Objective assessment of aesthetic outcomes of breast cancer treatment: Measuring ptosis from clinical photographs

Min Soon Kim\textsuperscript{a}, Gregory P. Reece\textsuperscript{b}, Elisabeth K. Beahm\textsuperscript{b}, Michael J. Miller\textsuperscript{b}, E. Neely Atkinson\textsuperscript{c}, Mia K. Markey\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a}Biomedical Engineering Department, The University of Texas at Austin, Austin, TX 78712, USA
\textsuperscript{b}Department of Plastic Surgery, The University of Texas M.D. Anderson Cancer Center, USA
\textsuperscript{c}Department of Biostatistics and Applied Mathematics, The University of Texas M.D. Anderson Cancer Center, USA

Received 15 October 2004; received in revised form 17 August 2005; accepted 10 October 2005

Abstract

The aesthetic outcome of breast cancer treatment is an important factor in breast cancer survivors’ quality of life. We investigated new quantitative, objective measurements of breast ptosis based on ratios of distances between fiducial points manually identified in oblique and lateral clinical photographs. Pto\-sis refers to the extent to which the nipple is lower than the inframammary fold. The new objective measures were compared to ratings made using an existing subjective scale. The variability in the objective measurements due to intra- and inter-observer variability in marking fiducial points was shown to be equivalent to less than one point on the subjective ptosis scale.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Breast neoplasms; Diagnosis, Computer-assisted; Reconstructive surgical procedures; Photogrammetry; Photography; Ptosis

1. Introduction

Breast cancer is the most common type of cancer among women in the US, excluding skin cancer. It is predicted that in 2005, there will be 212,118 new cases diagnosed [1]. The breast cancer survival rate has increased because of development of improved detection techniques and treatment methods [2,3]. Breast reconstructive surgery is a crucial part of breast cancer treatment, which can help breast cancer survivors regain a high quality of life. The American Society of Plastic Surgeons estimates that 74,000 breast reconstructive surgeries were performed in the US in 2002 [4]. Breast cancer survivors are now assured the benefit of breast reconstructive surgeries as well as breast conservation therapies according to the Women’s Health and Cancer Rights Act of 1998 (P.L. 105-277). Restored function and aesthetics are the goals of any reconstructive surgery. However, current techniques of breast reconstructive surgery cannot restore lactation and normal sensation. Therefore, the main goal of breast reconstructive surgery is restoration of aesthetics to assist breast cancer survivors in regaining a high quality of life. Aesthetics refers to physical characteristics of the breasts, such as shape, symmetry, and ptosis.

Adequate measures of surgical outcome are important to protect patients’ health and well-being. Healthcare providers, the insurance industry, and the government also need objective measures to set up reasonable guidelines for care. The lack of a generally accepted quantitative method for assessing breast aesthetics limits the effective assessment of the outcomes after breast reconstructive surgery.

Currently, physicians, patients, or other observers report subjective assessments of breast aesthetics based on printed photographs, digital images, or directly viewing patients [5–14]. However, these methods have been shown to have low intra- and inter-observer agreement. Quantitative, objective measures with high reliability are needed in order to investigate the relationships between patient and surgical variables and aesthetic outcomes.

The previous studies of photogrammetry to quantify breast aesthetics provide a strong foundation for our work...
However, there are certain limitations that must be overcome through additional investigations. First, prior studies have only used only AP views. Important aesthetic properties, such as ptosis described below, may not be adequately visible in the AP view. Second, the objective photogrammetry measures proposed in the literature are focused on symmetry characteristics to the point that some other properties such as ptosis are not addressed. Third, since prior photogrammetry studies were concerned with breast conservation therapy, assumptions were made that are not met in reconstructive surgery. In particular, measures were calculated assuming that only one breast was surgically altered, whereas even in unilateral reconstruction procedures it is often necessary to modify the unaffected breast to achieve symmetry.

An important aesthetic property that has not been addressed by prior photogrammetry studies is ptosis. Ptosis of the breast is classified according to the relationship between the nipple and the inframammary fold, the crease beneath the breast. Unacceptable ptosis occurs when nipple and lower pole of the breast descend lower than the level of inframammary fold [15]. Regnault [16] defined four grades of ptosis. A patient has no ptosis when the nipple and most of the breast gland is located above the level of submammary fold (Grade 0). In first degree or minor ptosis, the nipple is at the level of submammary fold and above the lower contour of the breast (Grade 1). Second degree or moderate ptosis is when the nipple lies below the fold but above the lower contour (Grade 2). Third degree or major ptosis is when the nipple lies at the lower breast contour and most of the breast is below the fold (Grade 3). The ptosis scale with example photographs is described in Table 1 and Fig. 1.

