knife: A systematic review of burns excision tools of randomised and non-randomised trials. *Burns* 2018;44(7): 1721-1737.
30. Kwa KAA, Goei H, Breederveld RS, Middelkoop E, van der Vlies CH, van Baar ME. A systematic review on surgical and nonsurgical debridement techniques of burn wounds. *J Plast Reconstr Aesthet Surg* 2019;72(11): 1752-1762.
31. https://nfukwaliteit.nl/pdf/Bouwstenen_voor_werken_aan_waardegedreven_zorg.pdf.
32. Young A, Brookes S, Rumsey N, Blazeby J. Agreement on what to measure in randomised controlled trials in burn care: study protocol for the development of a core outcome set. *BMJ Open* 2017;7(6): e017267.
33. ICHOM. International Consortium for Health Outcome Measurement; 2019.
34. Porter ME, Larsson S, Lee TH. Standardizing Patient Outcomes Measurement. *N Engl J Med* 2016;374(6): 504-506.
35. Rosenberg L, Krieger Y, Bogdanov-Berezovski A, Silberstein E, Shoham Y, Singer AJ. A novel rapid and selective enzymatic debridement agent for burn wound management: a multi-center RCT. *Burns* 2014;40(3): 466-474.
36. Ziegler B, Fischer S, Pieper D, Mathes T, Kneser U, Hirche C. Evidence and Trends in Burn Wound Debridement: An Evidence Map. *Plast Surg (Oakv)* 2020;28(4): 232-242.
37. Wong L, Rajandram R, Allorto N. Systematic review of excision and grafting in burns: Comparing outcomes of early and late surgery in low and high-income countries. *Burns* 2021.
38. Kirn DS, Luce EA. Early excision and grafting versus conservative management of burns in the elderly. *Plast Reconstr Surg* 1998;102(4): 1013-1017.
39. Cubison TC, Pape SA, Parkhouse N. Evidence for the link between healing time and the development of hypertrophic scars (HTS) in paediatric burns due to scald injury. *Burns* 2006;32(8): 992-999.
40. Anna Miroshnychenko KK, Bram Rochwerg, Sophocles Voineskos,. Comparison of early surgical intervention to delayed surgical intervention for treatment of thermal burns in adults: A systematic review and meta-analysis. *Burns Open* 2021;5(2): 67-77.
41. Wood F, Martin L, Lewis D, Rawlins J, McWilliams T, Burrows S, Rea S. A prospective randomised clinical pilot study to compare the effectiveness of Biobrane(R) synthetic wound dressing, with or without autologous cell suspension, to the local standard treatment regimen in paediatric scald injuries. *Burns* 2012;38(6): 830-839.


Appendices
Summary
Nederlandse samenvatting
PhD portfolio
Dankwoord
About the author

APPENDICES

Appendices

Appendix 1. Search strategy

Embase.com ('thermal injury'/de OR 'heat injury'/de OR 'burn'/exp OR 'burn patient'/de OR 'burn scar'/de OR 'burn nursing'/de OR 'burn unit'/de OR (((burn OR burned) NOT burn*-out*) OR burns OR ((thermal OR heat) NEAR/3 injur*)):ab,ti) AND ('quality of life'/exp OR 'health status'/exp OR 'general health status assessment'/exp OR 'health status indicator'/de OR 'disability'/exp OR 'work resumption'/de OR 'return to work'/de OR 'sexual function'/de OR 'daily life activity'/exp OR 'chronic pain'/exp OR 'functional disease'/de OR 'functional assessment'/de OR 'disabled person'/de OR 'physical activity'/exp OR 'physical performance'/exp OR 'independent living'/exp OR 'social participation'/de OR fitness/de OR ((qualit* NEAR/3 life*) OR hrql OR hrqol OR ((health OR functional* OR physic*) NEAR/3 (stat* OR limitation* OR outcome* OR recover* OR impair* OR result* OR fitness*)) OR disab* OR invalid* OR ((walk* OR work* OR mobilit*) NEAR/3 (difficult* OR limit*)) OR (work NEAR/3 (resum* OR return OR back)) OR ((sexual* OR disease* OR assess* OR hand OR adaptat*) NEAR/3 function*) OR (dail* NEAR/3 (life OR living) NEAR/3 activit*) OR adl OR adls OR iadl OR iadls OR badl OR badls OR ((chronic* OR long*-term* OR longterm*) NEAR/3 (pain*)) OR (loss NEAR/3 function*) OR (burn NEAR/3 outcome* NEAR/6 questionnaire*) OR boq OR ((burn OR burns) NEAR/6 health NEAR/3 scale*) OR bshs OR euroqol OR sf36 OR sf-36 OR short-form-36 OR eq-5d OR eq-6d OR eq5d OR eq6d OR itqol OR (physical* NEAR/3 (activ* OR perform*)) OR (independent* NEAR/3 living) OR (social* NEAR/3 participat*)):ab,ti) AND ('cohort analysis'/exp OR 'longitudinal study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'validation study'/exp OR 'validity'/exp OR 'sensitivity and specificity'/exp OR 'clinical article'/de OR 'observational study'/de OR 'major clinical study'/de OR 'follow up'/de OR 'epidemiological data'/de OR 'case control study'/exp OR 'cross-sectional study'/exp OR (cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR validat* OR validit* OR (observation* NEAR/3 stud*) OR sensitiv* OR specific* OR 'follow* up*' OR followup* OR epidemiolog* OR 'case control*' OR 'cross-section*'):ab,ti) NOT ([animals]/lim NOT [humans]/lim) NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim) AND [english]/lim
Medline Ovid (exp "burns"/ OR "burn units"/ OR (((burn OR burned) NOT burn*-out*) OR burns OR ((thermal OR heat) ADJ3 injur*)).ab,ti.) AND ("Quality of Life"/ OR "health status"/ OR "Health Status Indicators"/ OR "Disability Evaluation"/ OR "Return to Work"/ OR exp "Activities of Daily Living"/ OR "chronic pain"/ OR "Disabled Persons"/ OR "Motor Activity"/ OR "Physical Fitness"/ OR ((qualit* ADJ3 life*) OR hrql OR hrqol OR ((health OR functional* OR physic*) ADJ3 (stat* OR limitation* OR outcome* OR recover* OR impair* OR result* OR fitness*)) OR disab* OR invalid* OR ((walk* OR work* OR mobilit*) ADJ3 (difficult* OR limit*)) OR (work ADJ3 (resum* OR return OR back)) OR ((sexual* OR disease* OR assess* OR hand OR adaptat*) ADJ3 function*) OR (dail* ADJ3 (life OR living) ADJ3 activit*) OR adl OR adls OR iadl OR iadls OR badl OR badls OR ((chronic* OR long*-term* OR longterm*) ADJ3 (pain*)) OR (loss ADJ3 function*) OR (burn ADJ3 outcome* ADJ6 questionnaire*) OR boq OR ((burn OR burns) ADJ6 health ADJ3 scale*) OR bshs OR euroqol OR sf36 OR sf-36 OR short-form-36 OR eq-5d OR eq-6d OR eq5d OR eq6d OR itqol OR (physical* ADJ3 (activ* OR perform*)) OR (independent* ADJ3 living) OR (social* ADJ3 participat*)).ab,ti.) AND ("validation studies"/ OR "Sensitivity and Specificity"/ OR "observational study"/ OR exp "Epidemiologic Studies"/ OR (cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR validat* OR validit* OR (observation* ADJ3 stud*) OR sensitiv* OR specific* OR "follow* up*" OR followup* OR epidemiolog* OR "case control*" OR "cross-section*").ab,ti.) NOT (exp animals/ NOT humans/) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt. AND english.la.

