**Subjective scar assessment scales in orthopaedic surgery and determinants of patient satisfaction: a systematic review**

DECLARATIONS:

Declaration of competing interest: The authors declare that there are no conflicts of interest

Funding: this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ABBREVIATIONS[[1]](#footnote-1)

**Abstract**

*Introduction:* Scar assessment tools can be utilized during the post-operative period to monitor scar progress. The primary aim of this systematic review was to evaluate current subjective scar assessment scales utilized in orthopaedic surgery. The secondary aim of this review was to identify determinants of patients’ satisfaction with their scars and to evaluate current measurement scales.

*Materials and methods:* A systematic PRISMA compliant database search was conducted. Electronic databases, currently registered studies, conference proceedings and the reference lists of included studies were searched. There were no constraints based on language or publication status. A narrative synthesis provided a description and evaluation of scales utilized in orthopaedic surgery. Determinants of patient satisfaction were identified along with the scales used to measure satisfaction.

*Results*: A total of 7066 records were screened in the initial search. Twenty-six articles satisfied the inclusion criteria, assessing 7130 patients. Six validated subjective scar scales were identified in the literature. These were the Vancouver Scar Scale, Patient and Observer Scar Assessment Scale, Manchester Scar Scale, Stony Brook Scar Evaluation Scale, Visual Analogue Scale and the Hollander Wound Evaluation Scale. Studies utilizing these scales to evaluate scars following orthopaedic procedures did so successfully. These were total hip arthroplasty, total knee arthroplasty and limb reconstruction. The scales demonstrated satisfactory validity. Functional outcomes such as restoration of movement ranked among patients’ highest concerns. Scar cosmesis was found to be amongst patients’ lowest priorities.

*Conclusions:* Subjective scar assessment scales identified in the literature were not designed specifically for orthopaedic surgery. However, these were able to appropriately assess scars in the studies identified in this review. Current evidence suggests the effect of scar cosmesis on patient satisfaction with orthopaedic procedures is limited.

Key words: Scar; cosmesis; orthopaedics; patient satisfaction

**1.1 Introduction**

A scar can be defined as “a macroscopic disturbance of the normal structure and function of the skin architecture, resulting from the end product of a healed wound” [Ferguson et al1]. Scars can be aesthetically unpleasant for patients. They can cause anxiety, depression, and disruption of daily activities [Bayat et al2]. Objective scar assessment tools can be used to assess scar features such as colour, thickness, relief, pliability, and surface area [Verhaegen et al3]. Subjective tools provide a qualitative measurement by a patient or clinician [Fearmonti et al4]. These scales can be used to assess scars at different points during the post-operative period to monitor their progress.

Previous reviews of objective and subjective scar assessment tools have not been specific to orthopaedic surgery [Fearmonti et al4; Verhaegen et al3]. Scars resulting from orthopaedic procedures differ to those in other specialties. Orthopaedic surgery leads to notable tissue trauma, with severe scarring as a result [Grabowski et al5]. Scarring on the surface of joints can limit range of motion. Restriction of movement is less applicable in other specialties concerning tissues with low mobility (e.g., abdominal surgery). Additionally, mature scar strength is 80% that of intact skin, making it susceptible to dehiscence if loaded prematurely [Morton et al6]. Scar visibility, which is related to psychological distress [Brown et al7], varies according to position. Proximal scars covered by clothing are not visible and may lead to more distress than visible ones, since visible scars may allow patients to develop coping strategies [Brown et al7].

The primary aim of this systematic review was to evaluate current subjective scar assessment tools utilized in orthopaedic surgery. Objective scar assessment tools were excluded since these do not evaluate scar cosmesis [Singer et al8].

Patient satisfaction with scar cosmesis following surgery may determine choice of surgical approach. For instance, laparoscopic surgery is often preferred to open surgery due patient satisfaction associated with a minimally invasive approach [Rafiq et al9]. Knowledge of determinants of patient satisfaction regarding scars following orthopaedic surgery could therefore influence choice of surgical approach. The secondary aim of this systematic review was to identify determinants of patients’ satisfaction with their scars and to evaluate the scales used to measure it.

**1.2 Methods**

We aimed to evaluate current subjective scar assessment scales used in orthopaedic surgery using a PRISMA compliant search [Page et al10]. The secondary aim was to identify determinants of patients’ satisfaction with their scars and to evaluate scales that measure this parameter.

*1.2.1 Study eligibility:*

All studies utilizing subjective scar scales to evaluate scar cosmesis, studies utilizing scales to measure patients’ satisfaction with their scars, and studies reporting on determinants of patients’ satisfaction were included. Papers not reporting original data such as literature or systematic reviews were excluded, along with articles for which the full text was not available. Animal studies, studies utilizing objective scar assessment tools only, and those not concerning orthopaedic surgery were also excluded. There were no constraints based on language or publication status.

*1.2.2 Search strategy and data extraction:*

The following electronic databases were searched via OVID from their inception to 10/04/2022, using a PRISMA compliant search strategy: MEDLINE, Embase, Global Health, MIDIRS, PsycARTICLES and APA PsycInfo. Unpublished grey literature and currently registered studies were reviewed using the databases ISRCTN registry, the UK National Research Register Archive, the National Institute for Health Research Portfolio, OpenSIGLE (System for Information on Grey Literature in Europe) and the WHO International Clinical Trials Registry Platform. Conference proceedings from the British Trauma Society, British Orthopaedic Foot and Ankle Society, and British Orthopaedic Association were also searched. The reference lists of included studies were also searched.

The database search was conducted independently by the first and second authors to obtain a list of eligible studies (by applying pre-specified eligibility criteria). Searches were conducted twice for quality assurance. The first search was conducted on 20/09/2021. The search was repeated on 10/04/2022. The search strategy is described in ***Appendix A***.