<table>
<thead>
<tr>
<th>Ptosis grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = None</td>
<td>Nipple and most of gland are above IMF</td>
</tr>
<tr>
<td>1 = Minor</td>
<td>Nipple at IMF</td>
</tr>
<tr>
<td>2 = Moderate</td>
<td>Nipple is below IMF but above lower contour of breast</td>
</tr>
<tr>
<td>3 = Major</td>
<td>Nipple at lower breast contour &amp; breast below IMF</td>
</tr>
</tbody>
</table>

IMF: inframammary fold.

In this study, we designed quantitative, objective measurements of breast ptosis based on ratios of distances between fiducial points manually identified in lateral and oblique views of clinical photographs. The new objective measures were compared to ratings on a subjective scale made by three experienced clinical observers. Intra- and inter-observer variability of the new objective measures and subjective ratings were investigated.

2. Materials and methods

2.1. Datasets

The patient population for this study was women aged 21 years or older who underwent breast reconstruction surgery from January 1, 1990 to June 1, 2003. Anterior–posterior (AP), right and left lateral, and right and left oblique photographs were digitally taken with a Nikon 990 Coolpix or Cannon T90 35 mm SLR with 50 mm lens and digitized with a Nikon LS 2000 or Nikon Super Coolscan 4000ED (1.06) slide digitizer. An experienced plastic surgeon (GPR) selected pre-operative images for 52 patients. The dataset demonstrates a wide range of breast aesthetic characteristics including size, symmetry, ptosis, and projection.

2.2. Ptosis: subjective rating

The ptosis of each breast of all 52 patients was subjectively rated by three clinical observers (GPR, MJM, EKB) independently using a 4-point scale (Table 1, Fig. 1) described by Bostwick [15]. All five views of a patient were displayed simultaneously. The subjective assessment was repeated at three time points at least 2 weeks apart. Thus, it is unlikely that an observer recalled his/her previous rating. Kappa statistics of subjective scale measures between time points (first vs. second, second vs. third, and first vs. third) were analyzed to investigate intra-observer variability. Kappa statistics of the subjective scale measures between clinical observers (GPR vs. MJM, GPR vs. EKB, and MJM vs. EKB) were analyzed to investigate inter-observer variability.

![Fig. 1. Sample images (right lateral view) illustrating the ptosis grades described by Regnault [16] and Bostwick [15]. n: nipple point, i: lateral terminus of inframammary fold.](image-url)
2.3. Fiducial points

Out of the 52 patients whose names were unknown to any observer, a clinical observer (GPR) selected 10 patients that illustrate a wide range of aesthetic characteristics (size, ptosis, and projection) and ethnicity. Eight observers (three clinical (GPR, MJM, EKB) and five novices) marked fiducial points in images of the 10 patients on three occasions with 5 min breaks between assessments. The novice observer group was trained to mark fiducial points by a clinical expert (GPR) before the study was performed using a separate set of photographs. There were no additional guides or hints provided for marking fiducial points during the study. The four fiducial points marked were the sternal notch, lateral extent of the inframammary fold, the lowest visible point, and the nipple centroid. The sternal notch (s) or jugular notch point is the shallow indentation on the superior surface of the manubrium. It is located between two clavicular articulations [15,17]. The inframammary fold is a curvilinear structure which is generally hidden behind breast tissue since most women have some degree of ptosis [15]. The lateral terminus of the inframammary fold is the endpoint of the inframammary fold (i) where it intersects at the chest wall and it is typically visible in the lateral and oblique photographs. Therefore, the lateral terminus of the inframammary fold was used as the reference point. The lowest visible point (v) of the breast is located at the most inferior point of the breast. The centroid of nipple (n) is considered rather than the nipple-areola complex since many women have an irregular shaped areolae. The sternal notch (s), lateral terminus (i) and the nipple (n) are marked in the oblique views of the right and left breasts separately. The nipple (n), the lowest visible point (v), and the lateral terminus (i) are marked in the lateral views of the right and left breasts separately. We designed a program using MATLAB® (The MathWorks, Natick, MA) to automatically load the images and prompt the observer to manually identify locations of the fiducial points. The x- and y-coordinates and the time required per each fiducial point were recorded. We studied x- and y-localizations for the fiducial points for both clinical and non-clinical observers. Minimum and maximum values of variability of the fiducial points for both groups were calculated.