CINAHL EBSCOhost:

(MH "burns+" OR MH "burn units" OR TI (((burn OR burned) NOT burn*-out*) OR burns OR ((thermal OR heat) N2 injur*)) OR AB (((burn OR burned) NOT burn*-out*) OR burns OR ((thermal OR heat) N2 injur*))) AND ("Quality of Life+" OR MH "health status" OR MH "Health Status Indicators" OR MH "Disability Evaluation" OR MH "Job Re-Entry" OR MH "Activities of Daily Living+" OR MH "chronic pain" OR MH "Disabled+" OR MH "Motor Activity" OR MH "Physical Fitness" OR TI ((quality* N2 life*) OR hrql OR hrqol OR ((health OR functional* OR physic*) N2 (stat* OR limitation* OR outcome* OR recover* OR impair* OR result* OR fitness*)) OR disab* OR invalid* OR ((walk* OR work* OR mobilit*) N2 (difficult* OR limit*)) OR (work N2 (resum* OR return OR back)) OR ((sexual* OR disease* OR assess* OR hand OR adaptat*) N2 function*) OR (dail* N2 (life OR living) N2 activit*) OR adl OR adls OR iadl OR iadls OR badl OR badls OR ((chronic* OR long*-term* OR longterm*) N2 (pain*)) OR (loss N2 function*) OR (burn N2 outcome* N5 questionnaire*) OR boq OR ((burn OR burns) N5 health N2 scale*) OR bshs OR euroqol OR sf36 OR sf-36 OR short-form-36 OR eq-5d OR eq-6d OR eq5d OR eq6d OR itqol OR (physical* N2 (activ* OR perform*)) OR (independent* N2 living) OR (social* N2 participat*)) OR AB ((qualit* N2 life*) OR hrql OR hrqol OR ((health OR functional* OR physic*) N2 (stat* OR limitation* OR outcome* OR recover* OR impair* OR result* OR fitness*)) OR disab* OR invalid* OR ((walk* OR work* OR mobilit*) N2 (difficult* OR limit*)) OR (work N2 (resum* OR return OR back)) OR ((sexual* OR disease* OR assess* OR hand OR adaptat*) N2 function *) OR (dail* N2 (life OR living) N2 activit*) OR adl OR adls OR iadl OR iadls OR badl OR badls OR ((chronic* OR long*-term* OR longterm*) N2 (pain*)) OR (loss N2 function*) OR (burn N2 outcome* N5 questionnaire*) OR boq OR ((burn OR burns) N5 health N2 scale*) OR bshs OR euroqol OR sf36 OR sf-36 OR short-form-36 OR eq-5d OR eq-6d OR eq5d OR eq6d OR itqol OR (physical* N2 (activ* OR perform*)) OR (independent* N2 living) OR (social* N2 participat*))) AND (MH "validation studies" OR MH "Sensitivity and Specificity" OR MH "Nonexperimental Studies" OR MH "Epidemiological Research+" OR TI (cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR validat* OR validit* OR (observation* N2 stud*) OR sensitiv* OR specific* OR "follow* up*" OR followup* OR epidemiolog* OR "case control*" OR "cross-section*") OR AB (cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR validat* OR validit* OR (observation* N2 stud*) OR sensitiv* OR specific* OR "follow* up*" OR followup* OR epidemiolog* OR "case control*" OR "cross-section*")) NOT (MH animals+ NOT MH humans+) NOT PT (letter OR news OR comment OR editorial OR congresses OR abstracts) AND LA English

Cochrane

((((burn OR burned) NOT burn*-out*) OR burns OR ((thermal OR heat) NEAR/3 injur*)):ab,ti) AND (((qualit* NEAR/3 life*) OR hrql OR hrqol OR ((health OR functional* OR physic*) NEAR/3 (stat* OR limitation* OR outcome* OR recover* OR impair* OR result* OR fitness*)) OR disab* OR invalid* OR ((walk* OR work* OR mobilit*) NEAR/3 (difficult* OR limit*)) OR (work NEAR/3 (resum* OR return OR back)) OR ((sexual* OR disease* OR assess* OR hand OR adaptat*) NEAR/3 function*) OR (dail* NEAR/3 (life OR living) NEAR/3 activit*) OR adl OR adls OR iadl OR iadls OR badl OR badls OR ((chronic* OR long*-term* OR longterm*) NEAR/3 (pain*)) OR (loss NEAR/3 function*) OR (burn NEAR/3 outcome* NEAR/6 questionnaire*) OR boq OR ((burn OR burns) NEAR/6 health NEAR/3 scale*) OR bshs OR euroqol OR sf36 OR sf-36 OR short-form-36 OR eq-5d OR eq-6d OR eq5d OR eq6d OR itqol OR (physical* NEAR/3 (activ* OR perform*)) OR (independent* NEAR/3 living) OR (social* NEAR/3 participat*)):ab,ti) AND ((cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR validat* OR validit* OR (observation* NEAR/3 stud*) OR sensitiv* OR specific* OR 'follow* up*' OR followup* OR epidemiolog* OR 'case control*' OR 'cross-section*'):ab,ti)

Web of science

TS=(((((burn OR burned) NOT burn*-out*) OR burns OR ((thermal OR heat) NEAR/2 injur*))) AND (((qualit* NEAR/2 life*) OR hrql OR hrqol OR ((health OR functional* OR physic*) NEAR/2 (stat* OR limitation* OR outcome* OR recover* OR impair* OR result* OR fitness*)) OR disab* OR invalid* OR ((walk* OR work* OR mobilit*) NEAR/2 (difficult* OR limit*)) OR (work NEAR/2 (resum* OR return OR back)) OR ((sexual* OR disease* OR assess* OR hand OR adaptat*) NEAR/2 function*) OR (dail* NEAR/2 (life OR living) NEAR/2 activit*) OR adl OR adls OR iadl OR iadls OR badl OR badls OR ((chronic* OR long*-term* OR longterm*) NEAR/2 (pain*)) OR (loss NEAR/2 function*) OR (burn NEAR/2 outcome* NEAR/5 questionnaire*) OR boq OR ((burn OR burns) NEAR/5 health NEAR/2 scale*) OR bshs OR euroqol OR sf36 OR sf-36 OR short-form-36 OR eq-5d OR eq-6d OR eq5d OR eq6d OR itqol OR (physical* NEAR/2 (activ* OR perform*)) OR (independent* NEAR/2 living) OR (social* NEAR/2 participat*))) AND ((cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR validat* OR validit* OR (observation* NEAR/2 stud*) OR sensitiv* OR specific* OR "follow* up*" OR followup* OR epidemiolog* OR "case control*" OR "cross-section*"))) AND LA=(english) AND DT=(article)

Google scholar

burn | burns hrql | hrqol | "health | functional | physical status | limitation | outcome | recovery | impairment | fitness" | disability | adl | "chronic pain" | "loss * function" cohort | longitudinal | retrospective | prospective | validation | observational | "follow up" | epidemiologic

Appendix 2. Risk of bias assessed with the Quality in Prognostic Studies (QUIPS) risk of bias tool