Data were extracted by the first and second authors. Scales utilized to assess scar cosmesis and patient satisfaction were identified and evaluated. Determinants of patient satisfaction and the tools used to measure it were also identified. The studies’ findings when utilizing different scar assessment scales were summarized and presented.

*1.2.3 Methodological appraisal:*

Level of evidence and risk of bias of included studies were evaluated independently by the first and second authors. The level of evidence of the studies presented was determined with the March 2009 Oxford Centre for Evidence-Based Medicine: Levels of Evidence [Centre for Evidence-Based medicine11]. The Cochrane Collaboration's risk of bias tool was used to assess risk of bias in randomized controlled trials (RCTs) [Sterne et al12]. The Institute of Health Economics case series studies quality appraisal checklist was used to determine the quality of case series [Institute of Health Economics13].

**1.3 Results**

A total of 7066 records were screened in the initial search with 147 potentially eligible articles identified ***(figure C.1)***. One hundred twenty-three articles were excluded on the bases of the pre-specified exclusion criteria. A total of 26 studies, assessing 7130 patients, were reviewed. A narrative synthesis was performed given the heterogenous and qualitative nature of the data. The lack of numerical data prevented quantitative pooled analysis. The narrative synthesis involved a brief description of the scale in question, its advantages and disadvantages, and the identification of determinants of patient satisfaction.

*1.3.1 Study quality assessment:*

The studies included in this review revealed methodological limitations. The findings of the study quality assessment are presented in Table C.1. Of the 26 studies included, only nine carried a low risk of bias [Alvarez-Pinzon et al25; Beausang et al22; Bridwell et al37; Hollander et al26; Khan et al27; Lenzi et al18; Petis et al21; Soldado et al34; Tai et al36], with the rest exhibiting some concerns or a high risk.

Only seven were a RCT in design [Alvarez-Pinzon et al25; Duncan et al31; Khan et al27; Kundra et al35; Livesey et al15; Menkowitz et al19; Yuenyongviwat et al28]. Of these, only Alvarez-Pinzon et al25, Khan et al27 and Livesey et al15 clearly described their randomisation process and concealed the allocation. Menkowitz et al19 and Kundra et al35 did not clarify whether patients or assessors were blinded, whereas Yuenyongviwat et al28 could not blind the patients from the closure methods, nor could they blind the assessor for the Hollander evaluation score. However, the assessor for the patient satisfaction score was blinded . Accordingly, assessor bias could have impacted these three studies’ results. Duncan et al31 did not describe the randomization process.

All the studies with a RCT design reported all outcomes which were initially planned. Kundra et al35 randomized 100 patients, but only received responses from 70. They attempted to contact non-responders, but no further responses were received. Livesey et al15 randomized 90 patients, of which 12 were lost to follow-up due to cancellation or non-attendance. Of Menkowitz’s 41 randomized patients, two were lost to follow-up and one withdrew due to a cardiac arrest not related to the procedure. Alvarez-Pinzon et al25 reported three losses to follow-up of the 50 patients who were randomized. Khan et al27 and Yuenyongviwat et al28 reported no losses to follow-up. Moreover, five of seven studies with a RCT design based their sample sizes on a power calculation [Alvarez-Pinzon et al25; Khan et al27; Livesey et al15; Menkowitz et al19; Yuenyongviwat et al28]. Kundra et al35 and Duncan et al31 did not perform this exercise.

Of the 19 case series included in this review, only four were conducted prospectively [Hollander et al26; Singer et al8; Sundaram et al24; Tai et al36]. The rest involved a retrospective review of patient records, or the assessment of photographs of scars. In conclusion, the retrospective nature of the majority of the studies in this review, their low level of evidence, and concerns regarding their risk of bias hinder their quality.

Six subjective scar assessment scales were identified in the literature. These were the Vancouver Scar Scale (VSS), Patient and Observer Scar Assessment Scale (POSAS), Manchester Scar Scale (MSS), Stony Brook Scar Evaluation Scale (SBSES), Visual Analogue Scale (VAS) and the Hollander Wound Evaluation Scale (HWES) ***(table C.2)***.

*1.3.2 The Vancouver Scar Scale:*

Seventy-three patients with burn scars were assessed by three occupational therapists. Statistical analysis for interrater validity revealed a statistically significant agreement between observers for each parameter. These include skin pigmentation, vascularity, pliability and height (**table C.3**). The scope of each component is relatively small (e.g., pigmentation is measured in a 0-2 scale). A wide range would allow for a greater subjective description of the scar at the expense of interrater validity [Sullivan et al14]. The inter-rater reliability improved with increased familiarity of the scale.

Sullivan et al14 recommended a three to four month trial before implementing this scale in clinical practice. Livesey et al15 used the VSS to evaluate cosmetic appearance of scars following total hip arthroplasty (THA). They found no difference in cosmetic outcomes when closing skin with staples or adhesive. Karlen et al16 utilized it to evaluate scars following lower limb reconstruction. They recruited 25 patients who received femoral or tibial lengthening to reduce limb discrepancy and improve mobility. The mean Vancouver score for the femoral lengthening group was 4.5, and 1.2 for the tibial lengthening group. However, satisfaction with scars was higher in the femoral lengthening group than the tibial lengthening group, despite the former leading to a higher mean score in the VSS. Karlen et al16 attributed this discrepancy to clothing being able to cover thigh scars more easily.

