Published in final edited form as:

J Eval Clin Pract. 2011 December ; 17(6): 1075–1082. doi:10.1111/j.1365-2753.2010.01476.x.

Quantifying the Aesthetic Outcomes of Breast Cancer Treatment: Assessment of Surgical Scars from Clinical Photographs

Min Soon Kim¹, William N. Rodney², Gregory P. Reece³, Elisabeth K. Beahm³, Melissa A. Crosby³, and Mia K. Markey²

¹Center for Biomedical Informatics, Mount Sinai School of Medicine, New York, NY

²The University of Texas Department of Biomedical Engineering

³Department of Plastic Surgery, The University of Texas M. D. Anderson Cancer Center

Abstract

Accurate assessment of the degree of scarring that results from surgical intervention for breast cancer would enable more effective pre-operative counseling. The resultant scar that accompanies an open surgical intervention may be characterized by variance in thickness, color, and contour. These factors significantly impact the overall appearance of the breast. A number of studies have addressed the mechanical and pathologic aspects of scarring. The majority of these investigations have focused on the physiologic process of scar formation and means to improve the qualities of a scar. Few studies have focused on quantifying the visual impact of scars. This manuscript critically reviews current methods used to assess scars in terms of overall satisfaction after surgery. We introduce objective, quantitative measures for assessing linear breast surgical scars using digital photography. These new measurements of breast surgical scars are based on calculations of contrast and area. We demonstrate, using the intra-class correlation coefficient (ICC), that the new measures are robust to observer variability in annotating the scar region on clinical photographs. As an example of the utility of the new measures, we use them to quantify the aesthetic differences of reconstruction following skin-sparing mastectomy vs. conventional mastectomy.

Keywords

Aesthetics; Breast Neoplasm; Esthetics; Mastectomy; Outcomes; Prostheses and Implants; Reconstructive Surgical Procedures; Surgical Flaps; Surgical Scars; Treatment Outcome; Quality of Life; Breast Conservation Therapy

INTRODUCTION

One in eight women will be diagnosed with breast cancer in the United States, with an estimated 250,000 new cases in 2010¹. Currently, some form of surgical intervention remains an imperative in adequate breast cancer treatment, ranging from a breast conserving surgery with partial removal of the breast to a total mastectomy. The choice of surgical procedure depends on patient preference as well as the stage and location of the disease^{2, 3}. For many patients with breast cancer, breast reconstruction either immediately or at a later

Mia K. Markey, Ph.D., The University of Texas Department of Biomedical Engineering, 1 University Station C0800, Austin, TX 78712, +1.512.471.1711, +1.512.471.0616 (fax), mia.markey@mail.utexas.edu.

time after surgical intervention for the cancer is an option. Approximately 57,000 breast reconstructive surgeries are performed annually in the US ⁴.

There are a number of approaches to restoring a woman's appearance after partial or total mastectomy, ranging from prosthetic devices to surgical reconstruction with autologous tissues. The complexity and length of surgery vary, and different patients are suited to or desire different options. The reconstructive surgeon attempts to restore the breast appearance to one that is satisfying to the patient, taking into account the patient's disease state, general health, physiognomy, and preferences. Unfortunately, surgery, including reconstructive surgery, results in scarring that can be disfiguring, aesthetically unpleasant, and that can strongly affect a patient's quality of life ⁵⁻⁷. Scars have been linked to psychosocial problems such as development of post-traumatic stress reactions, loss of self-esteem, and stigmatization ^{6, 8}.

Cutaneous scars have different physical characteristics than unsullied skin. Scars manifest with different color, texture, and firmness than surrounding skin. Some scars may be thin and flat, while others are elevated, depressed, thick, or wide ⁹. Scars have been categorized by their characteristics. These categories in certain cases may reflect causality and have implications for the potential correction of the scars. The five main categories of scars are defined by the growth and characteristics of the scar tissue ⁵. Widened scars usually appear around six weeks after surgery, as the initial fine surgical scars are noted to increase and appear stretched and widened. Atrophic scars are generally small and often round with a concave center that lies below the surrounding skin and creates a contour deformity. Scar contracture is a function of healing by scarring that causes shortening of the scars and distortion. These scars can cause functional problems if they cross joints or skin creases as they may limit the range of motion. Hypertrophic scars are raised scars that remain within the boundaries of the original closed incision. Keloid scars are raised scars that spread beyond the margins of the original wound and invade the surrounding skin in a way that is site specific. The distinction between these latter two forms of scarring is important, as surgical correction of hypertrophic scars is usually possible with excision, whereas excision of a keloid can result in recurrence of a larger lesion that requires radiation or other treatments.