First author year	1. Study Participation	2. Study Attrition	3. Prognostic Factor Measurement	4. Outcome Measurement	5. Study Confounding	6. Statistical Analysis & Reporting
Daltroy 2000	Moderate	High		Low		Low
Dodd 2010	Moderate	High	Low	Low	Moderate	Low
Kazis 2002[29]	Moderate	Moderate	Moderate	Low	High	Low
Kazis 2012	Low	High	Moderate	Low	Moderate	Moderate
Kazis 2016	Low	High		Low		Low
Laitakari 2015	Moderate	High		Low		Low
Landgraf 2013	Low	Moderate		Low		Low
Landolt 2002	Low	High	Low	Low	Moderate	Low
Landolt 2009	Low	Moderate	Moderate	Low	Moderate	Low
Maskell 2013	Moderate	High		Low		Low
Maskell 2014	Moderate	High		Low		Moderate
Meyer 2012	Moderate	High	Moderate	Low	High	Low
Nicolosi 2013	Moderate	High	Moderate	Low	High	Low
Palmieri 2012	Moderate	High		Low		Low
Pan 2015	Low	High		Low		Low
Patrick 2007	High	High		Moderate		Low
Pope 2007	High	High		Low		High
Rosenberg 2013	Moderate	High		Low		Low
Spuijbroek 2011	Low	Moderate		Low		Low
Stubbs 2011	Moderate	High	Low	Low	Moderate	Low
Sveen 2012	Low	Moderate	Moderate	Low	Moderate	Low
Sveen 2014	Low	Moderate	Moderate	Low	Moderate	Low
Van Baar 2006	Low	Moderate		Low		Low
Van Baar 2006	Low	Moderate		Low		Low
Van Baar 2011	Low	High	Moderate	Low	Moderate	Low
Warner 2012	Moderate	High	Moderate	Low	Moderate	Low
Weedon 2011	Low	High	Moderate	Low	Moderate	Moderate

Note. Domain 3 and domain 5 were only assessed for the thirteen prediction studies.

Appendix 3a. Parent proxy scores versus children's self-ratings for Burns Outcome Questionnaire 5-18

First author, year, assessment time point	Upper extremity function	Physical function and sports	Transfers and mobility	Pain[1]	Itch[1]	Appearance	Compliance	Satisfaction with current state	Emotional health	Family disruption[1]	Parental concern[1]	School re-entry
Pan 2015[42] – 6 months												
Proxy-score	99	96	98	10	22	77	91	86	92	9	21	54
Child score	99	97	99	9	20	86	88	85	92	5	12	55
Difference	0	-1	-1	1	2	-9*	3	1	0	4*	9*	-1
Pan 2015 – 18 months												
Proxy score	99	98	99	3	16	81	93	89	94	6	15	53
Child score	100	99	99	4	17	84	84	87	95	6	8	55
Difference	-1	-1	0	-1	-1	-3	9	2	-1	0	7	-2
Meyer 2012[24] – baseline												
Proxy score	44,6	43,4	44,7	40,7	43,1	49,3	52,7	44,7	48,5	43,4	45,4	48,4
Child score	46,3	44,1	46,3	42,0	41,8	50,5	53,1	45,5	49,4	44,2	44,4	49,2
Difference	-1,7*	-0,7	-1,6*	-1,3	1,3	-1,2	-0,4	-0,8	-0,9	-0,8	1,0	-0,8
Difference 3-6 months	-0,3	-0,4	-1,0*	0,1	0,1	-0,6	-0,1	0,3	-0,3	-1,5*	0,4	-0,4
Difference 6-12 months	-0,1	0,1	-0,2	0,0	0,5	-1,3	0,7	-0,4	-0,2	-0,4	0,9	-2,5*
Difference 12-24 months	0,0	-0,3	-0,3	-0,2	-0,1	-2,2*	1,0	-0,3	-0,7	-1,1	0,2	-0,4
Difference 24-36 months	0,3	0,1	0,7	-0,2	0,1	-1,5*	0,2	-0,5	-1,2	0,1	-0,3	0,8
Difference 36-48 months	-0,1	0,0	-0,2	0,1	0,4	-1,5	0,5	0,5	-1,3	0,8	0,7	-0,5
Van Baar 2006[48] – 21,1 months												
Difference	0	0	0	0	-2	-4	0	2	-2	-1	1	-3

[1]The scoring for these domains is reversed to the other domains: a lower score implies a better score for these domains.

*Significant difference ($p<0.05$) reported by the authors. Pan et al. did not report whether differences were statically significant different at 18 months and Van Baar et al. found no statically significant differences.

Appendix 3b. Parent proxy scores versus children's self-ratings for Pediatric Quality of Life Inventory

First author, year, assessment time point	Physical health	Psychosocial health	Emotional functioning	Social functioning	School functioning	Perceived physical appearance	Total score
Maskell 2013 - 7,3 year							
Proxy score	84.1	81.2	81.2	83.1	78.3		82.3
Child score	87.8	81.8	79.2	85.0	81.3		83.9
Difference	-3.7	-0.6	2.0	-1.9	-3.0		-1.6
Maskell 2014 – pre/post intervention							
Proxy score	75.5	67.7	64.7	70.5	66.4	49.1	70.2
Child score	86.2	73.5	75.4	78.3	66.8	50.0	78.0
Difference	-10.7	-5.8	-10.6	-7.8	-0.4	-0.9	-7.8
Proxy score	80.5	72.1	71.4	78.7	67.1	53.1	75.2
Child score	87.8	75.8	73.8	85.7	68.2	54.5	79.9
Difference	-7.3	-3.7	-2.5	-7.0	-1.1	-1.4	-4.8
Proxy score	80.8	70.2	66.2	75.3	65.0	63.1	73.1
Child score	85.4	77.3	77.7	82.2	72.0	60.2	80.1
Difference	-4.6	-7.2	-11.5	-6.9	-7.0	2.9	-7.1
Proxy score	86.5	74.7	69.3	81.8	69.5	69.3	78.0
Child score	90.2	80.0	79.0	85.2	75.5	69.5	83.6
Difference	-3.7	-5.3	-9.7	-3.4	-6.0	-0.2	-5.5

Note. None of the studies reported whether differences were statically significant different

Appendix 4. Summary of predictors of HRQL in pediatric burn patients

	Univariable								Multivariable						
Author, year (HRQL instrument)	Landolt 2002 (TACQOL)	Sveen 2012 (BOQ 5-18)	Landolt 2009 (TACQOL)	Dodd (BOQ 5-18)	Dodd (BOQ 0-4)	Kazis 2002 (BOQ 0-4)	Weedon 2011 (PedsQL)	Meyer 2012 (BOQ 11-18)	Sveen 2014 (BOQ 5-18)	Stubbs (BOQ 5-18)	Stubbs (BOQ 0-4)	Van Baar 2011 (BOQ 5-18)	Warner 2012 (BOQ 5-18)	Kazis 2012 (BOQ 0-18)	Nicolosi 2013 (BSHS-R)
Patient characteristics															
Young age at burn	++		0						0				+	+	
Increasing age			+						-			+/-			
Male gender	0		0						+			+	-	0	0
Socioeconomic status	0		0												0
Comorbidity						-			-			0			
Burn-specific factors															
%TBSA burned	0	-	0			-			-	-	-	-		-	
Total full-thickness injury		-				-	-		-						
Longer length of hospital stay		-	0			-	0		-						
Visible scars	0								-						0
Facial burns			0							-	-	-	-	-	
Hand burns				-	-							0		-	
Time since burn/injury		+	0			+		+/-	+	+	+	+		+/-	
Aetiology			0				0								