*1.3.3 The Patient and Observer Scar Assessment Scale (POSAS):*

The POSAS was developed by Draaijers et al17 in 2004 for the assessment of burn scars [17,18] *(****table C.4****)*. Menkowitz et al19 used the POSAS to evaluate scar cosmesis following closure of total knee arthroplasty (TKA) incisions with zip or staples. The former consisted of two polyurethane strips attached at either side of the incisions. Closure was achieved with interconnecting nylon straps, enabling adjustable closure tension. At three weeks post-operatively, POSAS favoured zip for appearance, pain, itching, colour, thickness, stiffness, irregularity, vascularity, relief, pigmentation, pliability, surface area and overall observer opinion. However, there was no statistically significant difference in overall subject opinion when closing the skin with zip or staples. At six weeks, POSAS favoured Zip only for colour, stiffness, thickness, pigmentation, vascularity, relief, pliability, surface area and overall observer opinion. The difference between Zip and staple closure in terms of appearance, pain, itching and irregularity were no longer statistically significant at six weeks.

Wilson et al20 examined perception of surgical scars after direct anterior (DAA) or posterior approach (PA) for THA. Seventy-five DAA and 75 PA THA patients underwent scar assessment using the POSAS. All wounds were closed with subcuticular running closure and were secured with skin adhesive glue. Scars were graded closer to normal skin more often for DAA than for PA patients on the POSAS patient-reported scale. More irregularities were recorded for DAA scars than PA scars on the POSAS observer scale. Wilson’s results are contradicted by Petis et al21, who also compared cosmetic outcomes following a DAA or PA to THA. The scars were graded as closer to normal skin more often for the PA than the DAA. Age was not predictive of overall scar opinion on this scale.

*1.3.4 The Manchester Scar Scale (MSS):*

As opposed to the VSS and POSAS, the MSS does not break down colour into components such as redness and pigmentation. Instead, it compares scar colour to that of surrounding skin to increase its ease of use. (***table C.5***) Beausang et al22 found a high correlation between histologic scar assessment scores and clinical assessment scores obtained with the MSS. Ojima et al23 used the MSS to evaluate scars resulting from TKA after having performed longitudinal or transverse incisions. The latter led to lower MSS scores. Mean scar width and sensory disturbance were also smaller in the transverse group than the longitudinal group.

Sundaram et al24 performed TKA using a medial parapatellar incision and trivector arthrotomy in 91 patients. They also carried out this procedure using a midline incision and medial parapatellar arthrotomy in 76 patients. Resulting scar cosmesis was compared between both groups using the MSS, and the difference was not statistically significant.

*1.3.5 The Stony Brook Scar Evaluation Scale (SBSES):*

The SBSES was created by Singer et al8 by assessing scars following traumatic lacerations (***table C.6***). Singer also utilized a 100-mm VAS, marked worst to best scar, to evaluate scars following traumatic lacerations previously assessed by the SBSES. Resulting scores from each scale were compared. There was a statistically significant correlation between VAS and SBSES scores. Interobserver agreement was noted when using the SBSES. Singer excluded the presence or absence of distortion of surrounding tissues from the SBSES due to its poor interobserver reliability and infrequent occurrence. Petis et al21 used the SBSES to evaluate scars following THA. They compared outcomes when following a DAA or a PA. They found no difference for scar width, height, colour, presence of hatch marks and overall scar appearance between both approaches (p>0.05). Alvarez-Pinzon et al25 evaluated the use of a ring retractor in DAA THA scars. Forty-seven patients undergoing THA were randomized to surgery with (n = 23) or without (n = 24) ring retractor. The SBSES and MSS were used to assess cosmesis. No difference in cosmesis scores were noted between both groups at all points during follow-up, except for two weeks post-operatively, when MSS scores were superior in the group receiving a ring retractor.

*1.3.6 The Hollander Wound Evaluation Scale (HWES):*

The HWES was developed by Hollander et al26 evaluating scars due to traumatic wounds (***table C.7***). Hollander et al26 found this scale had substantial interobserver concordance for description, infection and overall cosmetic appearance of traumatic wounds. The HWES considers ‘step-off’ borders, but wound edges even out over time and are rarely seen in the long term [Petis et al21]. Overall satisfactory/unsatisfactory appearance is included despite its moderate interobserver reliability, as it allows the observer to downgrade the scar if one of the other items in the HWES is deficient [Hollander et al26].

Livesey et al15 used the HWES to evaluate cosmetic appearance of scars following THA, and compared cosmetic outcomes following closure of incisions with staples or adhesive. They found all scars scored highly on the HWES. Seventy-one (92%) of all scars were optimal (with a score of 3) but six (8%) were considered sub-optimal with a score of 2. No statistically significant difference was found for the scores between both groups at three months.

Khan et al27 performed a RCT comparing skin closure with 2-octylcyanoacrylate, subcuticular suture, or staples following 102 THAs and 85 TKAs. No statistically significant difference in HWES score was found between the three methods of skin closure at six weeks for neither THA nor TKA.

Yuenyongviwat et al28 randomized 70 patients undergoing TKA into two groups. Thirty-four patients had the upper half of the wound closed with skin staples, and the lower half with simple interrupted nylon suture. The opposite was performed on the remaining 36. This yielded 70 nylon-stitched wounds and 70 skin stapled-wounds for analysis. Difference in the HWES score between both closure methods was not statistically significant.

*1.3.7 The Visual Analogue Scale (VAS):*

The VAS is typically used in clinical practice to quantify pain [Emery et al29]. Quinn et al30 developed a VAS to assess cosmetic results of healed lacerations. This demonstrated good interobserver and intraobserver agreement. The reliability and validity of the VAS was verified by Duncan et al31. As opposed to the previous scales outlined, the VAS is not a standardised scale. The parameters assessed vary between studies.

*1.3.8 Scales measuring patient satisfaction:*

Few scales measuring patient satisfaction with their scars were identified in the literature. The patient component of the POSAS considers perception of the scar, but does not directly measure satisfaction. Durani et al32 constructed a 39-item scale to assess patients’ perceptions of their scars. They created the Patient Scar Assessment Questionnaire. This consisted of five subscales concerning appearance, symptoms, consciousness, and satisfaction with appearance and symptoms. It was designed by analysing scars following a large variety of procedures, including head and neck naevi excision, scar revision surgery, varicose vein removal and cardiothoracic surgery.