Adequate measures of aesthetic outcomes, including scar appearance, are important to protect patients' health and well-being ¹⁰. The majority of observer assessments of the appearance of surgical scars by clinicians ¹¹, patients ¹², or combinations of both ¹³ have only been studied as subscales of overall aesthetic outcome measures (Table 1). Visibility, the extent to which the surgical scar detracts from the aesthetic outcome, is the relevant property for assessing scarring following breast reconstructive surgery. The appearance of scars on the reconstructed breast mound is typically described using a 3-point scale (good, fair, poor) ^{11, 14}, 4-point scale (excellent, good, fair, poor) ¹⁵, or 5-point scale (excellent, good, satisfactory, poor, unacceptable). However, such broad, poorly-defined ratings have unacceptable inter-rater reliability ^{11, 16}. Moreover, using a single scale does not consider all of the relevant aesthetic features of scars.

Quantitative assessment methods have been introduced to study the mechanical properties, size, and color of the scar region (Table 2). While mechanical properties of scars influence the patient's comfort, they do not usually impact the visibility, and thus the aesthetics, of scars. The visibility of a scar is significantly influenced by its vascularity, pigmentation, and surface area ¹⁷. Thus, methodology for measuring mechanical properties is not reviewed in this paper, but instead current methods for evaluating scar color and surface area are summarized.

Color assessment

Scar color is influenced by the amounts of oxyhemoglobin (“vascularity”) as well as the bile and carotene pigments (“pigmentation”). Coloring changes as healing progresses, with immature scars appearing red due to increased vascularity. Therefore, formal evaluation of the appearance of the scar is usually deferred until 6-12 months after a surgical intervention. An evaluation and record of these color changes would be appropriate when considering aesthetic results, particularly once scar maturity has occurred¹⁸. However, it is sometimes difficult to separate the contribution of the two color parameters: pigmentation and vascularity. Following the example of Beausang et al.¹⁹, we propose the use of only one scale for evaluating discrepancies in color between scars and other skin. In other words, we used the assessment of the scar’s intensity gradient (contrast) with surrounding skin, rather than trying to fragment this parameter into its individual components of vascularity and pigmentation. This avoids the problem of trying to determine what proportion of the overall color is attributable to each component.

Area assessment

A few studies have investigated the applicability of simple methods for measuring the size of scars for management and research. The inter-observer reliability and accuracy (validity) was established for planimetry (area of a 2-D figure) by photography and by tracing the scar outline on a transparent sheet in a study by Zuijlen et al.²⁰. Drawings of 25, 50, and 75 cm² regions were created on 3 locations with increasing curvature (back, thigh, and forearm) on 20 healthy volunteers. While, both techniques showed a good reliability for 25 cm² areas (intraclass correlation coefficient, ICC $\geq$ 0.82), planimetry by photography was more reliable than planimetry by tracings for the 50 and 75 cm² cm areas. Also, planimetry by photography was more accurate than planimetry by tracing for all areas except for the area with the greatest curvature, the forearm. Thus, the work of Zuijlen et al. suggests that planimetry by photography is as suitable for surface area measurements as planimetry by tracing even for a surface with substantial curvature, such as the breast area.

Study overview

In this paper, we propose quantitative, objective measurements of the intensity gradient (contrast) and surface area of breast surgical scars that can be computed from digital clinical photographs that account for changes in physical characteristics of the breast, such as the size of the scar relative to the breast region size. This assessment employs ratios for patient-specific evaluations. We used the intra-class correlation coefficient to quantify the reliability of the new objective measures given observer variability in annotating the scar region on clinical photographs. Finally, as example of their practical utility, the proposed measures were used to quantify the aesthetic outcomes of breast reconstruction following skin-sparing mastectomy relative to those of breast reconstruction following conventional mastectomy.