Appendix 5a. Modified GRADE approach to rate the quality of evidence on content validity

Study design	Quality of evidence	Lower if
At least 1 content validity study	High	Risk of bias
No content validity studies	Moderate	-1 Serious
	Low	-2 Very serious
	Very low	
		Inconsistency
		-1 Serious
		-2 Very serious
		Indirectness
		-1 Serious
		-2 Very serious

Appendix 5b. Modified GRADE approach to rate quality of evidence on other measurement properties

Quality of evidence	Lower if
High	Risk of bias
Moderate	-1 Serious (no studies of very good quality or only one study of adequate quality)
Low	-2 Very serious (only one study of doubtful quality or only studies of inadequate quality)
Very low	-3 Extremely serious (only studies of inadequate quality)
	Imprecision
	-1 total (n = 50-100)
	-2 total (n < 50)
	Inconsistency
	-1 Serious (< 75% of studies did not display the same results)
	Indirectness
	-1 Serious (at least one study not addressing construct or target population of the review)

SUMMARY

The number of people surviving burns increased during the past few decades due to substantial advances in critical and surgical care management. Therefore, more patients have to deal with lifelong psychological problems, disabilities and disfigurements which are frequently a consequence of burn injury. This caused a shift in attention from short-term clinical outcomes (e.g. improvement of survival) to long-term patient-reported outcomes (PROMs) focusing on psychological sequelae and physical appearance. The studies described in this thesis aim to critically appraise available evidence of often used measurement instruments and contribute to improved surgical burn care through the evaluation of surgical techniques.

PART I HEALTH-RELATED QUALITY OF LIFE AFTER BURN INJURY

Health-related quality of life (HRQL) reflects a patients' perception of how a health condition affects his/her physical, psychological and social wellbeing after an injury or disease. HRQL is, therefore, an important PROM to understand, qualify and quantify the impact of burn injuries. In order to reveal and summarize the current knowledge on HRQL after burns, an extensive systematic literature review was performed. Results of Chapter 2 and Chapter 4 show that HRQL is increasingly studied after burn injury and no consensus consists on what specific HRQL instrument(s) should be used in adults or children after burn injuries. Many different instruments and a broad range of time assessment points were used, which complicated comparison and correlation of outcomes. HRQL was diminished shortly after sustaining the injury in most adults and children, but improved over time. Nevertheless, some HRQL domains and patient groups remained impaired, indicating that they need special attention during follow-up. In adults, participation restrictions due to physical functional limitations and emotional distress remained, including limitations of regular daily activities like work. In addition, burn patients seemed to be on average more anxious and/or depressed compared to the general population and a large number of patients continued experiencing pain and/or discomfort in the long term (Chapter 2). In children, especially appearance and parental concern remained a problem (Chapter 4).

These results show that burns have a prolonged impact on patients' daily and social life and indicate the need for a comprehensive approach, including both physical and psychological care in the aftermath of burns. Also, some patient groups were found to require special attention. Patients with major burns experienced evidently more problems than patients with mild burns, and patients with psychological problems before or shortly after the injury are frequently seen to maintain problems in the long-term and should therefore be carefully monitored (Chapter 3). In children, the psychological impact is barely studied. Paediatric patients with major burns, facial

burns, hand burns, and comorbidity experienced a poorer long-term HRQL and required special attention.

PART II CLINIMETRIC STUDIES ON OUTCOMES

The discipline of clinimetrics aims to improve the quality of measurements by assessment of the measurement properties of existing measurement instruments or by development of new instruments. In Chapter 5 measurement properties of instruments that have been used to assess HRQL in burn patients were critically assessed with the aim to guide the selection of an appropriate measurement instrument. Results showed that the Burns Specific Health Scale-brief (BSHS-B) and the Brisbane Burn Scar Impact Profile (BBSIP), have somewhat more favourable properties compared to other instruments. For now, it is recommended to use the BSHS-B to evaluate disease-specific HRQL in clinical burn research in adults based on the quality of measurement properties. The generic instruments EQ-5D and SF-36 can be used best to make comparisons with population norms or other patient groups. In children, none of the burn-specific instruments could be recommended based on the quality of the measurement properties or feasibility aspects and no evidence on measurement properties of generic instruments after burn injury exists. To further improve the understanding of HRQL after burn injury, consensus on uniformly validated instrument(s), time assessment points and data presentation is needed.

In Chapter 6, the minimal important change (MIC) and minimal clinically important difference (MCID), two important interpretability aspects, of the patient part of the Patient and Observer Scar Assessment Scale version 2.0 (POSAS 2.0) were explored. Our results suggest that patients consider minor differences (less than 0.75 on the 1–10 scale) in POSAS scores as clinically important scar quality changes. MCID values ranged from -0.39 to -0.08 and can be used to evaluate the effects of burn treatment at the same time (i.e. the difference between two trial arms) and to perform sample-size calculations.

PART III OUTCOMES OF BURN SURGERY

Part III focuses on clinical outcomes of surgical debridement and split skin grafting, which is the standard of care for dermal burn wounds that are not expected to heal in 14-21 days. Over the last decade, hydrosurgery has become available in burn surgery as an alternative technique for the gold standard of conventional tangential debridement by guarded knives. The retrospective analyses in Chapter 7 show that the use of the VERSAJET™ hydrosurgery system (Smith+Nephew, London UK) in specialized Dutch burn care is substantial and that it is more often used in younger patients, scalds, burns on irregularly contoured body areas and in patients

with a larger % total body surface area (TBSA) burned. To assess if hydrosurgical debridement leads to the long-standing assumption of maximal dermal preservation with a better scar outcome, we conducted the HyCon study (long-term scar quality after Hydrosurgical versus Conventional debridement for deep dermal burns). This double-blind within-patient randomized controlled trial (RCT) was conducted in the three Dutch Burn Centres. The study protocol of this trial is described in Chapter 8. Results of the trial are presented in Chapter 9 and show that scar quality and pliability outcomes were better for hydrosurgical debrided burns one year post surgery. This was probably the result of significant enhanced preservation of dermis during debridement, which was demonstrated with histology measures. Although clinical relevance of the outcomes might be limited as differences were small, hydrosurgery should be considered as a superior alternative for knife debridement. Further research is necessary to gain insight in the benefits of hydrosurgery for different patient- and burn categories.

Surgical debridement is followed by autologous split skin grafting, leaving a donor-site wound. Donor-site scarring is an important harm that patients should be informed about if split skin grafting is considered. During the DonorSite (DOSIS) study, we followed burn patients up to one year post-surgery to assess (patient-reported) scar quality of donor-sites after skin graft harvesting. In Chapter 10, changes in characteristics of donor-site scar quality over time and the opinion of patients and caregivers on donor-site scars were examined. Patients' and clinicians' perceptions of scar quality only slightly improved during scar maturation. The agreement between patients and observers generally increased between 3 and 12 months, but remained 'poor' for all items of the POSAS. Results of this study indicate that caregivers seem to underestimate the impact of scars on patients. Chapter 11 provides further insights in long-term patient reported scar quality of donor-sites. Donor-site scars differed from normal skin in a large part of the population 12 months post-surgery. Even one year after surgery the mean overall opinion of patients on donor-site scars was remarkably high (POSAS score 3.2 (scale 1-10)). Moreover, 37% of the patients reported a poor overall opinion on the donor-site scar (i.e. POSAS score ≥ 4). Especially colour of the donor site-scars was judged to remain deviant from normal skin. A younger age, female gender, a darker skin type, location on the lower leg and prolonged time to re-epithelization predict reduced patient-reported donor-site scar quality. Results of Chapter 10 and Chapter 11 can be used to inform patients on the long-term outcomes of their scars and to tailor preventive or therapeutic treatment options.