Simple categorical scales have been used. Jia et al33 reported on outcomes of open reduction in infants with developmental dysplasia of the hip, and evaluated the patients’ feelings towards their cosmetic appearance with a 4-item scale (very satisfied, satisfied, unsatisfied, very unsatisfied). However, it is not clear whether this scale evaluated scar cosmesis specifically or if it considered overall cosmetic appearance.

Soldado et al34 employed a 4-point Likert scale to rate parents’ level of satisfaction with the cosmetic appearance of their child’s scar following surgery for lateral condylar fracture of the elbow (0 = not satisfied, 1 = satisfied, 2 = very satisfied, 3 = extremely satisfied). All parents claimed to be “very satisfied” with their child’s scar.

In addition, Livesey et al15 reported that patients used a 5-point Likert scale to rate scar appearance compared to expected appearance following THA (1 = much better than expected, 2 = better than expected, 3 = as expected, 4 = worse than expected, 5 = much worse than expected). Skin closure with adhesive or staples were compared. There was no statistically significant difference in patient rating of the appearance of the wound compared to the expected appearance. Livesey et al15 also utilized a VAS to measure patient satisfaction with scars following THA. Patients completed a 100 mm VAS for satisfaction with their scar (0 = extreme dissatisfaction, 100 = complete satisfaction). There was no statistically significant difference in patient satisfaction between both groups.

Kundra et al35 used a linear VAS to assess patient satisfaction (0 = not satisfied, 100 = fully satisfied) in patients undergoing elective hand and wrist surgery. They compared cosmetic outcomes following skin closure using either absorbable or non-absorbable sutures. There was no statistically significant difference between the two groups in terms of patient satisfaction.

Yuenyongviwat et al28 compared patient satisfaction with cosmetic appearance of wounds following TKA with nylon sutures or skin staples. They utilized a 10-point verbal numeric rating scale (0 = least satisfied, 10 = most satisfied), and found no statistically significant difference in patient satisfaction between both closure methods.

*1.3.9 Factors determining patient satisfaction with scars:*

Tai et al36 investigated preoperative patient expectations for hallux valgus surgery in 153 patients. Improved walking was the most important expectation, followed by reduced pain and ability to wear shoes. Improved appearance (straighter toe) was the 10th most important expectation out of the 19 included in the questionnaire distributed to patients. However, this finding pertains bunion surgery only, and scar cosmesis was not included in their questionnaire. Afolayan et al39 contacted 125 patients who had hallux valgus surgery, with a 84% response rate. Of these, 30% experienced scar sensitivity following surgery. Despite this being a concern, 100% of patients would have the surgery again.

Bridwell et al37 recruited 91 sets of patients and their parents, and asked them to complete questionnaires regarding the patients’ upcoming scoliosis surgery. The greatest concern expressed by parents and patients was neurologic deficit. Location and appearance of the scar was the lowest concern (sixth out of the six concerns included in the questionnaire). The main reason for having surgery was to reduce future pain and disability. Eighty-two parents and patients provided answers regarding their preferred location of the scar. The majority preferred a posteriorly placed scar (52.44% of parents and 48.78% of patients) over an anteriorly placed scar (31.71% of parents and 36.59% of patients). 15.85% of parents and 14.63% of patients had no difference in preference. Incision length, location and ability to conceal with clothes were among the reasons for having a scar preference.

Moran et al38 explored the main concerns of 205 patients undergoing THA/TKA. These were graded on a scale of 1 (not concerned) to 4 (very concerned). The greatest concern was cancellation of surgery (mean 2.66), whereas concerns regarding scar problems were amongst the lowest (mean 1.21). Alvarez-Pinzon et al25 evaluated the use of a ring retractor in DAA THA scars. They asked patients to rank outcomes according to their importance (1 = most important, 5 = least important). Lack of hip pain was the most important outcome (mean 1.2), whereas scar cosmesis was the least important outcome (mean 3.9).

**1.4 Discussion**

Multiple subjective scar assessment scales utilized in orthopaedic surgery were identified in the literature. However, they were not created specifically for orthopaedic surgery, nor were they utilized in this specialty exclusively. The VSS and POSAS were designed for the assessment of burn scars. The SBSES, HWES and VAS (the latter as developed by Quinn et al30 in 1995) were constructed by assessing scars due to traumatic lacerations. Despite this, their validity and reliability have been verified, and they were able to successfully assess scars resulting from orthopaedic procedures. These were TKA, THA, and limb reconstruction. No other orthopaedic operations were identified in the literature for which the listed scar scales were used. This could be attributed to large scars resulting from these procedures (TKA, THA, and limb reconstruction) compared to less invasive interventions such as arthroscopy. Large scars could have psychological effects on patients, warranting an evaluation of cosmesis. However, there are multiple operations which yield large scars for which the listed scar assessment scales were not used. Examples include total shoulder arthroplasty, ankle arthroplasty, open reduction and internal fixation of fractures, and total elbow arthroplasty. The scar scales evaluated in this review should be utilized in other procedures to further test their reliability.

Patient satisfaction is commonly used to measure the quality of health care. It can affect clinical outcomes, patient retention and medical malpractice claims. Patient satisfaction is an indicator of treatment success [Prakash et al40]. Despite this, studies identified in this review revealed that cosmetic appearance of scars did not rank highly among patients’ concerns with orthopaedic surgery outcomes [Alvarez-Pinzon et al25; Bridwell et al37; Moran et al38; Tai et al36]. Their concordance strengthens this claim. In addition, one study found that scores in the VSS following femoral or tibial lengthening did not correlate with patient satisfaction [Karlen et al16]. Patients receiving femoral lengthening had higher VSS scores than those undergoing tibial lengthening. However, patient satisfaction was higher in the former. Karlen et al16 attributed this discrepancy to clothing being able to cover thigh scars more easily. This attests to the inability of current scar scales to adequately assess patient satisfaction, which can be affected by scar positioning. However, Karlen’s study is a case series of a small sample size (25 patients), which hinders the validity of its findings.