MATERIALS AND METHODS

Dataset

This retrospective analysis of clinical photographs was approved by the Institutional Review Boards (IRB) of The University of Texas at Austin and The University of Texas M. D. Anderson Cancer Center (UT MDACC). The patient population for this study was women aged 21 years or older who underwent breast reconstruction with a lower abdominal flap surgery including a Transverse Rectus Abdominis Myocutaneous (TRAM) flap, a DIEP (Deep Inferior Epigastric Perforator) or SIEA (Superficial Inferior Epigastric Artery) from January 1, 1990 to June 1, 2003 at UT MDACC. Since all three procedures employed a “lower abdominal flap”, the shorter term “flap” is used throughout this paper. The data pool

consisted of digitized conventional 35mm photographs and digital photographs. The conventional photographs were digitized with a Nikon LS 2000 (Nikon, Tokyo, Japan) or Nikon Super Coolscan 4000ED (1.06) slide digitizer (Nikon, Tokyo, Japan). The digital cameras used were a Nikon 990 Coolpix (Nikon, Tokyo, Japan) or Cannon T90 35mm SLR with 50mm lens (Canon, Tokyo, Japan). All photographs were taken against a sky blue background, following standard guidelines for clinical photography in plastic surgery²¹. A forward facing flash unit, but no studio umbrella flash, was used in a room lit by overhead fluorescent lights.

The data set for this study consisted of 40 post-operative anterior-posterior views. The images were selected by an experienced plastic surgeon (GPR) to demonstrate a large variety of breast aesthetic characteristics, such as size, symmetry, ptosis, and projection. Out of the 40 anterior/posterior images used, 16 patients underwent a conventional mastectomy followed by flap breast reconstruction and 24 patients had skin-sparing mastectomy followed by flap breast reconstruction. Photographs of patients who undergo flap breast reconstruction following a conventional mastectomy generally do not show the complete scar area. In most cases, significant parts of scars are hidden below the breast mound or on the lateral side. Patients who undergo a flap breast reconstruction after a skin sparing mastectomy show very little or no skin of the flap or scarring in an anterior-posterior photograph because the vast majority of breast incisions are made around the areola and extended if required. Very few incisions are made along the inframammary fold. The native breast skin, therefore, covers almost the entire flap and the pigments placed when the reconstructed nipple and areola are tattooed camouflages the exposed flap skin (Figure 1).

Quantitative measures of intensity gradient (contrast) and area

Two measurements chosen for their substantial relevance to scar assessment were calculated from Regions of Interest (ROIs) that were manually delineated. The size of the scar was computed as the ratio of the area of the scar region to the total area of the affected breast (AR) using ImageJ²², an open image analysis software used to delineate the ROIs on the digital photographs. ROIs were manually identified on each image by two non-clinician observers (MSK, WNR) on the same image set on two occasions two weeks apart under the supervision of an experienced clinician (GPR). They manually delineated ROIs using the polygon selection tool in ImageJ by looking at the images at 400% zoom ratio on the computer screen with consistent resolution and size (1490×900) throughout the study.

The intensity gradient (contrast) of the scar versus the surrounding skin was computed as the ratio of the absolute value of the difference in mean grayscale intensity of the scar and surrounding skin regions to the average intensity of the surrounding skin region (SIG). For patients who underwent a flap breast reconstruction after a conventional mastectomy, the intensity of the scar region was also compared to that of the exposed skin of the flap (FIG) in order to investigate the contrast difference between these two regions.

$$SIG = \frac{|I_{scar} - I_{surrounding}|}{I_{surrounding}} \quad (1)$$

$$FIG = \frac{|I_{scar} - I_{flap}|}{I_{flap}} \quad (2)$$

The size of the scar is computed as a ratio of the area of the scar region to the total area of the affected breast (AR) based on the manual delineation of the ROIs. In other words, the proportion of the breast affected by scarring is quantified.

$$AR = \frac{A_{scar}}{A_{affected}} \quad (3)$$

Intra- and inter- observer variability: intraclass correlation coefficient

To quantify the intra-observer agreement within measurements made by each of two observers, the Intraclass Correlation Coefficient (ICC) was computed using SPSS (SPSS Inc). Because the set of images was a random subset of images from the class of representative cases of post-operative images and the observers were randomly selected from the population of observers, a single-score two-way random model was used.

$$ICC = \frac{MS_R - MS_E}{MS_R + (k - 1) MS_E + \frac{k}{n} (MS_C - MS_E)}$$

where n denotes the number of photographs, k denotes the number of measurement, MS_R is the mean square error, MS_E is the residual mean square error, and MS_C is the mean square error between measurements.