NEDERLANDSE SAMENVATTING

Ontwikkelingen in de acute brandwondenzorg hebben de overlevingskans van patiënten met ernstige brandwonden de afgelopen jaren aanzienlijk verbeterd. Als gevolg hiervan moeten ook meer patiënten leven met de langetermijngevolgen van dit letsel. Dit heeft geleid tot een verandering in veel gebruikte uitkomsten in de brandwondenzorg. Waar eerder de focus lag op kortetermijnuitkomsten, zoals sterfte en IC opname, is er nu meer aandacht voor patiënt gerapporteerde langetermijnuitkomsten (PROMS, patient-reported outcome measures), zoals kwaliteit van leven en littekenkwaliteit. De in dit proefschrift beschreven studies hebben het doel al het beschikbare wetenschappelijke bewijs over de meest gebruikte meetinstrumenten samen te vatten en kritisch te beoordelen (Deel I en II) en bij te dragen aan de verbetering van chirurgische zorg voor brandwonden door de vergelijking van verschillende operatietechnieken (Deel III).

DEEL I KWALITEIT VAN LEVEN NA EEN BRANDWOND

Gezondheidsgerelateerde kwaliteit van leven (GKvL) geeft informatie over hoe een persoon zijn of haar lichamelijk, psychisch en sociaal functioneren ervaart en is een belangrijke patiënt gerapporteerde uitkomstmaat om de impact van een ongeluk of ziekte te kunnen beoordelen en kwantificeren. Door middel van een systematisch literatuuronderzoek werd al het beschikbare onderzoek naar GKvL na brandwonden bestudeerd. Al deze kennis is samengevat in het eerste deel van dit proefschrift. De resultaten van hoofdstuk 2 en hoofdstuk 4 laten zien dat GKvL steeds vaker bestudeerd wordt, maar dat er voor zowel volwassenen als kinderen geen consensus bestaat over welk meetinstrument (d.w.z. vragenlijst) het beste gebruikt kan worden om GKvL te meten. Het samenvatten en vergelijken van de resultaten van alle verschillende studies was complex, omdat er verschillende vragenlijsten werden gebruikt die ook nog eens op verschillende tijdstippen na het brandwondenongeluk waren afgenomen. Over het algemeen leek GKvL kort na het oplopen van de brandwonden op veel fysieke en psychosociale domeinen verminderd, maar dit herstelde zich grotendeels na verloop van tijd. Patiënten met brandwonden zijn echter wel vaker angstig en/of depressief in vergelijking met de algemene bevolking en een groot deel ervaart pijn en/of discomfort op de lange termijn. Volwassenen ervaren op de lange termijn met name beperkingen bij dagelijkse activiteiten, zoals werk. Kinderen ervaren op lange termijn vooral problemen met hun uiterlijk. Daarnaast blijkt dat de bezorgdheid van ouders op lange termijn groot blijft. Deze resultaten geven aan dat brandwonden een langdurige impact hebben op het dagelijkse functioneren en sociale leven van patiënten. Daarom moet er tijdens de nazorg aandacht zijn voor zowel fysieke als psychosociale zorg en is het belangrijk om ook aandacht te hebben voor de familie van de patiënt.

Om te exploreren welke patiënten mogelijk extra (na)zorg nodig hebben werden voorspellende factoren van verminderde GKvL beschreven in hoofdstuk 3 (volwassenen) en hoofdstuk 4 (kinderen). Niet geheel onverwacht bleken ernst van de brandwonden en het hebben van post-traumatische psychische problemen sterke voorspellers voor een verminderde GKvL op lange termijn. Bij kinderen bleek daarnaast dat het hebben van brandwonden op zichtbare lichaamsdelen, zoals het gezicht en de handen, een verminderde GKvL voorspelt. Hoe jonger kinderen de brandwonden oplopen, hoe minder last zij hiervan op lange termijn lijken te hebben. Dit komt mogelijk doordat jonge kinderen zich minder herinneren van het trauma en zich makkelijker aanpassen aan littekens en een veranderd uiterlijk.

Om onderzoeksresultaten in de toekomst beter samen te kunnen voegen en vergelijken moet er een internationale consensus komen over welk instrument gebruikt moet worden, op welke momenten GKvL gemeten dient te worden en hoe uitkomsten gepresenteerd moeten worden.

DEEL II KLINIMETRISCHE STUDIES NAAR UITKOMSTEN

Dit deel van het proefschrift richt zich op het onderzoeken van de eigenschappen van verschillende meetinstrumenten voor patiënt gerapporteerde uitkomsten.

Vragenlijsten die gebruikt worden om GKvL te meten zijn ziekte-specifiek of generiek. Ziekte-specifieke vragenlijsten zijn speciaal ontwikkeld voor mensen met brandwonden en bevatten bijvoorbeeld vragen over uiterlijk en (over)gevoeligheid voor hitte. Generieke vragenlijsten bevatten meer algemene vragen en kunnen voor elke ziekte of populatie gebruikt worden. In hoofdstuk 5 worden de meeteigenschappen van ziekte-specifieke en generieke vragenlijsten die gebruikt zijn in onderzoeken naar GKvL bij patiënten met brandwonden beschreven. Het doel hiervan was te bepalen welke vragenlijst de meest betrouwbare en valide resultaten geeft en dus het beste kan worden gebruikt. De ziekte specifieke 'Burns Specific Health Scale-Brief' (BSHS-B) en 'Brisbane Burn Scar Impact Profile' (BBSIP) bleken de beste meeteigenschappen te hebben. Omdat de BSHS-B in de meeste talen beschikbaar is, raden wij op basis van ons onderzoek aan deze te gebruiken. De generieke vragenlijsten EuroQol-5 Dimension (EQ-5D) en 36-item Short Form Survey (SF-36) kunnen het best gebruikt worden wanneer men GKvL van brandwonden wil vergelijken met de algemene populatie of andere patiëntengroepen. We konden op basis van kwaliteit van meeteigenschappen geen aanbeveling doen voor het gebruik van specifieke vragenlijsten voor kinderen.

In hoofdstuk 6 zijn de Minimal Important Change (MIC) en Minimal Cinically Important Difference (MCID) van de Patient and Observer Scar Assessment Scale versie 2.0 (POSAS 2.0) onderzocht. Met de POSAS kan men patiënt gerapporteerde littekenkwaliteit meten doordat de patiënt een

score invult voor pijn, jeuk, kleur, dikte, reliëf en soepelheid van het litteken ten opzichte van de normale huid. De MIC en MCID geven aan welk verschil in POSAS uitkomst voor patiënten belangrijk is en zijn daarom belangrijk om de uitkomsten van onderzoeken naar littekenkwaliteit te interpreteren. Uit de resultaten blijkt dat patiënten kleine veranderingen (<0.75 op een schaal van 1 – 10) en kleine verschillen (<0.39 en op een schaal van 1 – 10) al als belangrijk ervaren.