The extent of the effect of scar cosmesis on patient satisfaction with orthopaedic procedures could be limited. This specialty mostly involves operating on joints. This leads to certain factors carrying a higher weight than they would in other specialties, such as scars restricting movement. In addition, patients allocate a high importance to restoration of function. This can lead to them assigning a low priority to scar cosmesis, as long as movement impairment is reduced. This is particularly relevant for TKA or THA. These are invasive operations, typically performed in elderly patients. With potentially low aesthetic expectations, and a high importance placed on function restoration, outcomes regarding satisfaction would be positive despite a large scar. This balancing act could explain scar cosmesis ranking among the lowest concerns regarding orthopaedic surgery.

Certain considerations regarding scars following orthopaedic procedures do not apply to other specialties. Examples include movement restriction and variable scar location and visibility. Therefore, scales specific to orthopaedic surgery measuring patient satisfaction with scars are necessary. This review identified a small number of scales that achieve this. The patient component of the POSAS does not measure level of satisfaction. The Patient Scar Assessment Questionnaire is limited to linear scars and is not specific to orthopaedic surgery. The use of simple Likert scales and the VAS exemplifies the lack of a standardized approach to assessing patient satisfaction with their scars. Understanding its determinants can aid the creation of scales measuring this parameter.

Kim et al41 claim that multiple factors can contribute to an undesirable scar. These include the patient’s ethnicity, surgical techniques used, postoperative infections and anatomical location of the incision. For instance, a scar on a patient’s face can deeply influence patients, disturbing their well-being and reducing their social role. Further research is required to validate these claims. This must come in the form of questionnaires answered by a large number of patients. The findings of such research can be utilized to create higher quality scales that measure patient satisfaction with scars following orthopaedic surgery.

**1.5 Conclusion**

The assessment of scars in orthopaedic surgery relies on scales that were not designed specifically for this specialty. However, these were able to appropriately assess scars in the studies identified in this review. The scar assessment scales included in this review must be utilized in other orthopaedic operations to further test their reliability. Current evidence suggests the effect of scar cosmesis on patient satisfaction with orthopaedic procedures is limited. This could be attributed to factors such as restored mobility and functionality carrying a higher weight than scar cosmesis.

**1.6 References:**

1. Ferguson MWJ, Whitby DJ, Mamta S, Armstrong J, Siebert JW, Longaker MT. Scar formation: The spectral nature of fetal and adult wound repair. Plastic Reconstr Surg. 1995;97(4):854-860.

2. Bayat A, McGrouther DA, Ferguson MWJ. Skin scarring. BMJ. 2003;326(7380):88-92. doi: [10.1136/bmj.326.7380.88](https://dx.doi.org/10.1136%2Fbmj.326.7380.88)

3. Verhaegen PDHM, van der Wal MBA, Middelkoop E, van Zuijlen PPM. Objective scar assessment tools: a clinimetric appraisal. Plast Reconstr Surg. 2011;127(4):1561-1570. doi: [10.1097/PRS.0b013e31820a641a](https://doi.org/10.1097/prs.0b013e31820a641a)

4. Fearmonti R, Bond J, Erdmann D, Levinson H. A Review of Scar Scales and Scar Measuring Devices. Eplasty. 2010;10:e43.

5. Grabowski G, Pacana MJ, Chen E. Keloid and Hypertrophic Scar Formation, Prevention, and Management: Standard Review of Abnormal Scarring in Orthopaedic Surgery. J Am Acad Orthop Surg. 2020;28(10):e408-e414. doi: 10.5435/JAAOS-D-19-00690.

6. Morton LM, Phillips TJ. Wound healing and treating wounds: Differential diagnosis and evaluation of chronic wounds. J Am Acad Dermatol. 2016;74(4):589-605. doi: <https://doi.org/10.1016/j.jaad.2015.08.068>

7. Brown BC, Moss TP, McGrouther DA, Bayat A. Skin scar preconceptions must be challenged: Importance of self-perception in skin scarring. J Plast Reconstr Aesthet Surg. 2010;63(6):1022-1029. doi: <https://doi.org/10.1016/j.bjps.2009.03.019>

8. Singer AJ, Arora B, Dagum A, Valentine SRN, Hollander JE. Development and Validation of a Novel Scar Evaluation Scale. Plast Reconstr Surg. 2007;120(7):1892-1897. doi: 10.1097/01.prs.0000287275.15511.10

9. Rafiq MS, Khan MM. Scar Pain, Cosmesis and Patient Satisfaction in Laparoscopic and Open Cholecystectomy. J Coll Physicians Surg Pak. 2016;26(3):216-219.

10. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372(71). <https://doi.org/10.1136/bmj.n71>

11. Centre for Evidence-Based Medicine (2020) Oxford Centre for Evidence-Based Medicine: Levels of Evidence (March 2009). Centre for Evidence-Based Medicine. [https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009. Accessed 14 July 202](https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009.%20Accessed%2014%20July%202021)2

12. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366**:**l4898

13. Institute of Health Economics (2014) Case Series Studies Quality Appraisal Checklist. Institute of Health Economics. [https://www.ihe.ca/research-programs/rmd/cssqac/cssqac-about. Accessed 14 July 20](https://www.ihe.ca/research-programs/rmd/cssqac/cssqac-about.%20Accessed%2014%20July%2020)22

14. Sullivan T, Smith J, Kermode J, McIver E, Courtemanche DJ. Rating the burn scar. J Burn Care Rehabil. 1990;11:256–60.