An average-score two way random model ICC was used for analysis of inter-observer agreement, i.e., the agreement between the two observers. The average option was chosen because the averages of two measurements made by each observer were used.

$$ICC = \frac{MS_R - MS_E}{MS_R + \frac{MS_C - MS_E}{n}}$$

where n denotes the number of photographs, MS_R is the mean square error, MS_E is the residual mean square error, and MS_C is the mean square error between measurements.

The guidelines for interpretation of the ICC used are as follows: an ICC value of less than 0.40 indicates poor reproducibility, ICC values in the range 0.40 to 0.75 indicate fair to good reproducibility, and an ICC value of greater than 0.75 shows excellent reproducibility²³.

Flap breast reconstruction following a conventional mastectomy vs. flap breast reconstruction following a skin-sparing mastectomy

As an example of their practical utility, the new objective measures were used to quantify the aesthetic outcomes of flap breast reconstruction following skin-sparing mastectomy vs. flap breast reconstruction following conventional mastectomy. An F-test was performed to test for the equality of variances of the two groups. Based on the F-test result, a suitable two-sample t-test was performed to analyze the difference between the mean values of the objective measures of the two groups.

RESULTS

Intensity gradient (contrast) assessment: intra-observer agreement

The intra-observer agreement as assessed by ICC between the two sets of measurements made by each of two observers on the cases of flap breast reconstruction following a

conventional mastectomy was very good (Table 3). For Observer 1 (MSK), the intra-observer agreement was “excellent” for FIG (ICC = 0.94) and “fair to good” for SIG (ICC = 0.68). Similarly, for Observer 2 (WNR), the intra-observer agreement was “excellent” for both FIG (ICC = 0.95) and SIG (ICC = 0.84). Likewise, the intra-observer agreement on the cases of flap breast reconstruction following skin-sparing mastectomy was very good (Table 4). The intra-observer agreement for SIG on the flap breast reconstruction following skin-sparing mastectomy cases was “fair to good” for Observer 1 (MSK) (ICC = 0.61) and “excellent” for Observer 2 (WNR) (ICC = 0.83).

Area assessment: intra-observer variability

The intra-observer agreement between the two sets of area measurements as evaluated by using ICC was very good. For the area measurements made by Observer 1 (MSK) on the cases of flap breast reconstruction following a conventional mastectomy, the agreement was “fair to good” (ICC = 0.72). Likewise, the ICC was “excellent” for Observer 2 (WNR) (ICC = 0.83) (Table 3). The intra-observer agreement between the two sets of measurements made by each observer on the cases of flap breast reconstruction following a skin-sparing mastectomy using the ICC coefficient showed mixed variability (Table 4). The reproducibility for Observer 1 (MSK)’s AR was “poor” (ICC = 0.18), while the intra-observer agreement was “good” for Observer 2 (WNR)’s AR (ICC = 0.63).

Intensity gradient (contrast) assessment, inter-observer variability

The inter-observer agreement between the intensity gradient (contrast) measurements made by Observer 1 (MSK) and Observer 2 (WNR) on the cases of flap breast reconstruction following a conventional mastectomy was “fair to good” for SIG (ICC = 0.50) and “excellent” for FIG (ICC = 0.95). Similarly, for flap breast reconstruction following skin-sparing mastectomy cases, the inter-observer agreement was “fair to good” for SIG (ICC = 0.62).

Area assessment, inter-observer variability

As evaluated by ICC, the inter-observer agreement between the area measurements made by Observer 1 (MSK) and Observer 2 (WNR) on the cases of flap breast reconstruction following a conventional mastectomy was “excellent” (ICC = 0.81). “Fair to good” agreement in the area measurements was achieved for flap breast reconstruction following a skin-sparing mastectomy (ICC = 0.68).

Flap breast reconstruction following a conventional mastectomy vs. following a skin-sparing mastectomy

Based on the proposed AR measurements, the proportion of the breast area affected by scarring is statistically significantly larger ($p = 0.01$) for flap breast reconstruction following a conventional mastectomy (overall mean = 0.05) than for flap breast reconstruction following a skin-sparing mastectomy (overall mean = 0.03), according to a two sample t-test with unequal variance ($p = 0.04$, F-test). This result is consistent with our expectation that removal of a larger amount of skin in the mastectomy will result larger scar regions.