DEEL III UITKOMSTEN VAN BRANDWONDENCHIRURGIE

In dit deel ligt de focus op klinische uitkomsten van huidtransplantaties, de standaard behandeling voor brandwonden die niet binnen 14-21 dagen genezen. Voordat een huidtransplantatie kan plaatsvinden moet eerst het aangedane weefsel van de brandwond worden gedebrideerd. Conventioneel wordt dit gedaan met een scherp chirurgisch mes, waarmee het dode weefsel laagje voor laagje wordt weggeneden (tangentiële excisie). De laatste jaren is hydrochirurgie populair geworden als alternatief. Hierbij wordt met behulp van water onder hoge druk het dode weefsel verwijderd. De retrospectieve analyse in hoofdstuk 7 laat zien dat het VERSAJET™ hydrosurgery system (Smith+Nephew, London UK) in de Nederlandse brandwondencentra veel wordt gebruikt en dat het vaker wordt ingezet bij jongere patiënten, heet water verbandingen, brandwonden op irreguliere oppervlakten en bij patiënten met een groter percentage totaal verbrand lichaamsoppervlak (TVLO).

Men denkt dat hydrochirurgisch débridement voor een huidtransplantatie lijdt tot een betere littekenkwaliteit doordat er meer gezonde huid wordt gespaard dan bij tangentiële excisie met een mes. Om dit te onderzoeken werd de HyCon studie (lange termijn uitkomsten van littekenkwaliteit na Hydrochirurgisch versus Conventioneel débridement van brandwonden) opgezet. Dit dubbelblind gerandomiseerde onderzoek werd uitgevoerd bij 137 patiënten in de drie Nederlandse brandwondencentra. Het primaire eindpunt was littekenkwaliteit, gemeten met de POSAS. In hoofdstuk 8 is het studieprotocol samengevat. De resultaten van de studie worden in hoofdstuk 9 beschreven. Deze laten zien dat een jaar na de operatie kwaliteit en soepelheid van het litteken beter zijn na hydrochirurgisch débridement van de brandwond dan na tangentiële excisie. Er werd geen verschil in nadelige effecten, zoals infectie of wondgenezingsstoornissen, van de ingreep gezien. Uit histologisch onderzoek bleek ook dat er, microscopisch gezien, inderdaad meer dermis (lederhuid) gespaard werd met hydrochirurgie. Ondanks dat verschillen tussen de littekenmetingen klein waren, kan hydrochirurgie daarom als een beter alternatief voor débridement met het mes gezien worden.

Wanneer er een huidtransplantatie verricht wordt, ontstaat er ook altijd een wond op de plaats van het lichaam waar de gezonde huid vandaan wordt gehaald. Hier vormt zich ook een litteken. Voordat patiënten de operatie ondergaan moeten zij hier dan ook over worden ingelicht. Om

meer inzicht te krijgen in patiënt gerapporteerde kwaliteit van de littekens van deze donorplaatsen volgden wij patiënten tot één jaar na de huidtransplantatie tijdens de DonorSite (DOSIS) studie. In hoofdstuk 10 wordt beschreven hoe patiënten en artsen verschillende kenmerken (o.a. kleur, dikte, reliëf en soepelheid) van littekens gedurende het jaar beoordeelden. Twaalf maanden na de operatie werden de meeste kenmerken niet veel beter gescoord dan 3 maanden na de operatie, terwijl littekenkwaliteit over het algemeen vaak over de tijd verbetert als gevolg van rijping van het litteken. Daarnaast bleek dat de mening van artsen en patiënten over de kenmerken slecht overeenkwam en artsen de impact van het litteken voor de patiënt onderschatten. In hoofdstuk 11 gaan we verder in op de patiënt gerapporteerde kwaliteit van donorplaats littekens. Een jaar na het ongeluk werden alle POSAS littekenkenmerken door patiënten nog gescoord als 'afwijkend van de normale huid' en was de gemiddelde score voor 'algehele indruk' over het litteken opmerkelijk slecht; van alle patiënten rapporteerde 37% een 'slechte' algehele indruk (POSAS score ≥ 4). Patiënten gave met name een slechte score voor de littekeneigenschap 'kleur'. Voorspellende factoren van verminderde patiënt gerapporteerde littekenkwaliteit waren een lagere leeftijd, vrouwelijk geslacht, donker huidtype, locatie op het onderbeen en langere genezingsduur van de initiële wond. De resultaten van hoofdstuk 10 en hoofdstuk 11 kunnen worden gebruikt om patiënten voor te lichten over de langetermijnuitkomsten van littekens op de donorplaats en om de behandeling daarop aan te passen.

PHD PORTFOLIO

PhD student: C.M. Legemate
PhD period: August 2016 – November 2021
Supervisors: Prof. dr. E. Middelkoop
dr. C.H. van der Vies
dr. M.E. van Baar

Courses	Year	EC
Research integrity	2021	2.0
Longitudinale data-analyse	2018	3.0
Methodologische advisering	2018	3.0
Multilevel analyse	2018	2.0
Klinische predictiemodellen	2018	2.0
Missing data: Consequences and solutions	2018	2.0
Basiscursus Regelgeving en Organisatie voor Klinisch Onderzoekers (BROK)	2018	1.0
Praktische epidemiologie: Opzetten van een onderzoek	2017	4.0
Klinimetrie: Het ontwikkelen en evalueren van meetinstrumenten	2017	3.0
Regressietechnieken	2016	5.0
Emergency Management of Severe Burns	2016	2.7
Principes van epidemiologische data-analyse	2016	3.0
Writing a scientific article	2016	3.0
Epidemiologisch onderzoek: Opzet en interpretatie	2016	4.0

(Inter)national conferences attended		
Amsterdam Movement Sciences, PhD day	2020	1.0
Najaarsvergagering, NVPC, Leeuwarden	2019	1.0
18th European Burn Association congress, Helsinki, Finland	2019	2.0
Voorjaarsvergadering, NVPC, Amsterdam	2019	1.0
2nd European Burn Association educational course, Rotterdam	2018	2.0
19th International Society for Burn Injury congress, New Delhi, India	2018	2.0
Scar Academy and the European Tissue Repair Society congress, Amsterdam	2018	2.0
17th European Burn Association congress, Barcelona, Spain	2017	2.0
Najaarsvergagering, NVPC, Hengelo	2017	1.0
Amsterdam Movement Sciences, PhD day	2017	1.0
Voorjaarsvergadering, NVPC, Amsterdam	2017	2.0