15. Livesey C, Wylde V, Descamps S, Estela CM, Bannister GC, Learmonth ID, Blom AW. Skin closure after total hip replacement: a randomised controlled trial of skin adhesive versus surgical staples. J Bone Jt Surg Br. 2009;91(6):725-729. doi: <https://doi.org/10.1302/0301-620X.91B6.21831>

16. Karlen LKP, Yinusa W, Yan LS, Wang KW, Hoi LY, John LCY. Analysis of Scar Formation After Lower Limb Lengthening: Influence on Cosmesis and Patient Satisfcation. J Pediatr Orthop. 2004;24(6):706-710.

17. Draaijers LJ, Tempelman FRH, Botman YAM, Tuinebreijer WE, Middelkoop E, Kreis R, van Zuijlen PPM. The Patient and Observer Scar Assessment Scale: A Reliable and Feasible Tool for Scar Evaluation. Plast Reconstr Surg. 2004;113(7):1960-1965. doi: 10.1097/01.PRS.0000122207.28773.56.

18. Lenzi LGS, Santos JBG, Raduan Neto J, Fernandes CH, Faloppa F. The Patient and Observer Scar Assessment Scale: Translation for portuguese language, cultural adaptation, and validation. Int Wound J. 2019;16(6):1513-1520. doi: <https://doi.org/10.1111/iwj.13228>

# 19. Menkowitz B, Olivieri G, Belson O. Patient Satisfaction and Cosmetic Outcome in a Randomized, Prospective Study of Total Knee Arthroplasty Skin Closure Comparing Zip Surgical Skin Closure with Staples. Cureus. 2020;12(1):e6705. doi: [10.7759/cureus.6705](https://dx.doi.org/10.7759%2Fcureus.6705).

20. Wilson JM, Petis SM, Pagnano MW, Sierra RJ, Trousdale RT, Taunton MJ. Scar Perception After Two Surgical Approaches for Total Hip Arthroplasty. Arthroplast Today. 2022;14:96-99. doi: <https://doi.org/10.1016/j.artd.2022.01.028>

21. Petis SM, Brown TS, Pagnano MW, Sierra RJ, Trousdale RT, Taunton MJ. Scar perception following direct anterior versus miniposterior approach for total hip arthroplasty. Orthop proc. 2018;100(13).

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

23. Ojima T, Yoshimura M, Katsuo SI, Mizuno K, Yamakado K, Hayashi S, Tsuchiya H. Transverse incision advantages for total knee arthroplasty. J Orthop Sci. 2011;16(5):524-530. doi: 10.1007/s00776-011-0133-4.

24. Sundaram RO, Ramakrishnan M, Harvey RA, Parkinson RW. Comparison of scars and resulting hypoaesthesia between the medial parapatellar and midline skin incisions in total knee arthroplasty. Knee. 2007;14(5):375-378. doi: <https://doi.org/10.1016/j.knee.2007.06.002>.

25. Alvarez-Pinzon AM, Mutnal A, Suarez JC, Jack M, Friedman D, Barsoum WK, Patel PD. Evaluation of wound healing after direct anterior total hip arthroplasty with use of a novel retraction device. Am J Orthop. 2015;44(1):e17-24.

26. Hollander JE, Singer AJ, Valentine S, Henry MC. Wound registry: Development and validation. Ann Emerg Med. 1995;25(5):675-684. doi: <https://doi.org/10.1016/S0196-0644(95)70183-4>

27. Khan RJK, Fick D, Yao F, Tang K, Hurworth M, Nivbrant B, Wood D. A comparison of three methods of wound closure following arthroplasty. A prospective, randomised, controlled trial. The J Bone Jt Surg Br. 2006;88(2). doi: <https://doi.org/10.1302/0301-620X.88B2.16923>

28. Yuenyongviwat V, Iamthanaporn K, Hongnaparak T, Tangtrakulwanich B. A randomised controlled trial comparing skin closure in total knee arthroplasty in the same knee: nylon sutures versus skin staples. Bone Jt Res. 2016;5(5). doi: <https://doi.org/10.1302/2046-3758.55.2000629>

29. Emery EE, Woodhead EL, Molinari V, Hunt MG. Handbook of Assessment in Clinical Gerontology. 2nd ed. Lichtenberg, PA: Elsevier Academic Press; 2010

30. Quinn JV, Drzewiecki AE, Steill IG, Elmslie TJ. Appearance scales to measure the cosmetic outcomes of healed lacerations. Am J Emerg Med. 1995;13(2):229-231.

31. Duncan JAL, Bond JS, Mason T, Ludlow A, Cridland P, O’Kane S, Ferguson MWJ. Visual Analogue Scale scoring and ranking: a suitable and sensitive method for assessing scar quality? Plast Reconstr Surg. 2006 118(4):909–918.

32. Durani P, McGrouther DA, Ferguson MWF. The Patient Scar Assessment Questionnaire: A Reliable and Valid Patient-Reported Outcomes Measure for Linear Scars. Plast Reconstr Surg. 2009;123(5):1481-1489. doi: 10.1097/PRS.0b013e3181a205de

33. Jia G, Wang E, Lian P, Liu T, Zhao S, Zhao Q. Anterior approach with mini-bikini incision in open reduction in infants with developmental dysplasia of the hip. J Orthop Surg Res. 2020 15(1):180.