The intensity gradient (contrast) between the scars and surrounding skin was statistically significantly greater ($p < 0.01$) in flap breast reconstruction following skin-sparing mastectomy (overall SIG = 0.21) than that for flap breast reconstruction following a conventional mastectomy (overall SIG = 0.11), based on a two sample t-test with equal variance ($p = 0.08$, F-test). The reason for the difference in the intensity gradient between the two types of mastectomy may be the amount of pigmentation of the skin used for the reconstruction. Most women do not expose the majority of their breast skin to the sun;

however, many women do expose their abdominal skin to the sun. Thus, a Caucasian patient who undergoes skin sparing mastectomy may have more obvious scars because reconstruction is performed using her non-tanned native breast skin. In comparison, a Caucasian patient who undergoes conventional mastectomy may have less obvious scars because reconstruction is performed using her tanned abdominal skin.

When the two intensity gradient (contrast) measurements within flap breast reconstruction following a conventional mastectomy were compared, the average intensity gradient between scars and flap skin (overall FIG = 0.17) was greater than that between scars and the surrounding skin area (overall SIG = 0.11). This difference was not statistically significant ($p = 0.078$), according to two sample t-test with unequal variance ($p = 0.001$, F-test), though the failure to find a significant difference may be due to the limited sample size of this study (post hoc power = 45%).

DISCUSSION

Breast surgery leaves scars that can strongly affect a patient's feelings about her appearance, and so impact her quality of life. Current methods for assessing surgical scars require employment of mechanical, electronic, or chemical measurements that are expensive and time consuming. In addition, the employment of qualitative observer ratings based on direct viewing of patients has low intra- and inter-observer agreement. Effective, quantitative measures are required to investigate the relationships between surgical variables and aesthetic outcomes. In this study, we developed, evaluated, and employed new measures of intensity gradient (contrast) and area of linear breast surgical scars as computed from digital/digitized photographs.

Digital photography is widely utilized in the plastic and reconstructive surgery field for preoperative planning and record keeping purposes. However, two important points for future studies of surgical scar assessment by digital photography should be noted. First, standardized photography is critical for reliable photographic assessment of scars. Even though the data used in this study were collected following current photographic standards²¹, the authors still had difficulty in delineating scar regions on some patient images because of varying light intensity, distance and direction from the light to the subject, camera lighting settings, background color, and a subject's pose in relation to a camera. Second, all pertinent information on the photographic conditions, including time point of photographs being taken, should be recorded for further analysis and comparison of patients' aesthetic outcomes. Likewise, thorough clinical information accompanying the photographs will help surgeons understand the measurement results.

We propose new measures of intensity gradient (contrast) and area of linear breast surgical scars that can be simply computed from digital photographs. Since the calculation is dependent on the manual delineation of regions of interest in the image, we studied the impact of variation in the identification of the regions on the scar measures. We demonstrated that the new objective measures of scar intensity gradient (contrast) and area showed encouraging inter- and intra-observer agreement despite variability in the manual delineation of the regions of interest on the digital photographs. This suggests that the new measures could be used in practice to provide a standardized means of quantifying the appearance of scars that result from surgical treatment of breast cancer.

As an example of the utility of the new objective measures, we used them to compare the aesthetic outcomes of flap breast reconstruction following skin-sparing mastectomy to those of flap breast reconstruction following conventional mastectomy. The AR measurement suggests that breast area affected by scarring is larger for a flap breast reconstruction

following a conventional mastectomy than for a flap breast reconstruction following a skin-sparing mastectomy, which is consistent with expectations that the skin-sparing procedure covers most of the flap and is performed through a smaller breast incision. We also compared the intensity gradient (contrast) of the scars relative to the surrounding skin for flap breast reconstruction following skin-sparing mastectomy vs. conventional mastectomy. The average intensity gradient (contrast) between the scar and remaining native breast skin was higher for breast reconstruction following a skin-sparing mastectomy than for breast reconstruction following a conventional mastectomy. This contrast difference could be due to differences in pigmentation, such as that resulting from typical differences in the amount of sun exposure of breast and abdominal skin.

