Methods for the Aesthetic Evaluation of Breast Cancer Conservation Treatment: A Technological Review

Hélder P. Oliveira1*, Jaime S. Cardoso1, André Magalhães2 and Maria J. Cardoso3

1INESC TEC and Faculdade de Engenharia, Universidade do Porto, Portugal
2Faculdade de Medicina, Universidade do Porto, Portugal
3Breast Unit, Champalimaud Cancer Center, Champalimaud Foundation, Portugal

Abstract: Breast-conserving approaches aim to attain better aesthetic results in addition to local control and achieving survival rates equivalent to mastectomy in patients with breast cancer. While the oncologic outcome of breast conservation procedures is easily estimated objectively by disease-free and overall survival rates, the cosmetic outcome has no standard of evaluation. Although breast conservation techniques have been widely studied, different forms of evaluation and heterogeneous working practices have contributed to different aesthetic results. As this scenario suggests, the evaluation of aesthetic results should be mandatory in any institution performing breast cancer treatment, contributing to the improvement of current strategies by enabling the identification of variables which have a significant impact on the final aesthetic result.

In the process of assessing cosmetic outcomes there are several important issues that should be considered: which factors have a crucial impact on the cosmetic outcome of Breast Cancer Conservation Treatment (BCCT); which parameters or features should be evaluated in the cosmetic assessment of BCCT; how patients are evaluated; which scales are used in this evaluation; which methods and technological solutions are available for the evaluation of cosmetic results of BCCT.

In this paper we try to discuss all these questions, with an emphasis on the objective methods and corresponding technologies used in the aesthetic evaluation of BCCT. The most relevant publications related to the mentioned topics are presented, critically analysed and put in chronological perspective. Current and future trends are also discussed.

Key Words: Assessment of aesthetic results, Breast cancer conservative treatment, Objective methodsm Subjective methods, Three-Dimensional Methods, Two-dimensional Patient Data.

I. INTRODUCTION

Breast cancer is considered a public health problem, and is the most frequently diagnosed cancer and the leading cause of cancer death in women worldwide. One in ten women will develop breast cancer at some point in their lifetime. According to recent worldwide data, this disease accounted for 23% (1.38 million) of the total new cancer cases and for 14% (458 400) of the total cancer deaths in 2008 [1]. The incidence of breast cancer has increased in most countries worldwide in the last decades, due to changes in reproductive behaviour and the use of exogenous hormones, as well as differences in weight, exercise, diet and alcohol consumption [2].

Approximately 90% of breast cancers are curable if detected in their initial phase and treated properly. It is a very frequent disease and remains one of the most publicised malignancies not only because of its high incidence and prevalence, but also because of the impact that the aesthetic appearance of the breast has on women. Routine screening promotes early diagnosis with regular medical prevention consultations depending on the presence or absence of risk factors.

The breast cancer surgical treatment is essentially based on only one of two procedures: mastectomy or BCCT. Contrary to mastectomy where the entire breast is removed, in conservative treatment, the tumour is excised macroscopically together with a margin of cancer free breast tissue. Subsequently the patient undergoes radiotherapy to the remaining breast tissue. This conservative approach, in its various forms, has made local control of the disease possible, with survival rates similar to those obtained with mastectomy, but with better cosmetic results [3, 4].

The aesthetic result and, patients’ Quality of Life (QOL), have become fundamental objectives in this type of treatment. However, results depend on several factors (patient, tumour, surgical technique and factors related to treatment, and of course radiotherapy) which contribute to very different outcomes [5, 6].

The assessment of the cosmetic result, as a mean to evaluate one of the aspects of treatment quality, has become essential to any institution performing breast cancer treatment [7]. This evaluation helps in the refinement of current techniques and to identify variables with a significant impact on the final aesthetic result and susceptible of improvement [5]. With the development of new oncoplastic techniques in breast conservation, it is even more important to have means to compare cosmetic results. The quantification of these results will rend possible to tailor the spectrum of techniques.
II. DETERMINANT FACTORS TO COSMETIC OUTCOME

Breast conservation surgery, radiotherapy included, has not a definite standard for every patient. There are technical variations related both to surgery and radiotherapy and even though BCCT has an identical meaning worldwide, it is not a uniform procedure [6]. To obtain a good aesthetic outcome, it is essential to have a clear notion of the factors that can influence results [8]. Historically, factors with an impact on aesthetic outcomes are divided into patient, tumour and treatment-related factors [9] (see Table 1).