2.4. Ptosis: objective measures

We explored objective, quantitative measures of aesthetic properties computed from relative distances between fiducial points. It is not possible to compare absolute distances since the magnification of the images was not standardized over the data set images. The pixel coordinate system was used since it preserves the unique coordinates for each fiducial point when the images are resized. We designed two measures relating to ptosis using the oblique and lateral clinical photographs.

Measure (1) is computed from the oblique views of the right and left breasts separately (using the proximal breast, which is nearest to observer’s viewpoint). The vertical level of the sternal notch (s) is taken to be zero and the vertical displacements of the lateral terminus (i) and the nipple (n) are calculated from it. If the nipple is higher than the lateral terminus, then by definition there is no ptosis and the value is set to one; otherwise, measure (1) is calculated. Values near one indicate little ptosis while values near zero indicate significant ptosis (Fig. 2).

\[ s - i \]
\[ s - n \]

Measure (2) is calculated from the lateral views of the right and left breasts separately. The vertical level of the lowest visible point of the breast (v) is taken to be zero, and the vertical displacements of the nipple (n) and the lateral terminus (i) are calculated from it. If the nipple is higher than the lateral terminus, then by definition there is no ptosis and the value is set to one; otherwise, measure (2) is calculated. Values near one indicate little ptosis and values near zero indicate significant ptosis (Fig. 3).

\[ n - v \]
\[ i - v \]

Our objective measures for ptosis as defined above are expected to have values in the range from 0 to 1, but the subjective scale ranges from 0 to 3. Thus, we rescaled our objective measure values to be more similar to the subjective scales to simplify interpretation of our objective measures. For the right and left breast separately of 10 cases, the subjective ratings and objective measures were averaged across the three clinical observers (GPR, MJM, EKB) across the three time points. A simple linear regression was then used to relate the objective measures to the subjective ratings. The regression model was used to transform the objective measures of the novice group for same 10 patient cases to evaluate the impact of the variability in manual identification of the fiducial points.
3. Results

3.1. Ptosis: subjective rating

Kappa statistics of subjective scale measures between time points (first vs. second, second vs. third, and first vs. third) and between the three clinical observers (GPR vs. MJM, GPR vs. EKB, and MJM vs. EKB) were studied. Kappa results showed good to excellent intra-observer agreement (0.52–0.84).

However, the inter-observer agreement varied across observers (0.23–0.49) and was lower in the case of MJM vs. EKB (0.07–0.15). The average values of the subjective ratings across the three clinical observers for 10 cases were used to calibrate our objective, quantitative measures (Tables 2 and 3).

3.2. Fiducial point localization

We studied the variability of the three clinical observers over time, variability between the clinical observer group and novice group, and the time required for each fiducial points between the clinical observer group and novice group.

Fig. 4 illustrates examples of the variability of a clinical observer (GPR) over three time points and Fig. 5 illustrates the variability of three clinical observers (GPR, MJM, EKB) for one time point (Tables 4 and 5). The clinical observer group was especially consistent in marking the locations of the nipples relative to novice group. In the oblique view, the maximum standard deviation across time points among all patients, both breasts was 1.64 pixels in x and 0.71 pixels in y for the nipples. In the lateral view, the maximum standard deviation across time points among all patients, both breasts was 1.27 pixels in x and 1.41 pixels in y for the nipples. The variability across time points among all patients, both breasts was likewise very small for the y coordinate of the lowest visible point in the lateral view (1.15 pixels). The variability was moderately small (standard deviation ≈ 1.5 pixels) for most of the other fiducial points, the exception being the y coordinate of the lateral terminus of the inframammary fold (9.12 pixels). Over all, the clinical observers were extremely consistent over time in marking all of the fiducial points used in this study.