In summary, this study presented novel measures of the appearance of surgical scars that can be computed from clinical digital photographs. The measures quantify the proportion of the breast area affected by scarring and the intensity gradient (contrast) between the scars and the surrounding skin. We demonstrated that the measures are robust to intra- and inter-observer variability in delineating the regions of interest on the images. Finally, we presented a practical example of how the new measures of scar appearance could be used to quantitatively compare different surgeries.

Acknowledgments

The authors also would like to thank Chris Kite for technical assistance. This work was supported in part by a grant from The University of Texas at Austin Office of the Vice President for Research, grant number 1 R21 CA109040 – 01A1 from The National Cancer Institute, and grant RSGPB-09-157-01-CPPB from the American Cancer Society.

References

1. Cancer Facts & Figures 2008. American Cancer Society; Atlanta: 2008.
2. Bold RJ. Surgical management of breast cancer: today and tomorrow. *Cancer Biother Radiopharm.* 2002; 17(1):1–9. [PubMed: 11915165]
3. Hunt, KK.; Robb, GL.; Strom, EA.; Ueno, NT. *Breast Cancer.* Springer; New York: 2001.
4. 2000/2006/2007 National Plastic Surgery Statistics. American Society of Plastic Surgeons; Arlington Heights, IL: 2008.
5. Bayat A, McGrouther DA, Ferguson MWJ. Skin scarring. *BMJ (Clinical Research Ed).* 2003; 326(7380):88–92. Electronic.
6. Atiyeh BS, Ioannovich J, Al-Amm CA, El-Musa KA, Dham R. Improving scar quality: a prospective clinical study. *Aesthetic Plast Surg.* 2002; 26(6):470–6. [PubMed: 12621572]
7. Godwin Y. A comparison of the patient and surgeon opinion on the long-term aesthetic outcome of reduction mammoplasty. *British Journal of Plastic Surgery.* 1998; 51:444–9. [PubMed: 9849364]
8. Bayat A, McGrouther DA, Ferguson MW. Skin scarring. *BMJ.* 2003; 326(7380):88–92. [PubMed: 12521975]
9. van Zuijlen PPM, Angeles AP, Kreis RW, Bos KE, Middelkoop E. Scar assessment tools: implications for current research. *Plastic and Reconstructive Surgery.* 2002; 109(3):1108–22. Print. [PubMed: 11884845]
10. Kim MS, Sbalchiero JC, Reece GP, Miller MJ, Beahm EK, Markey MK. Assessment of breast aesthetics. *Plast Reconstr Surg.* 2008; 121(4):186e–94e.
11. Lowery JC, Wilkins EG, Kuzon WM, Davis JA. Evaluations of aesthetic results in breast reconstruction: an analysis of reliability. *Ann Plast Surg.* 1996; 36(6):601–6. discussion 7. [PubMed: 8792969]
12. Andrade WN, Semple JL. Patient self-assessment of the cosmetic results of breast reconstruction. *Plast Reconstr Surg.* 2006; 117(1):44–7. discussion 8-9. [PubMed: 16404246]