Factors most frequently reported as being important to the final aesthetic result are patient weight [10], breast size [10-13], tumour location [14-16], tumour size [10, 11, 17-19], specimen weight or volume [5, 10, 11, 13, 16, 19-21], placement of incisions [5, 10, 13, 19], chemotherapy [17, 22-24] and radiotherapy to the breast [16, 20, 25].

Today with upcoming oncoplastic techniques, new factors related essentially to surgery will arise. If the specimen size and weight are seen as the most consensual of the factors, and if this concept is applied to the therapeutic reduction mammaplasty technique, where large volumes are removed along with the tumour, there will be a paradoxical effect of a better aesthetic outcome because both breasts are interventioned simultaneously [26]. Nevertheless, having a clear idea of which factors influence the cosmetic result is essential to achieve a reproducible way of measuring these outcomes, regardless of the technique used. Only then can we correlate results with practice and subsequently improve results.

III. COSMETIC ASSESSMENT PARAMETERS

To evaluate the aesthetic result of BCCT, an observer usually identifies and evaluates colour, shape, geometry, irregularity and roughness of the treated breast, compared to the untreated breast. Ideally, the comparison of the aesthetic result of BCCT would be performed using pre-treatment images of the patient. Unfortunately, the habit of capturing images before initiating treatment is not a rule in the majority of centres treating breast cancer. As a consequence, published works about the aesthetic results of BCCT use the comparison between both breasts, assuming that better results correspond to more similar breasts. This is indeed a much more practical approach and, in the era of new emerging oncoplastic techniques where both breasts are frequently submitted to surgical procedures, comparison is even more demanding [8]. The ultimate objective is to obtain identical breasts taking into account not only the surgical aspect, but also the adjuvant radiotherapy treatment applied in the majority of patients.

There are several factors contributing to global aesthetic result. Asymmetry in size is probably the most important factor which impacts the global cosmetic result and primarily depends on the amount of excised tissues, when surgery is unilateral [5, 12, 13, 21, 27-31]. Surgery and radiotherapy derived fibrosis, with upward retraction of the inferior mammary sulcus and the nipple-areola complex (NAC) can also have a consequence in the evaluated symmetry without impairing the size of the breast [12, 21]. There are other fac-

<table>
<thead>
<tr>
<th>Patient-related</th>
<th>Tumour and treatment-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Incision size (cm), tumour size (mm)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>Incision type (radial, sulcus, periareolar, curved)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>Specimen weight (g)</td>
</tr>
<tr>
<td>Thorax perimeter (cm)</td>
<td>Side (left, right)</td>
</tr>
<tr>
<td>Body mass index (BMI)</td>
<td>Tumour location (UOQ, UIQ, LOQ, LIQ, OQT, SQT, IQt, LQqt)</td>
</tr>
<tr>
<td>Breast size (radius cm)</td>
<td>Type of intervention (tumorectomy or tum, with sentinel node biopsy or with axillary dissection)</td>
</tr>
<tr>
<td>Bra cup size (A, B, C, D)</td>
<td>Scar visibility (not, slightly, very visible)</td>
</tr>
<tr>
<td>Age of menarche</td>
<td>Surgeon (general, breast → 150 cases / year)</td>
</tr>
<tr>
<td>No. pregnancies</td>
<td>Chemotherapy (yes, no)</td>
</tr>
<tr>
<td>Age of first pregnancy</td>
<td>Hormone therapy (yes, no)</td>
</tr>
<tr>
<td>Oral contraceptives (yes, no)</td>
<td>Follow-up (months)</td>
</tr>
<tr>
<td>Menopause (yes, no)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Patient, Tumour and Treatment-related Factors (from [9])
tors with great impact on aesthetic result. Scar visibility and length can also influence the aesthetic result [5, 28, 31, 32]. Finally, other aspects that need to be considered, generally attributed to radiotherapy, are differences in colour, both the hyperpigmentation of the treated breast or the hypopigmentation of the NAC complex, and to a lesser extent telangiectasias [5, 32-34].