Figs. 6 and 7 illustrate examples of the variability between a clinical observer (GPR) and non-clinical observers. As is shown

![Fig. 3. Measure #2 in right lateral view. The level of lowest visible point (v) is extended with dotted line. The distances between nipple (n) to the lowest visible point (v) and between the point of the lateral terminus of inframammary fold (i) to the lowest visible point (v) are indicated with arrowed lines.](image)

Table 2
Kappa statistics of subjective ratings over the time by three clinical experts

<table>
<thead>
<tr>
<th></th>
<th>First vs. second</th>
<th>First vs. third</th>
<th>Second vs. third</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GPR</td>
<td>MJM</td>
<td>EKB</td>
</tr>
<tr>
<td>Right breast</td>
<td>0.64</td>
<td>0.59</td>
<td>0.63</td>
</tr>
<tr>
<td>Left breast</td>
<td>0.72</td>
<td>0.71</td>
<td>0.59</td>
</tr>
</tbody>
</table>

All 52 patients images were rated for ptosis by three clinical observers (GPR, MJM, and EKB) using a 4-point scale (Table 1) based on Regnault [16] and Bostwick [15] at three points, approximately 2 weeks apart. The amount of time between ratings was chosen to minimize the likelihood that the clinical observer would recall his previous ratings. First vs. second, first vs. third, and second vs. third ratings are compared.

Table 3
Kappa statistics of subjective ratings across the three clinical experts

<table>
<thead>
<tr>
<th></th>
<th>First</th>
<th>Second</th>
<th>Third</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GPR vs. MJM</td>
<td>GPR vs. EKB</td>
<td>MJM vs. EKB</td>
</tr>
<tr>
<td>Right breast</td>
<td>0.39</td>
<td>0.37</td>
<td>0.08</td>
</tr>
<tr>
<td>Left breast</td>
<td>0.26</td>
<td>0.37</td>
<td>0.08</td>
</tr>
</tbody>
</table>

All 52 patients images were rated for ptosis by three clinical observers (GPR, MJM, and EKB) using a 4-point scale (Table 1) based on Regnault [16] and Bostwick [15] at three points, approximately 2 weeks apart. The amount of time between ratings was chosen to minimize the likelihood that the clinical observer would recall his previous ratings. GPR vs. MJM, GPR vs. EKB, and MJM vs. EKB ratings are compared.
Fig. 4. Variability of fiducial point localization by expert over time (+: first trial, ●: second trial, *: third trial). Expert observer (GPR) marked the fiducial points (left top: sternal notch, right top: left nipple, left bottom: the lowest visible point, right bottom: lateral terminus of inframammary fold). The sternal notch and left nipple are marked with white markers and the lowest visible point and the lateral terminus of the inframammary fold are marked with black markers Four left oblique views of one patient were generated for example photographs.

Fig. 5. Variability of fiducial point localization by three experts at first trial (+: GPR, ●: MIM, *: EKB). Three expert observers (GPR, MIM, EKB) marked the fiducial points (left top: sternal notch, right top: left nipple, left bottom: the lowest visible point, right bottom: lateral terminus of inframammary fold). The sternal notch and left nipple are marked with white markers and the lowest visible point and the lateral terminus of the inframammary fold are marked with black markers Four left oblique views of one patient were generated for example photographs.
Eight observers (three clinical and five novices) marked the fiducial points using our new objective tools. Standard deviations of the fiducial points from Measure #1 (right and left oblique views) are tabulated. Minimum and maximum values of standard deviations from x- and y-coordinate are calculated across the three time points for each patient and each breast.

Eight observers (three clinical and five novices) marked the fiducial points using our new objective tools. Standard deviations of every fiducial points from Measure #2 (right and left lateral views) are tabulated. Minimum and maximum values of standard deviations from x- and y-coordinate are calculated across the three time points for each patient and each breast.

Fig. 6. Variability of fiducial point localization comparing an expert and novice groups for Measure #1. Observers marked the fiducial points (left: sternal notch, middle: left nipple, right: lateral terminus of inframammary fold). Fiducial points marked by six observers are indicated with markers; +: expert observer and five novice observers: ○, ×, □, ○, △. The sternal notch and left nipple are marked in white and the lateral terminus of the inframammary fold is marked in black. Three left oblique views of one patient were generated for example photographs.