13. Edsander-Nord A, Brandberg Y, Wickman M. Quality of life, patients' satisfaction, and aesthetic outcome after pedicled or free TRAM flap breast surgery. *Plast Reconstr Surg.* 2001; 107(5): 1142–53. discussion 54–5. [PubMed: 11373553]
14. Veiga DF, Neto MS, Garcia EB, Filho JV, Juliano Y, Ferreira LM, Rocha JL. Evaluations of the aesthetic results and patient satisfaction with the late pedicled TRAM flap breast reconstruction. *Ann Plast Surg.* 2002; 48(5):515–20. [PubMed: 11981193]
15. Sneeuw KC, Aaronson NK, Yarnold JR, Broderick M, Regan J, Ross G, Goddard A. Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 2. Relationship with psychosocial functioning. *Radiother Oncol.* 1992; 25(3):160–6. [PubMed: 1470692]
16. Kerrigan CL, Homa K. Visual Assessment of Linear Scars: A New Tool. *Plastic and Reconstructive Surgery.* 2009; 124(5):1513–9. 10.097/PRS.0b013e3181b9898e. [PubMed: 20009837]
17. 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):514–22. [PubMed: 16079683]
18. 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):103–7. [PubMed: 15019115]
19. Beausang E, Floyd H, Dunn KW, Orton CI, Ferguson MW. A new quantitative scale for clinical scar assessment. *Plast Reconstr Surg.* 1998; 102(6):1954–61. [PubMed: 9810991]
20. van Zuijlen PP, Angeles AP, Kreis RW, Bos KE, Middelkoop E. Scar assessment tools: implications for current research. *Plast Reconstr Surg.* 2002; 109(3):1108–22. [PubMed: 11884845]
21. Yavuzer R, Smirnes S, Jackson IT. Guidelines for standard photography in plastic surgery. *Ann Plast Surg.* 2001; 46(3):293–300. [PubMed: 11293523]
22. Rasband, WS. ImageJ. U. S. National Institute of Health; Bethesda: 1997-2005.
23. Rosner, B. *Fundamentals of Biostatistics.* Duxbury Press; 2005.
24. Sullivan T, Smith J, Kermode J, McIver E, Courtemanche DJ. Rating the burn scar. *J Burn Care Rehabil.* 1990; 11(3):256–60. [PubMed: 2373734]
25. Truong PT, Abnoui F, Yong CM, Hayashi A, Runkel JA, Phillips T, Olivotto IA. Standardized assessment of breast cancer surgical scars integrating the Vancouver Scar Scale, Short-Form McGill Pain Questionnaire, and patients' perspectives. *Plast Reconstr Surg.* 2005; 116(5):1291–9. [PubMed: 16217470]
26. Sneeuw KC, Aaronson NK, Yarnold JR, Broderick M, Regan J, Ross G, Goddard A. Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 1. Comparison of patients' ratings, observers' ratings, and objective assessments. *Radiother Oncol.* 1992; 25(3):153–9. [PubMed: 1470691]
27. Godwin Y, Wood SH, O'Neill TJ. A comparison of the patient and surgeon opinion on the long-term aesthetic outcome of reduction mammoplasty. *Br J Plast Surg.* 1998; 51(6):444–9. [PubMed: 9849364]
28. Setaro M, Sparavigna A. Quantification of erythema using digital camera and computer-based colour image analysis: a multicentre study. *Skin Res Technol.* 2002; 8(2):84–8. [PubMed: 12060471]
29. Coelho SG, Miller SA, Zmudzka BZ, Beer JZ. Quantification of UV-induced erythema and pigmentation using computer-assisted digital image evaluation. *Photochem Photobiol.* 2006; 82(3):651–5. [PubMed: 16522135]
30. van Zuijlen PP, Angeles AP, Suijker MH, Kreis RW, Middelkoop E. Reliability and accuracy of techniques for surface area measurements of wounds and scars. *Int J Low Extrem Wounds.* 2004; 3(1):7–11. [PubMed: 15866782]
31. Hurtut, T.; Cheriet, F.; Joncas, J.; Dansereau, J. Enhancement and Segmentation of Scar Color Images after a Scoliosis Surgery; Digital Imaging Computing: Technique and Applications (DICTA), Sydney IAPR & IEEE Conference; Sydney. 2003; p. 233-42.

Figure 1.

Sample images illustrating the manual delineation of regions of interest for quantitative assessment of surgical scars. Top images show the areas of scar, flap region (left), surrounding skin region, and affected breast area (AR) (right) of patient who underwent flap breast reconstruction following a conventional mastectomy. Bottom images show the areas of scar (left) and affected breast area (AR) (right) of patient who underwent flap breast reconstruction following a skin-sparing mastectomy. For the bottom images, only one of the patient's two affected breasts is marked for clarity.