Supervising	
Scientific internship students	2.0
F. Klein, P.J. Ooms	

Oral presentations

- Patiënt gerapporteerde uitkomsten van littekenkwaliteit van donorsites na huidtransplantaties: uitkomsten van een prospectieve cohort studie.
  Wetenschapsdag Maasstad Ziekenhuis, 2021, Rotterdam
- Hydrochirurgisch debridement versus Conventioneel debridement voor split skin grafting: resultaten van een intra-patient gerandomiseerde multi-centre trial.
  Wetenschapsdag Bandwondenstichting, 2021, online
- Patient satisfaction regarding donor site scars after skin graft harvesting: a prospective cohort study.
  18th European Burn Association congress, 2019, Helsinki, Finland
- Recommendations on the most suitable health related quality-of-life measurement instrument after burn injury: a systematic review of measurement properties.
  18th European Burn Association congress, 2019, Helsinki, Finland
- Predictors of patient satisfaction regarding donor site scars after skin graft harvesting: a prospective cohort study.
  19th International Society for Burn Injury congress, 2018, New Delhi, India
- Scar quality of donor sites following skin graft harvesting in burn patients: an explorative cohort study.
  Scar Academy and the European Tissue Repair Society congress, 2018, Amsterdam
- Een multicenter gerandomiseerde studie naar lange termijn littekenkwaliteit na hydrochirurgische versus conventionele tangentiële excisie van diep dermale brandwonden.
  Wetenschapsdag Bandwondenstichting, 2017, Amersfoort
- De rol van hydrochirurgie in de Nederlandse brandwondenzorg. De HyCon study: een multicentre RCT.
  Wetenschapsdag Maasstad Ziekenhuis, 2016, Rotterdam

Poster presentations

- The application of hydrosurgery for burn wound debridement: an 8 year cohort analysis
  2nd European Burn Association educational course, 2018, Rotterdam
- Evaluation of hydrosurgery in Dutch burn care: 'Which patients benefit most?'
  17th European Burn Association congress, 2017, Barcelona, Spain

LIST OF PUBLICATIONS

This thesis

Legemate CM, Middelkoop E, Carrière ME, van Zuijlen PPM, van Baar ME, van der Vlies CH, the Hycon Study group. Clinically relevant changes and differences in scar outcome for burn patients measured by the Patient and Observer Scar Assessment Scale. *Submitted.*

Legemate CM, Kwa KAA, Goei H, Pijpe A, Middelkoop E, van Zuijlen PPM, Beerthuizen GIJM, Nieuwenhuis MK, van Baar ME, van der Vlies CH, the HyCon Study Group. Scar quality after hydrosurgical versus conventional debridement for burns: a within-patient randomised double-blind trial. *Accepted for publication in the British Journal of Surgery.*

Legemate CM, Ooms PJ, Trommel N, Middelkoop E, van Baar ME, Goei H, et al. Patient-reported scar quality of donor-sites following split-skin grafting in burn patients: Long-term results of a prospective cohort study. *Burns. 2021;47(2):315-21.*

Legemate CM, Ooms PJ, Trommel N, Goei H, Lucas Y, Middelkoop E, et al. Course of scar quality of donor sites following split skin graft harvesting: Comparison between patients and observers. *Wound Repair Regen. 2020;28(5):696-703.*

Legemate CM, Spronk I, Mokkink LB, Middelkoop E, Polinder S, van Baar ME, et al. Evaluation of measurement properties of health-related quality of life instruments for burns: A systematic review. *J Trauma Acute Care Surg. 2020;88(4):555-71.*

Legemate CM, Goei H, Gostelie OFE, Nijhuis THJ, van Baar ME, van der Vlies CH, et al. Application of hydrosurgery for burn wound debridement: An 8-year cohort analysis. *Burns. 2019;45(1):88-96.*

Spronk I, **Legemate CM**, Polinder S, van Baar ME. Health-related quality of life in children after burn injuries: A systematic review. *J Trauma Acute Care Surg. 2018;85(6):1110-8.*

Spronk I, **Legemate CM**, Dokter J, van Loey NEE, van Baar ME, Polinder S. Predictors of health-related quality of life after burn injuries: a systematic review. *Crit Care. 2018;22(1):160.*

Spronk I, **Legemate C**, Oen I, van Loey N, Polinder S, van Baar M. Health related quality of life in adults after burn injuries: A systematic review. *PLoS One. 2018;13(5):e0197507.*

Legemate CM, Goei H, Middelkoop E, Oen I, Nijhuis THJ, Kwa KAA, et al. Long-term scar quality after hydrosurgical versus conventional debridement of deep dermal burns (HyCon trial): study protocol for a randomized controlled trial. *Trials. 2018;19(1):239.*

Other publications

Fledderus AC, Boom T, **Legemate CM**, van der Horst CMAM, Spuls PI. Measurement instruments for the core outcome set of congenital melanocytic naevi and an assessment of the measurement properties according to COSMIN: a systematic review. *Submitted.*

Legemate CM, van der Vlies CH. Donor site morbidity - Where is the evidence to guide therapeutic decision making? *Burns. 2021;47(6):1456-7.*

Legemate CM, de Rooij FPW, Bouman MB, Pigot GL, van der Sluis WB. Surgical outcomes of testicular prostheses implantation in transgender men with a history of prosthesis extrusion or infection. *Int J Transgend Health. 2021;22(3):330-6.*

Kwa KAA, **Legemate CM**, Pijpe A, Meij-de Vries A, Middelkoop E, van Baar ME, et al. Doxepin cream is not effective in reducing itch in burn scar patients: A multicenter triple-blind randomized clinical crossover trial. *Burns. 2020;46(2):340-6.*

Legemate CM, Lucas Y, Oen I, van der Vlies CH. Regrafting of the Split-Thickness Skin Graft Donor-Site: Is It Beneficial? *J Burn Care Res. 2020;41(1):211-4.*

Legemate CM, Middelkoop E, van Baar ME, van der Vlies CH. Is the time right to put down the knife? A call for evidence-based decision making. *Burns. 2018;44(7):1859-60.*

Legemate CM, Hop MJ, Nieuwenhuis MK, Middelkoop E, van Baar ME, van der Vlies CH. Determining depth of burns using laser Doppler imaging. *Ned Tijdschr Geneeskd. 2018;162:D2374.*

Legemate CM, van der Kwaak M, Gobets D, Huikeshoven M, van Zuijlen PPM. The pedicled internal pudendal artery perforator (PIPAP) flap for ischial pressure sore reconstruction: Technique and long-term outcome of a cohort study. *J Plast Reconstr Aesthet Surg. 2018;71(6):889-94.*

DANKWOORD

In de eerste plaats gaat mijn dank uit naar alle patiënten en hun families die mee hebben willen doen aan de verschillende onderzoeken. Hun geduld tijdens polibezoeken en inspanningen om bij te dragen aan het onderzoek naar brandwonden waren indrukwekkend. Ook ben ik alle medewerkers van het brandwondencentrum in het Maasstad Ziekenhuis zeer dankbaar voor hun inzet, geduld en hulp. Mede door hen heb ik met groot plezier aan dit proefschrift gewerkt.

Dr. C.H. van der Vlies, beste Kees, je legde de basis voor mijn promotietraject en gaf me de vrijheid om dit deels zelf in te vullen. Dat je me ondersteunde bij het volgen van de master epidemiologie en de ruimte gaf om een periode in de kliniek te werken is zeer waardevol voor mij geweest. Heel veel dank voor het vertrouwen, je persoonlijke betrokkenheid en het creëren van zo'n fijne werkplek.

Dr. M.E. van Baar, beste Margriet, je algehele kennis over (brandwonden)onderzoek en vlotte reacties zorgden ervoor dat ik direct weer verder kon als ik het zelf even niet meer overzag. De korte lijntjes en je toegankelijkheid waren voor mij als beginnend onderzoeker heel prettig. Bedankt voor je betrokkenheid, open kritische houding en de vele (ad-hoc) overlegmomenten.