IV. PATIENT EVALUATION

Patients are usually evaluated either by direct observation [15, 18, 35-41], using conventional photographs or slides [5, 12, 25, 31, 32, 42-44], special cameras - telecameras [30, 45] or more recently using digital images [46-48].

Direct observation of patients is promoted by the vast majority of personnel involved in BCCT as the most complete form of cosmetic evaluation, allowing physicians not only to get a better global appreciation of the results but also of other factors, which are not visualised in a photograph, such as skin atrophy and oedema of the breast and arm [5, 44]. However, photographs are more practical. Images can be saved permanently and visualised when necessary, and can be easily observed by different physicians [32, 46]. More recently, digital photographs have replaced paper prints and slides, making the entire process even easier and less expensive, while maintaining picture quality [46-48].

Eadie et al. [46] compared the aesthetic evaluation of breast conserving treatment in 10 patients by direct observation, printed analogical photographs, printed digital photographs and digital images on screen. Although printed analogical photographs presented better results, the correlation of results between direct observation and images did not depend on the type of images used [46].

V. SCALES OF ASSESSMENT

The most widespread scale used since the beginning of conservation breast procedures, until today, is the Harvard scale introduced by Jay Harris in 1979 [25]. It classifies the overall cosmetic result in four classes: Excellent (treated breast nearly identical to the untreated breast), Good (treated breast slightly different from the untreated one), Fair (treated breast clearly different from the untreated breast but not seriously distorted) and Poor (treated breast seriously distorted). Harris also introduced a scoring system to classify both the results of the surgical procedure (0-scar unapparent, 1-scar apparent, 2-major tissue loss) and of radiotherapy (0-none, 1-slight, 2-moderate, 3-severe).

Several other authors have tried to obtain a more complete and discriminative classification which included other factors such as skin oedema and thickening, mammary fibrosis, retraction and telangiectasias [29, 49, 50]. They used scores for each of the considered parameters and added the values to obtain a final score that would allocate the patients in the four classes described by Harris [5, 50, 51] or into other similar scales, some using more than four categories [12] and others using less [52].

More sophisticated scores have been described with the objective of measuring BCCT aesthetic result [29, 52]. However, none of these have been established as a preferred alternative to the Harvard scale. In the absence of more valuable alternatives, these scales, while subjective and without discrimination power, provide useful information on the cosmetic result of BCCT [36, 38, 53].

VI. METHODS OF ASSESSMENT

Generally there are two groups of methods used on the cosmetic evaluation of BCCT: subjective and objective methods. Subjective methods include patients’ self-evaluation, evaluation by a single observer or by a panel of observers. Objective evaluation depends essentially on the type of quantitative measurements used. For both groups evaluation can be made directly on the patient or using photographs (prints, slides or digital images).

A. Subjective Methods

1) Self-Evaluation: Patient self-evaluation is probably the easiest way to analyse the cosmetic outcome of BCCT. It is surely the one that best translates the psychosocial adaptation of patients to the result [54]. However, its reproducibility is low because it depends on several factors not amenable to quantification, such as age and socioeconomic status [29, 30], each one of them having a direct impact on how women see themselves after treatment. For those who support this method, the self-evaluation process and satisfaction with the result are the most important factors [55]. In works where self-evaluations are compared with external observer evaluations, patients invariably evaluate themselves more favourably [5, 10, 22, 25, 30, 53, 56, 57]. One plausible reason for this better evaluation is related to the fear that patients have to express criticism towards treatment or responsible caregivers [53]. This feeling is even stronger when a self-evaluation questionnaire is answered during a follow-up visit [43]. However, this fear can exist even with questionnaires sent by e-mail [55]. Another important point is the fact that BCCT patients tend to consider this as an alternative to mastectomy and even when results are worse, they are still seen as a better option than having the entire breast removed [43, 55].