# 34. Soldado FF, Domenech-Fernandez P, Barrera-Ochoa S, Bergua-Domingo JM, Diaz-Gallardo P, Hodgson F, Knorr J. Transverse Anterior Approach to the Elbow for Pediatric Displaced Lateral Humeral Condyle Fractures. Arch Bone Jt Surg. 2020;8(2):142-146. doi: [10.22038/abjs.2019.30756.1797](https://dx.doi.org/10.22038%2Fabjs.2019.30756.1797)

# 35. Kundra RK, Newman S, Saithna A, Lewis AC, Srinivasan S, Srinivasan K. Absorbable or non-absorbable sutures? A prospective, randomised evaluation of aesthetic outcomes in patients undergoing elective day-case hand and wrist surgery. Ann R Coll Surg Engl. 2010;92(8):665-667. doi: [10.1308/003588410X12699663905113](https://dx.doi.org/10.1308%2F003588410X12699663905113)

36. Tai CC, Ridgeway S, Ramachandran M, Ng VA, Devic N, Singh D. Patient Expectations for Hallux Valgus Surgery. J Orthop Surg. 2008;16(1):91-95. doi: [https://doi.org/10.1177/230949900801600121](https://doi.org/10.1177%2F230949900801600121)

37. Bridwell KH, Shufflebarger HL, Lenke LG, Lowe TG, Betz RR, Bassett GS. Parents’ and Patients’ Preferences and Concerns in Idiopathic Adolescent Scoliosis. Spine. 2000;25(18):2392-2399.

38. Moran M, Khan A, Sochart DH, Andrew G. Evaluation of patient concerns before total knee and hip arthroplasty. J of Arthroplast. 2003;18(4):442-445. doi: <https://doi.org/10.1016/S0883-5403(03)00061-5>

39. Afolayan J, Ieong E, Akere C, Little N, Pearce C, Solan M. The incidence and natural history of scar sensitivity following hallux valgus surgery; addressing patients’ concerns. Int J Surg. 2011;9(7):530. doi: <https://doi.org/10.1016/j.ijsu.2011.07.174>

40. Prakash B. Patient Satisfaction. J Cutan Aesthet Surg. 2010;3(3):151-155. doi: [10.4103/0974-2077.74491](https://dx.doi.org/10.4103%2F0974-2077.74491).

41. Kim SH, Lee SJ, Lee JW, Jeong HS, Suh IS. Clinical trial to evaluate the efficacy of botulinum toxin type A injection for reducing scars in patients with forehead laceration: A double-blinded, randomized controlled study. Med. 2019;98(34):e16952. doi: 10.1097/MD.0000000000016952

Appendix A (Search strategy):

$ Signifies truncation

Surg$ OR Operat$ OR Procedure

AND

Tool\* OR Scale\*

AND

Trauma OR Orthopedic$ OR Orthopaedic$

AND

Cosmesis OR Cosmetic$ OR Appearance

AND

Scar$

Deduplicate

Appendix B (List of legends):

**Fig. C.1** PRISMA diagram depicting the study collection process

**Table C.1**: Level of evidence and risk of bias of studies included in this review

**Table C.2**: A comparison of subjective scar assessment tools

**Table C.3**: The Vancouver Scar Scale (Sullivan et al, 1990) [14]

**Table C.4**: The Patient and Observer Scar Assessment Scale (Draaijers et al, 2004) [17]

**Table C.5**: The Manchester Scar Scale (Beausang et al, 1998) [22]

**Table C.6**: The Stony Brook Scar Evaluation Scale (Singer et al, 2007) [8]

**Table C.7**: The Hollander Wound Evaluation Scale (Hollander et al, 1995) [26]

Appendix C (figures and tables):

**Fig. C.1** PRISMA diagram depicting the study collection process

Records removed *before screening*:

Duplicate records removed

(n = 1848)

Records marked as ineligible by automation tools (n = 0)

Records removed for other reasons (n = 0)

Records identified from:

Databases (n = 5759)

Registers (n = 222)

Conference proceedings (n = 1501)

Citation searching (n = 425)

**Identification**

Records excluded:

Scar cosmesis not evaluated (n = 23)

Full text not available (n = 12)

Not orthopaedic surgery (n = 64)

Used objective scar scales (n = 3)

Did not report original data (n = 20)

Animal study (n = 1)

Records screened

(n = 6059)

**Screening**

Records assessed for eligibility

(n = 149)

Studies included in review

(n = 26)

**Included**

**Table C.1**: Level of evidence and risk of bias of studies included in this review

|  |  |  |
| --- | --- | --- |
| Study | Level of evidence | Risk of bias |
| Singer et al8 | 4 | Some concerns |
| Quinn et al30 | 4 | Some concerns |
| Sullivan et al14 | 4 | Some concerns |
| Livesey et al15 | 1b | Some concerns |
| Karlen et al16 | 4 | Some concerns |
| Draaijers et al17 | 4 | Some concerns |
| Menkowitz et al19 | 2b | High |
| Petis et al21 | 4 | Low |
| Beausang et al22 | 4 | Low |
| Ojima et al23 | 4 | Some concerns |
| Hollander et al26 | 4 | Low |
| Duncan et al31 | 2b | Some concerns |
| Durani et al32 | 4 | Some concerns |
| Jia et al33 | 4 | Some concerns |
| Soldado et al34 | 4 | Low |
| Kundra et al35 | 2b | High |
| Tai et al36 | 4 | Low |
| Bridwell et al37 | 4 | Low |
| Afolayan et al39 | 4 | Some concerns |
| Khan et al27 | 1b | Low |
| Yuenyongviwat et al28 | 2b | Some concerns |
| Sundaram et al24 | 4 | High |
| Moran et al38 | 4 | Some concerns |
| Wilson et al20 | 4 | Some concerns |
| Lenzi et al18 | 4 | Low |
| Alvarez-Pinzon et al25 | 2b | Low |