Table 1

Observer ratings of scars reported in the literature

Reference	Scale	Test Population	Findings
Sullivan 24	3, 4, and 5 point scale	Direct assessment of 73 burn patients	Developed Vancouver Scar Scale (VSS), where pigmentation, vascularity, pliability, and scar height are assessed independently, with increasing score indicating greater change from unaffected skin.
Van De Kar 17	10-point scale	Direct assessment of 100 patients who underwent different surgeries, including breast reduction and hand surgery	Tested the application of the Patient and Observer Scar Assessment Scale, which was developed for burns scars, to linear surgical scars. Internal consistency of observer and patient scales was good (Cronbach's $\alpha = 0.86$ and 0.90 , respectively) and reliability of the observer scale was good for the total score (Pearson $r = 0.96$, $p < 0.001$).
Truong 25	5, and six point scales	Direct assessment of 59 patients who underwent breast conserving surgery and mastectomy	Tested VSS, patient self-rating scale, and Short-Form McGill Pain Questionnaire. They obtained acceptable internal consistency for all three (Cronbach's $\alpha = 0.79, 0.64, 0.72$ respectively). Patient satisfaction was not significantly associated with VSS scores.
Viega 14	10-point scale	Direct assessment of 20 patients who underwent TRAM reconstruction after unilateral mastectomy	Compared patients' own assessments with that of two senior plastic surgeons using global rating scales as well as the five subscales of the modified Garbay system. The inter-rater and intra-rater agreement was poor to fair for the majority of subscales ($0.01 < \kappa < 0.61$).
Lowery 11	4-point, scales	Photographs of 50 patients who underwent a variety of breast reconstruction procedures	Suggested explicit criteria and to separate various components of the aesthetic result to improve the reliability of the assessment ($\kappa = 0.31$).
Andrade 12	5-point scale	214 questionnaires returned by patients who underwent a variety of breast reconstruction procedures	Observed a significant association between overall aesthetic evaluation and scar assessment.
Edsander-Nord 13	7-point, subscales including scar on the breast,	Photographs of 53 patients who underwent TRAM and pedicled flap reconstruction	Found a strong correlation between the patients' and the panel's evaluations of the aesthetic outcome (t -test, $p < 0.0001$).
Sneeuw 26	4-point scale	Photographs of 53 patients who underwent breast conservation therapy	Observed high inter-rater agreement between the nurse and the oncologist ($\kappa = 0.64$) but lower inter-rater agreements between ratings by patients and clinical observers ($\kappa < 0.10$).
Godwin 27	5-point scale	34 photographs of patients who underwent	Found that scarring was the most frequent cause of

Reference	Scale	Test Population	Findings
		breast reduction	dissatisfaction for both surgeons (23%) and patients (25%) ($p = 0.07$).

Table 2

Quantitative assessments and measures of scars reported in the literature

Reference	Test Population and Subject	Findings
Draaijers 18	49 burn scar areas in 20 patients	Found that measurements with the DermaSpectrometer and Minolta Chromameter (Pearson $r = 0.72$) were reliable. At least three observers necessary with for reliable pigmentation measurement (Pearson $r > \text{or} = 0.77$). Agreement for observers for pigmentation was unacceptable ($kappa = 0.349$).
Setaro 28	348 photographs of Caucasian adults with erythema	Found that digital measurement of absent, slight, and moderate erythema using normalized r, g, b color coordinate system was clinically useful.
Coelho 29	Photographs of 9 erythema patients of White, Asian, or Black or African-American ethnicity	Demonstrated the feasibility of using digital photography for objective evaluation of UV erythema and pigment changes due to UV exposure in different racial/ethnic groups even with light to heavy UV exposure schedules (Pearson $r = 0.66, 0.96, 0.93$).
van Zuijlen 30	20 healthy volunteers	Demonstrated that planimetry by photography is more reliable than planimetry by tracings for areas that are not extremely curved ($0.72 < ICC < 0.87$).
Hurtut 31	Undefined number of photographs of scoliosis patients	Demonstrated encouraging results with an average deviation using automatic contours compared to manual contours of $0.67 \pm 0.24 \text{ mm}$ with a camera resolution of 1.8 mm . This is less than the deviation between two manual contours ($0.75 \pm 0.21 \text{ mm}$).

Table 3

Intraclass correlation (ICC) coefficient for SIG, FIG, and AR ratio ratios between the first and second set of measurements made by each observer for the cases of flap breast reconstruction following conventional mastectomy. The intra-observer agreement was excellent for most measures.

Number of cases (N=16)	SIG	FIG	AR
Observer 1 (MSK)	ICC = 0.68 fair to good	ICC = 0.94 excellent	ICC = 0.72 fair to good
Observer 2 (WNR)	ICC = 0.84 excellent	ICC = 0.95 excellent	ICC = 0.83 excellent

Table 4

Intraclass correlation (ICC) coefficient for SIG and AR ratio between the between the first and second set of measurements made by each observer for the cases of flap breast reconstruction flap following skin sparing mastectomy. The intra-observer agreement was better for SIG than for AR.

Number of cases (N = 24)	AR	SIG
Observer 1 (MSK)	ICC = 0.18 poor	ICC = 0.61 fair to good
Observer 2 (WNR)	ICC = 0.63 good	ICC = 0.83 excellent