It is undoubtedly important to have feedback from patients on aesthetic results; however, the low reproducibility of this method makes it impossible to use it for comparison purposes [5, 30, 32, 53, 56, 57].

2) Observer Evaluation: Currently, the most used evaluation of aesthetic results is the subjective assessment made by one or more observers, who focus directly on patients or on photographic representations of them [22], [40]. The final aesthetic result is assessed using, usually, the Harvard scale [25], where results are ranked according to comparisons between the operated breast and the untreated breast. This evaluation is undertaken by one [18, 22, 36-38, 49-51, 57-60] or several observers [5, 10, 17, 30-33, 39, 41, 46, 53, 61-64].

Even though this subjective evaluation has been criticised, it is still the most frequently used method [36, 38, 51, 53, 61, 65]. However, there are some problems regarding the interpretation of results from studies which use this type of assessment [66]: for example, impartiality is not always guaranteed since it is often performed by professionals involved in the treatment [18, 20, 22, 38, 49, 58, 59, 67, 68].
Moreover, its reproducibility is usually not high and agreement, when measured between observers, is only low or moderate [5, 32, 44]. The impracticality and the invasion of patients’ privacy inherent to this type of evaluation is not negligible: if the evaluation is carried out by one or more observers involved in the treatment process and at different monitoring stages, the discomfort caused to patients is high, especially when the evaluation is performed directly and not through the use of photographs.

Another negative aspect related to this type of evaluation is the impractical side of the method, especially if we think that patients should be evaluated by multiple observers in different phases of treatment and follow-up. Even using digital images and electronic mail evaluation, this process can take months or even years, when performed in a large number of patients with multiple observers [69].

B. Objective Methods

Pezner et al. [44] published an evaluation of a series of frontal views from 14 patients performed by 44 observers that challenged the value of subjective evaluations in terms of reproducibility. They also proposed the first measure to objectively assess the cosmetic outcome of BCCT: the Breast Retraction Assessment (BRA) [21], which determines the treated breast retraction comparatively to the untreated breast. Using a marked acrylic sheet over the thorax of the patients, they calculated upward and inward retraction of the treated versus the non-treated breast. Higher BRA values corresponded to worse cosmetic results. BRA was subsequently correlated with tumour size, chemotherapy and radiation fields [21].

BRA and the measures suggested afterwards in the literature (solely capturing the breast asymmetry) were often complemented by a subjective assessment made by observers [5, 12, 29-33, 45, 56, 70]. Even with these additions, all available methods were subject to significant intra-observer and inter-observer variability. More than trying to evaluate the aesthetic result from the measurements, these methods correlate them with subjective evaluation. Current methodologies in use continue to show a significant lack of standardisation, not only in the type of assessment used, but also in terms of the factors that should be considered in this evaluation and even in the instruments used for this analysis.

VII. OBJECTIVE ASSESSMENT OF BREAST AESTHETIC OUTCOME

Objective methods were introduced in an attempt to overcome the lack of objectivity and reproducibility. There are two different methodologies to objectively evaluate the aesthetic result: patients’ 2D photographs and 3D information. Fig. 1 shows some of the most significant works related to the objective assessment of breast aesthetic outcome in breast cancer treatment.

A. Two-dimensional Patient Data

These methods consist of comparing the two breasts with simple measurements marked on patients’ photographs [5, 12, 21]. These methods are essentially based on asymmetries between treated and nontreated breasts that improved the assessment reproducibility (see Table 2). However, it has been argued that they do not take into account the global appearance of the aesthetic results, due to the lack of the third dimension [5, 29, 30, 32].

1) First-generation Methods: As stated before, the first measure proposed for the objective assessment of the cosmetic outcome of BCCT was the BRA [21] to determine retraction in the treated breast compared to the untreated breast. Using identical reasoning applied to patient photographs, Van Limbergen et al. [12] came up with two new asymmetry measurements, besides the BRA: the Lower Breast Contour (LBC) and the Upward Nipple Retraction (UNR). They also found a strong correlation between the obtained values and a subjective classification performed by observers.