Fig. 7. Variability of fiducial point localization comparing an expert and novice groups for Measure #2. Observers marked the fiducial points (left: left nipple, middle: lowest visible point, right: lateral terminus of inframammary fold). Fiducial points marked by six observers are indicated with markers; +: expert observer and five novice observers: ○, ×, □, ○, △. The left nipple is marked in white and the lowest visible point and the lateral terminus of the inframammary fold is marked in black. Three left oblique views of one patient were generated for example photographs.
Table 6

Time required for marking fiducial points between clinical and novice groups for 10 patients

<table>
<thead>
<tr>
<th></th>
<th>Clinical group first/second/third/overall</th>
<th>Novice group first/second/third/overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nipples (n)</td>
<td>4.56/3.84/3.56/3.99</td>
<td>3.75/3.11/2.95/3.27</td>
</tr>
<tr>
<td>Lateral terminus of inframammary fold (i)</td>
<td>4.27/3.43/3.01/3.51</td>
<td>5.35/3.95/3.35/4.22</td>
</tr>
<tr>
<td>Lowest visible point (v)</td>
<td>3.73/3.45/3.35/3.51</td>
<td>4.20/3.85/3.15/3.74</td>
</tr>
<tr>
<td>Sternal notch (s)</td>
<td>5.34/4.12/4.33/4.59</td>
<td>5.25/3.99/3.65/4.30</td>
</tr>
<tr>
<td>Time required for all fiducial points in Measures #1 and #2 per patient (four nipple points, four lateral terminus of Inframammary folds, two lowest visible points, two sternal notches)</td>
<td>46.44</td>
<td>46.01</td>
</tr>
</tbody>
</table>

Eight observers (three clinical and five novices) marked the fiducial points using our new objective tools. Mean values of time for each fiducial point and each trial from Measures #1 and #2 are computed. (unit: seconds).

![Graph](image)

Fig. 8. Simple linear regression between ptosis ratings made by the clinical observer and Measure #1. A simple linear regression was used to relate average values computed for the objective measures to the subjective ratings for 10 patient cases provided by three clinical observers (GPR, MJM, EKB). The regression data was used to transform the objective measure data of novices group for same patient cases to evaluate the impact of the variability from manual identification of the fiducial points.

in these examples, overall the novice observers in this study were able to accurately mark the locations on the fiducial points if the clinical observer’s marks are taken as the ground truth. Tables 4 and 5 report the maximum standard deviation across time for each fiducial point among all novice observers viewing all patients, both breasts. As was the case for the clinical observer group, the novice group shows the least variability of all the fiducial points for localizing the nipples in either view (standard deviations ≈ 1 pixels). In fact, the variability in localizing most of the fiducial points was similar for the clinical and novice observers, with two exceptions. In the oblique view, the novice observers showed considerably more variability for the x- and y-coordinates of the lateral terminus of the inframammary fold and the y coordinate of the sternal notch. These structures are among the more subtle fiducial points used in this study and these results indicate that additional training may be desirable for novice observers or that automated methods should be developed.

The time required for marking fiducial points by both the clinical and novice groups for 10 patients was studied (Table 6). The mean time required to locate each fiducial point in each trial was computed. The mean times required for marking all fiducial points for Measures #1 and #2 were 46.44 and 46.01 s for the clinical and novice groups, respectively. This result indicates that non-surgeons can identify the fiducial points needed for our objective measures as quickly as surgeons can. It was also observed that the time needed to locate the fiducial points typically decreased with each trial for both the clinical and novice groups, presumably as a result of the users becoming more accustomed to the software interface.

3.3. Ptosis: objective measures

Our objective measures for ptosis are expected to have values in the range from 0 to 1, but the subjective scale ranges from 0 to 3. We rescaled the values of our new measures to be more
similar to the subjective scales to simplify interpretation of our objective measures. The mean of the three clinical observers’ subjective ratings at three time points were calculated over right and left breasts separately for 10 patient cases. Simple linear regression was used to relate the ptosis ratings obtained using the subjective scale to our objective measures (Figs. 8 and 9). The simple linear regression model from clinical group was used to transform the objective measures from novice group’s distance ratios to a scale more similar to that of the subjective scale to allow for easier interpretation of the objective measures. After rescaling the objective measures, we determined that the level of variability of novice group in the objective ptosis measure resulted from variability in the fiducial point marking. Maximum variability from different observers for the same case at different time point was about 0.47 ptosis grade on Measure #1 (Table 7) and 0.42 ptosis grade on Measure #2 (Table 8).