**Table C.2**: A comparison of subjective scar assessment tools

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Scar Scale | Scoring System | Parameters Assessed | Advantages | Disadvantages |
| VSS | 0 (best) to 13 (worst) | Skin pigmentation, vascularity, pliability and height | Ease of use | Does not indicate the absolute severity of the pathologic condition of a burn scar |
| POSAS | Observer scale: 5 (best) to 50 (worst)  Patient scale: 6 (best) to 60 (worst) | Observer scale: vascularization, pigmentation, thickness, relief and pliability  Patient scale: scar pain, pruritus, colour, stiffness, thickness and regularity | Takes patients’ perspectives into account, use of a numeric scale with multiple values | Parameters assessed in patient scale may not represent the concerns of every patient |
| MSS | 5 (best) to 28 (worst) | Scar colour, contour, distortion, texture. Matte or shiny appearance, VAS. | Takes multiple cosmetic parameters into account | The high weight of the VAS reduces the impact that specific scar parameters have on the overall score |
| SBSES | 0 (worst) to 5 (best) | Width, height, colour, presence of hatch marks or suture marks and overall appearance | Ease of use (composed of 5 unequivocal items), fast completion | Takes “overall appearance” into account, which is less specific than the other parameters assessed in the SBSES |
| HWES | 0 (best) to 6 (worst) | Step-off borders, contour, margin separation, edge inversion, distortion of surrounding tissues and overall scar appearance | Ease of use, fast completion. Provides technical feedback to practitioner on the quality of their wound repair | The presence or absence of distortion of surrounding tissues has poor interobserver reliability and occurs infrequently. Not suitable for long-term assessment of scars [Singer et al8] |
| VAS | 0 to 10 | Different parameters are assigned to each end of the scale | Ease of use | Not a standardised, difficult to interpret results [Quinn et al30] |

**Table C.3**: The Vancouver Scar Scale [Sullivan et al4]

|  |  |  |
| --- | --- | --- |
| Parameter | Scar characteristic | Score |
| Vascularity | Normal | 0 |
| Pink | 1 |
| Red | 2 |
| Purple | 3 |
| Pigmentation | Normal | 0 |
| Hypopigmentation | 1 |
| Hyperpigmentation | 2 |
| Pliability | Normal | 0 |
| Supple | 1 |
| Yielding | 2 |
| Firm | 3 |
| Banding | 4 |
| Contracture | 5 |
| Height | Flat | 0 |
| <2 mm | 1 |
| 2 - 5 mm | 2 |
| >5 mm | 3 |

**Table C.4**: The Patient and Observer Scar Assessment Scale [Draaijers et al17]

|  |  |
| --- | --- |
| Observer Scar Assessment Scale | |
| Parameter | 1 (normal skin) to 10 (worst scar imaginable) |
| Vascularization | Rate 1 to 10 |
| Pigmentation | Rate 1 to 10 |
| Thickness | Rate 1 to 10 |
| Relief | Rate 1 to 10 |
| Pliability | Rate 1 to 10 |
| Total Observer Score | 5 to 50 |
| Patient Scar Assessment Scale | |
| Is the Scar Painful? | 1 (no complaints) to 10 (worst imaginable) |
| Is the Scar Itching? | 1 (no complaints) to 10 (worst imaginable) |
| Is the colour of the scar different? | 1 (as normal skin) to 10 (very different) |
| Is the scar more stiff? | 1 (as normal skin) to 10 (very different) |
| Is the thickness of the scar different? | 1 (as normal skin) to 10 (very different) |
| Is the scar irregular | 1 (as normal skin) to 10 (very different) |
| Total Patient Score | 6 to 60 |

**Table C.5**: The Manchester Scar Scale [Beausang et al22]

|  |  |  |
| --- | --- | --- |
| Parameter | Scar Characteristic | Score |
| Colour (compared to surrounding skin) | Perfect | 1 |
| Slight mismatch | 2 |
| Obvious mismatch | 3 |
| Gross mismatch | 4 |
| Matte or Shiny | Matte | 1 |
| Shiny | 2 |
| Contour | Flush with surrounding skin | 1 |
| Slightly proud/indented | 2 |
| Hypertrophic | 3 |
| Keloid | 4 |
| Distortion | None | 1 |
| Mild | 2 |
| Moderate | 3 |
| Severe | 4 |
| Texture | Normal | 1 |
| Just palpable | 2 |
| Firm | 3 |
| Hard | 4 |
| Visual Analogue Scale | 0 (excellent) to 10 (poor) | |

**Table C.6**: The Stony Brook Scar Evaluation Scale [Singer et al8]

|  |  |  |
| --- | --- | --- |
| Parameter | Scar Characteristic | Score |
| Width | >2 mm | 0 |
| ≤2 mm | 1 |
| Height | Elevated or depressed | 0 |
| Flat | 1 |
| Colour | Darker than surrounding skin | 0 |
| Same colour or lighter than surrounding skin | 1 |
| Hatch marks or suture marks | Present | 0 |
| Absent | 1 |
| Overall appearance | Poor | 0 |
| Good | 1 |

**Table C.7**: The Hollander Wound Evaluation Scale [Hollander et al26]

|  |  |  |
| --- | --- | --- |
| Scar attribute | Scar characteristic | Score |
| Step-off borders | Absent | 0 |
| Present | 1 |
| Contour irregularities | Absent | 0 |
| Present | 1 |
| Margin Separation | Absent | 0 |
| Present | 1 |
| Edge Inversion | Absent | 0 |
| Present | 1 |
| Excessive distortion of surrounding tissue | Absent | 0 |
| Present | 1 |
| Overall appearance | Satisfactory | 0 |
| Unsatisfactory | 1 |

Appendix D (footnotes):

Footnote 1 (Abbreviations): VSS: Vancouver Scar Scale; POSAS: Patient and Observer Scar Assessment Scale; MSS: Manchester Scar Scale; SBSES: Stony Brook Scar Evaluation Scale; VAS: Visual Analogue Scale; HWES: Hollander Wound Evaluation Scale

1. VSS: Vancouver Scar Scale; POSAS: Patient and Observer Scar Assessment Scale; MSS: Manchester Scar Scale; SBSES: Stony Brook Scar Evaluation Scale; VAS: Visual Analogue Scale; HWES: Hollander Wound Evaluation Scale [↑](#footnote-ref-1)