Tsouskas and Fentiman [70] described the Breast Compliance Evaluation (BCE) as a measurement of objective cosmetic outcome. Compliance is the difference between the distance from the nipple to the Infra-Mammary Fold (IMF) with the patient in the erect and supine positions. The same line of thought was followed by several other authors in more recent work [5, 30, 32, 33, 56].

Noguchi et al. [45] introduced an innovative sum of objective and subjective evaluations. The objective part was undertaken with a Moiré topographic camera, which solved the differences between the displayed curves in both breasts. Other parameters were evaluated subjectively by observers (skin changes and scar) and the final result was the sum of both evaluations.

The same path was followed by others who used both types of method and the final result was always the sum of the two obtained scores [29, 31]. The idea of adding the results of both subjective and objective evaluations resides in the fact that, for the majority of authors, neither method is successful in translating the complete evaluation of aesthetic results. When a bad aesthetic result is obtained, for example, due to a very visible and badly placed scar, the correlation
between asymmetry measurements and subjective classification is obviously lost [12].

A study to examine the convergence between objective and subjective indices of cosmetic and functional status after patients undergo BCCT was developed by Krishnan et al. [29], where the relations between the objective indicators and QOL were tested. They found that these scores were well correlated with the patients’ self-reported cosmesis questionnaires. The findings revealed positive cosmetic and functional treatment outcomes, showing good or excellent cosmetic results.

The fifth edition of the EORTC manual for clinical research in breast cancer published in 2005 [71] also stressed the idea of adding qualitative and quantitative evaluations. The subjective evaluation should be made by a panel of at least five observers classifying results according to the Harris Scale [25] and an objective evaluation using the measurements of asymmetry described by Pezner [21], and finally adding Turesson Skin damage classification [72, 73], graded as the area of telangiectasias and skin necrosis, due to radiotherapy.

The first-generation of models introduced objective quantification of individual aspects of the overall cosmetic outcome. These developments represent significant advances from a complete reliance on expert judgement; however, we have yet to see these capabilities substantially differentiate operational evaluation methodology. The simple sum of individual indices, objective and subjective, fails to acknowledge that different aspects have different contributions to the overall result; the reliance on the subjective assessment of individual aspects makes the final decision subjective and irreproducible; even for the objective measures, the manual measurement makes them partially subjective; as a consequence the process is still time consuming and hard to implement under routine clinic conditions. In many ways, the development of these models has been a step in the right direction, but we are still far from an acceptable system.

2) Second-generation Methods: Two approaches emerged in 2007 in an attempt to bring something new to the evaluation of aesthetic results in breast conservation [47, 48]. These methods were introduced to make it possible to predict the global aesthetical result. The software tools are based on different individual characteristics automatically and objectively extracted from patient photographs. These approaches aimed to overcome the acute shortage of such tools and to exploit the unique ability that computational methods have to provide an effective, easy, fast, reliable and reproducible tool to evaluate the outcomes of breast cancer patient care. In both works, the team started by creating a database of patient images evaluated subjectively by a panel of experts, obtaining classifications which served as a ground truth the developing new methods and comparing results.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Paper</th>
<th>Year</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turesson and Notter</td>
<td>[72-73]</td>
<td>1984</td>
<td>Turesson Skin damage classification graded as the area of telangiectasias and skin necrosis, due to radiotherapy</td>
</tr>
<tr>
<td>Pezner et al.</td>
<td>[21]</td>
<td>1985</td>
<td>Breast Retraction Assessment (BRA), which determines the treated breast retraction comparatively to the untreated breast</td>
</tr>
<tr>
<td>Linbergen et al.</td>
<td>[12]</td>
<td>1989</td>
<td>Lower Breast Contour (LBC) and the Upward Nipple Retraction, which determines the nipple displacement and breast contour retraction of the treated breast respectively, comparatively to the untreated breast</td>
</tr>
<tr>
<td>Tsouskar and Fentiman</td>
<td>[70]</td>
<td>1990</td>
<td>Breast Compliance Evaluation (BCE), which determines the difference between the distance from the nipple to the Infra-Mammary Fold (IMF) of the treated breast, comparatively to the untreated breast</td>
</tr>
<tr>
<td>Noguchi et al.</td>
<td>[45]</td>
<td>1991</td>
<td>Moiré topographic camera to compute the differences between the displayed curves in both breasts</td>
</tr>
<tr>
<td>Cardoso and Cardoso</td>
<td>[47]</td>
<td>2007</td>
<td>BCCT.core software for the aesthetic outcome assessment entailed a semi-automatic extraction of features considered to have an impact on the overall cosmetic result (asymmetry, color differences and scar visibility features)</td>
</tr>
<tr>
<td>Fitzal et al.</td>
<td>[48]</td>
<td>2007</td>
<td>BAT© software and Breast Symmetry Index (BSI) definition to evaluate the cosmetic outcome of BCCT</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>[82]</td>
<td>2007</td>
<td>Breast ptosis measurement based on ratios of distances between fiducial points manually identified in oblique and lateral patient photographs</td>
</tr>
<tr>
<td>Oliveira et al.</td>
<td>[81]</td>
<td>2010</td>
<td>Features extracted from lateral patients photographs for the aesthetic outcome assessment</td>
</tr>
<tr>
<td>Lee et al.</td>
<td>[83]</td>
<td>2012</td>
<td>Quantitative measure of breast curvature based on catenary.</td>
</tr>
</tbody>
</table>
Fitzal et al. [48] described a Breast Symmetry Index (BSI) to evaluate the cosmetic outcome of BCCT. They created software (Breast Analysing Tool - BAT©) to measure differences between left and right breast sizes from a patient’s digital picture (see Fig. 2).