Two additional simple linear regression analyses were performed for comparison, which we will refer to as models B and C. Regression model B was built on all 52 cases using the subjective ratings from clinical group vs. the objective measurements based on the fiducial point markings of a single novice observer (MKS). The same modeling strategy was used for model C, with the difference that the 10 patient cases which were used for testing the rescaling of the novice group measurements were excluded. The results indicated that the maximum variability from different novice observers from model B were 0.50 ptosis grade on Measure #1 and 0.46 grade on Measure #2 and for model C were 0.48 ptosis grade on Measure #1 and 1.25 ptosis grade on Measure #2. Thus, the results of using model B, which used a larger number of cases but a smaller number of observers than the main model, were essentially equivalent to that of the main model. The results of using model C, in which the 10 cases being assessed were held out when the model was built, showed an increased variability as was expected.

4. Discussion

Breast reconstruction is a critical component of breast cancer treatment for patients who undergo mastectomy. Accurate
assessment of the aesthetic outcome is important to guide health care providers and patients in evaluating treatment options.

Currently, physicians, patients, or other observers report subjective assessments of breast aesthetics based on printed photographs, digital images, or by directly viewing patients [5–14]. One disadvantage of subjective assessments is that they have low intra- and inter-observer agreement [5,7–9,18–22]. The use of subjective, qualitative scales possessing only four and five gradations limits the analyses that can be performed. Quantitative, objective measures with high reliability are required to investigate the relationships between patient and surgical variables and aesthetic outcomes.

One approach to quantifying breast aesthetics is anthropometry, in which a tape measure is used to make “linear measurements” between fiducial points (anatomical landmarks such as the sternal notch) [23–29]. However, this method has disadvantages in that linear measurements are not routinely collected and the relationships to subjective scales is unclear [26,28–31]. Linear measurements on a photograph of the patient instead of on the patient directly have shown potential for quantifying breast aesthetics. Photogrammetry has three advantages over linear measurements on the patient directly. First, taking a photograph is not as invasive. Second, it is possible to retrospectively make several measurements on the photographs. Third, photographs are generally available for analysis since they are routinely collected for documentation purposes. Both photogrammetry and direct anthropometry have better inter-observer agreement [5] and allow for more quantification than subjective assessment by human observers.

Clinical photographs of patients undergoing breast reconstruction are usually taken from five different views: anterior–posterior, right and left lateral, and right and left oblique views. An anterior–posterior (AP) view includes clavicles and shoulders above, pubis below, arms at side. An oblique view is taken with the patient turned 45° and with the distal arm back slightly. Lateral views are taken directly from the side such that only the proximal breast is typically visible. Efforts have been made to standardized clinical photography, but variability remains a challenge [32–35].

Three research groups have proposed making measurements on photographs, either manually or aided by a computer, to quantify breast aesthetics outcomes of breast-conservation therapy [18,20,36–39]. Sacchini et al. [18] used a computer program to make four measurements on digital images of the AP view of 148 patients who underwent breast-conservation therapy. They measured the difference in height between the two nipples, the difference in height between the inferior poles, the difference of the distance from the midline to the nipples, and the difference of the distance from the sternal notch to the nipples. In a separate study, Sacchini et al. [36] computed a nipple offset distance using digital photographs of 101 patients who underwent breast-conservation therapy. In both studies, the authors concluded that subjective ratings (4-point scale) showed the same trends as their measurements, but they did not report a statistical analysis.

Pezner et al. [37] defined breast retraction assessment (BRA) as a nipple offset distance calculated with the aid of a clear acrylic sheet marked as a grid. This is functionally equivalent to using a photograph. They reported baseline values of BRA for 29 normal volunteers and compared these to the values for 27 patients who underwent breast-conservation therapy. The mean BRA value was significantly larger for the breast cancer patients than for the normal volunteers. Further, patients who underwent more extensive resection of the tumor had significantly larger BRA values than those who underwent more limited resection.