Prof. dr. E. Middelkoop, beste Esther, onze voortgangsgesprekken maar ook je (kritische) commentaar op mijn stukken gaven vaak nieuwe inzichten om het onderzoek naar een hoger niveau te tillen. Dank voor je betrokkenheid op afstand, het veelvuldig geven van het zetje naar 'definitieve' beslissingen en je steun bij het vaart zetten achter de laatste loodjes.

Nicole Trommel, lieve Nicole, als onderzoeksverpleegkundige op het BWC ben jij van onschatbare waarde geweest voor mijn projecten. Van jou heb ik geleerd dat je met doorzettingsvermogen, hard werken en creatief denken veel kan bereiken in de onderzoekswereld. Het kan niet, bestaat niet. Wanneer ik weg was voor mijn studie of werk in de kliniek kon ik mijn onderzoeken met een gerust hart bij jou achterlaten. Bedankt voor al je hulp en gezelligheid op de afdeling en tijdens congressen. Ik wens je heel veel succes bij de vervolgstappen in je loopbaan op het BWC.

Geachte leden van de promotiecommissie: prof. dr. M.B. Bouman, prof. dr. C.C. Breugem, prof. dr. M.H.J. Verhofstad, prof. dr. A.E.A.M. Weel-Koenders en dr. E. Bosma, hartelijk dank voor het kritisch lezen van mijn manuscript en uw goedkeuring om dit werk te mogen verdedigen.

Onderzoek doe je niet alleen. Met name aan de klinische onderzoeken hebben veel mensen, door heel Nederland, meegewerkt. Ik wil bij deze alle leden van de HyCon studiegroep en medeauteurs van verschillende manuscripten bedanken. In het bijzonder:

Jan Dokter en Irma Oen, als brandwondenartsen hebben jullie veel betekend voor het uitrollen van verschillende onderzoeken en het verzamelen van data. Bedankt voor de extra inspanningen die jullie hiervoor moesten leveren en jullie persoonlijke betrokkenheid.

Helma Hofland, je ervaring als onderzoeks- en brandwondenverpleegkundige in binnen- en buitenland is indrukwekkend. Bedankt voor de ondersteuning bij de HyCon en DOSIS studie en je inspirerende verhalen.

Ymke Lucas, lieve Ym, wat hebben we een leuke tijd gehad op het BWC, maar zeker ook daarbuiten tijdens sportklasjes in Amsterdam, lunchen in Rotterdam en natuurlijk onze reis door India na de ISBI in New Delhi. Bedankt voor het zetten van vele cappuccino's en het verdelen van studiewonden op de OK. Ik ben blij dat jij vandaag als collega, vriendin en paranimf naast me wil staan.

Er zijn veel mensen die niet zozeer direct hebben bijgedragen aan de wetenschappelijke inhoud van dit proefschrift maar die wel van belang zijn geweest op weg naar deze promotie en het behoud van een goede work-life balance. Hiertoe horen mijn familie, vrienden, vriendinnen en (oud)collega's. In het bijzonder:

Stafleden en ANIOS van de afdeling plastische chirurgie van het OLVG, het was inspirerend om een tijdje in zo'n jong (voornamelijk vrouwelijk) team te werken. Dank voor de fijne werksfeer en alles wat ik van jullie geleerd heb.

Stafleden en AIOS van de afdeling plastische chirurgie van het Amsterdam UMC, bedankt voor de leuke en leerzame tijd. Ik kijk ernaar uit om terug te komen voor het vervolg van mijn opleiding. Dr. M.C. Obdeijn en dr. H.A.H. Winters, hartelijk dank voor het in mij gestelde vertrouwen bij de aanname voor de opleiding tot plastisch chirurg.

Loes Huygen, lieve Loes, van zweten in de tentamenzalen op de VU en borrels in Amsterdam naar worstenbroodjes in Den Bosch. Ik vind het bijzonder knap dat jij inmiddels al radioloog bent met een mooi gezin terwijl ik 'net begin'. Dank voor alle gezelligheid en de goede gesprekken als het even tegenzat.

Brigitte, Charlie, Fleur, Joelle, Lara, Maartje, Maud, Vere en Rosanne, mijn vriendinnen van de BC (de afkorting lijkt me voldoende ;)). Mijn verhuizing naar de Nassaukade viel ongeveer gelijk met de start van mijn promotietraject. Ik denk met veel plezier terug aan onze borrels, diners en feestjes, die zorgden voor de nodige afleiding naast werk. Bedankt voor de bijzondere band die we hebben opgebouwd. Jullie zijn de besten om het leven mee te vieren maar ook te delen als het wat moeilijker is.

Lara Jongbloets en Fleur ter Kuile, dank voor jullie onvoorwaardelijke vriendschap.

Maartje de Nie, het met jou samen buigen over sollicitatiebrieven, oefenen van gesprekken en denken over carrièremogelijkheden heeft zeker bijgedragen aan waar ik nu sta. Onze voornamelijk overeenkomende ideeën over het leven zorgen voor een dierbare vriendschap. Bedankt voor al je nuchtere adviezen en dat je vandaag naast me wilt staan als paranimf.

Mijn broers en zus, Jaap, Sophie en Rik. Ons warme (steeds groter wordende) gezin zorgt voor een stabiele basis waar jullie elk op een eigen manier aan bijdragen. Ik ben trots op jullie!

Lieve papa en mama, jullie hebben me gestimuleerd om door te zetten en gesteund als ik het nodig had. De echte basis, dat zijn jullie. Een rustige, fijne thuishaven waar ons hele gezin altijd welkom is. Dank voor de onvoorwaardelijke steun en liefde. Ik ben blij dat jullie samen aan het roer staan.

Mijn Jip, bedankt voor alles.

ABOUT THE AUTHOR

Nine Legemate was born on November 5th, 1989 in Utrecht. She grew up in the nearby town Vleuten together with her parents, two brothers and sister. She finished secondary education at the Niftarlake college in Maarssen in 2007. Subsequently she started her medical training at the VU University in Amsterdam.

Halfway through medical school, Nine developed an interest in reconstructive surgery. In her final year she was especially intrigued by the field of burn surgery during a clinical elective at the plastic surgery department of the Red Cross Hospital in Beverwijk. After graduating medical school she started as a surgical resident not in training at the St. Antonius Hospital in Nieuwegein. In 2016, she got the opportunity to work fulltime as a PhD candidate on multiple research activities at the burn centre of the Maasstad Hospital in Rotterdam (Associations of Dutch Burn Centres), which led to the formation of this thesis. During this period she finished the post-graduate master Epidemiology and gained more clinical experience as a resident not in training at the plastic surgery departments of the ISALA hospital (Zwolle), OLVG (Amsterdam) and Amsterdam UMC.

In May 2021, Nine started her surgical training at the Jeroen Bosch Hospital in Den Bosch, after which she will continue her residency plastic, reconstructive and hand surgery at the Amsterdam UMC. With the completion of her PhD fellowship she will be able to register as a senior scientific epidemiological researcher (Epidemiologist B).

She is married to Jip Tolenaar and they are expecting their first child during this spring.

Amsterdam Movement Sciences

VU VRIJE UNIVERSITEIT AMSTERDAM

Amsterdam UMC University Medical Centers

Amsterdam Movement Sciences conducts scientific research to optimize physical performance in health and disease based on a fundamental understanding of human movement in order to contribute to the fulfillment of a meaningful life.