Van Limbergen et al. [38,39] defined four measurements to be made on AP photographs: the vertical displacement of the nipple from the level of the sternal notch (A), the vertical displacement of the lower pole of the breast from the level of the sternal notch (I), the horizontal displacement of the nipple from the midline (M), and the horizontal displacement of the lateral breast contour from the midline (L). The differences in each measure between the left and right breasts provided measures of symmetry (e.g., AI). The authors also calculated the BRA measure defined by Pezner et al. [37], which is closely related to their measures. In Van Limbergen et al.’s notation: 

\[
BRA = \sqrt{(\Delta M)^2 + (\Delta A)^2}
\]

The measures were calculated from the photographs of 142 patients who underwent breast-conservation therapy. A panel subjectively rated the aesthetics of each photograph on a five-point scale, and the mean panel score was taken as the overall subjective rating for the patient. The authors reported significant correlations between each of the measures and the mean panel score. The difference in the vertical displacement of the lower pole of the breasts from the level of the sternal notch, \(\Delta I\), was most strongly correlated with the subjective assessment. More recently, the same group evaluated the BRA measurement in a large study of breast-conservation therapy [20]. Recognizing the difficulties in standardizing the magnification of photographs, they introduced a relative version of BRA, which they referred to as pBRA, that is equal to BRA divided by the distance from the sternal notch to the nipple of the unaffected breast. The pBRA measure was compared to the mean panel assessment of two male surgeons, two male radiation oncologists, and one female data manager, who rated 647 photographs on a four-point scale (based on [40]). There was a significant high correlation between the mean panel score and the pBRA measurement. The authors reported mixed results regarding the correlation of pBRA with subjective assessments for inferomedially located tumors [20,39].

In this study, we designed quantitative, objective measurements of breast ptosis based on ratios of distances between fiducial points manually identified in lateral and oblique views of clinical photographs. Three clinical observers rated preoperative clinical photographs of 52 patients who underwent breast reconstruction for ptosis using a qualitative, subjective scale at three time points (at least 2 weeks). Kappa statistics between time points indicated good to excellent intra-observer agreement (0.52–0.84), which encouraged our use of the average of the three observers’ ratings as a guide in assessing our new quantitative, objective measurements of breast ptosis.

In our current implementation, fiducial points are manually identified. While still not ideal, it would be more practical to use manually identified fiducial points if the task could be accurately performed by a non-clinical observer (e.g., a research
assistant). In this study we found that novice observers could reliably locate the required fiducial points as compared to the identifications made by clinical observers. Moreover, both novices and clinical observers demonstrated low variability in repeated marking of fiducial points at different time points. Thus, it is practical to use our approach with novice observers. However, a more automated method would be better for large outcome studies or for developing intra-operative tools. It takes about 40s per patient to mark the 12 fiducial points required for the two ptosis measures presented here, though additional fiducial points would likely be needed to address other aesthetic properties (e.g., symmetry).

Our quantitative, objective measurements of breast ptosis are based on ratios of distances between fiducial points. Typically, only the y coordinate, but not both are needed for any particular fiducial point. For measure #2, it was encouraging that for the lateral terminus of the inframammary fold the coordinate used (y) was less variable than the one that was not needed from the lateral view (x). A similar, but weaker trend was seen for the inframammary fold in the oblique view used for measure #1. In future studies, the effect of the patient’s body-mass index should be examined since the small sample used in this study suggests that some landmarks such as the sternal notch are more difficult to locate on heavier patients.

A simple linear regression model was developed to relate the subjective ratings and objective measurements for 10 patient cases by the clinical observers. The regression model was used to rescale the objective measurements of the novice group to allow for easier interpretation of the variability in the measures due to variability in fiducial point identification. The objective measures from novice group showed that the maximum variability from different novice observers for the same cases at different time point was about 0.47 ptosis grade on Measure #1 (Table 7) and 0.42 ptosis grade on Measure #2 (Table 8).

Over all there was good agreement between the novice and clinical observers, but some differences were seen. While both clinical and novice observers showed the same trend of some fiducial points being easier to reliably locate than others (e.g., lateral terminus of inframammary fold vs. nipples), the differences between the easier and more difficult points were larger for the novice observers. These observations suggest that additional training may be desirable for novice observers, particularly for more subtle structures, such as the sternal notch.

One potential limitation of objective measurements of breast ptosis based on ratios of distances between fiducial points is that they could be difficult to interpret. We addressed this problem by building a linear regression model that relates the objective measurements to subjective ratings. The linear regression model was then used to rescale each ratio to provide a final quantitative, objective measure that can be interpreted in the same manner as the more familiar subjective rating scale for ptosis. The level of variability of the objective measures of novice observers after rescaling was found to be equivalent to less than half of a point on the subjective ptosis scale. This result indicates our new objective measures show a high level of correspondence with subjective ratings of clinical observers.

5. Summary

Our long-term goal is to develop decision aids that will improve breast cancer treatment. Quantitative, objective measures with high reliability are needed to meaningfully relate patient and surgical variables to aesthetic outcomes and to compare the outcomes of different breast cancer treatment strategies. Our approach to quantifying breast aesthetics is to measure distances between anatomical landmarks (fiducial points) on standardized clinical photographs. We found that an existing subjective scale for rating ptosis showed high intra-observer agreement, but the inter-observer agreement was lower. New objective measures of ptosis showed encouraging levels of concordance with ratings made by experienced surgeons using the subjective scale. The objective measures were found to be robust to intra- and inter-observer variability in marking fiducial points, including “novice” observers.

Acknowledgments

The authors thank five lab members who participated in this study as the non-clinical observers. The authors also thank Zack Mahdavi and Chris Kite, the system administrators in Biomedical Informatics Laboratory, for technical assistance. This work was funded in part by a grant from the Office of the Vice President for Research at The University of Texas at Austin and by Grant 1R21 CA109040-01A1 from the National Institutes of Health.

References


Min Soon Kim is a Graduate Research Assistant in Biomedical Engineering at The University of Texas at Austin. He received B.S. degree in Mechanical Engineering at Chungnam National University in Korea and M.S. degree in Biomedical Engineering at UT Austin and is pursuing his Ph.D. in Biomedical Engineering. He is currently working in the Biomedical Informatics Lab and his research focuses in quantifying aesthetic outcomes after breast reconstruction surgery. Long-term goal of the research is to develop decision aids that will improve breast cancer treatment.

Gregory P. Reece, is a Professor of Plastic Surgery at the University of Texas M.D. Anderson Cancer Center in Houston, Texas. He obtained a B.S. degree in Biology at McNeese State University and a Medical Degree at Louisiana State University in Shreveport, Louisiana. He is a board certified general surgeon and plastic surgeon who practices at the University of Texas M.D. Anderson Cancer Center in Houston, Texas. The mission of the Department of Plastic Surgery at M.D. Anderson Cancer Center is to provide state of the art reconstructive services to their patients and to develop new advances in reconstructive surgery and regenerative medicine.

Elisabeth K. Beahm completed her Plastic Surgery training at the University of Chicago and undertook subsequent fellowships in microsurgery and aesthetic and facial reconstructive surgery. Dr. Beahm is currently an Associate Professor of Surgery in the Department of Plastic Surgery at the University of Texas M.D. Anderson Cancer Center, where she serves as Director of Resident Education. Dr. Beahm has developed a clinical focus devoted to the refinement of surgical strategies for breast and nasal reconstruction and she maintains a longitudinal interest in tissue engineering research.

Michael J. Miller received his Medical Degree from the University of Massachusetts, completed his plastic surgery training at Ohio State University and a fellowship in microsurgery at Tulane University Medical Center in New Orleans. A professor of plastic surgery, Dr. Miller serves as Deputy Chairman of the Department of Plastic Surgery at the University of Texas M.D. Anderson Cancer Center. He also holds adjunct appointments at Biomedical Engineering at Rice University and University of Texas Health Science Center. His research interests include tissue modeling, tissue replacement and surgical reconstruction for physical deformities related to cancer.

Edward N. Atkinson is a Professor of Statistics at the University of Texas M.D. Anderson Cancer Center. He received his Ph.D. in Mathematical Sciences from Rice University. He is active in the design and evaluation of clinical trials in cancer research and in optical methods for the early detection of cancer.

Mia K. Markey is currently an Assistant Professor in Biomedical Engineering at The University of Texas at Austin. She earned her B.S. in Computational Biology from Carnegie Mellon and her Ph.D. in Biomedical Engineering, as well a Certificate in Bioinformatics, from Duke University. The mission of the Biomedical Informatics Lab (www.bme.utexas.edu/research/informatics) is to design cost-effective, computational medical decision aids that will help physicians better diagnose, treat, and manage cancer.