The sum of all differences results in the breast symmetry index (BSI) score, which is capable of measuring differences in size between breasts. They were able to demonstrate BAT© accuracy and additionally showed that BSI is reproducible and may be used by experts and non-experts with the same results. The BSI index measured from frontal pictures correlates highly with subjective votes from experts and significantly differentiates between good and fair cosmetic outcomes. However, there was no correlation between the BSI index and the patients’ self-evaluation. Moreover, the BSI did not further differentiate between excellent and good or fair and poor cosmetic outcomes. These data show that the BSI measured with the BAT© may be used to some extent to evaluate cosmetic results of BCCT.

Cardoso and Cardoso [47] introduced a computer-aided medical system named BCCT.core (see Fig. 3). The development of BCCT.core entailed the semi-automatic extraction of several features from patient photographs, capturing some of the factors which are considered to have an impact on the overall cosmetic result. To accomplish cosmetic categorisation, a concise representation of a BCCT image is first obtained based on asymmetry, colour differences and scar visibility features.

The asymmetry between both breasts was conveyed under a large set of indices. To extract the colour features, the colour content of each breast is first characterised by a colour histogram; the dissimilarity between both histograms serves as the colour features. The scar visibility was translated into local colour dissimilarity, by comparing corresponding sectors of the breasts. The measurements are preceded by the automatic localisation of fiducial points (nipple complex, breast contour and jugular notch of the sternum) [74] on the photographs. Measurements for asymmetry, skin colour change and surgical scar appearance are then supported on these fiducial points. The representation is then further analysed by a pattern classifier performing the categorisation [75]. The pattern classifier finds the combination of individual features that approximate best the overall classification as given by the panel of experts. A correct classification rate of about 70% was obtained, when comparing with the subjective consensual classification from the panel [69].

3) Summary: The BAT© and the BCCT.core have been recently compared [76] regarding the same set of cases. The conclusion of this study, was that the BAT© software maintains a moderate performance regardless of picture quality using only asymmetry measurements. On the other hand, the BCCT.core, using other parameters such as colour differences and scar appearance, obtains different results with good performance with better photographs but only moderate performance in lower quality pictures [77]. Moreover, the BCCT.core has been used in many works related with BCCT cosmetic evaluation [78-80].

Fig. (2). BAT© software interface from Fitzal et al. [48].

Fig. (3). BCCT.core software interface.

These tools, while innovative and reproducible, have several drawbacks that need to be addressed. Both use only frontal photographs of the patients, which do not include lateral or oblique views. This approach is rare in the history of the aesthetic outcome of BCCT [81] and only a small number of researchers have used lateral features in their work. The exceptions include Kim et al. [82], who compare objective measurements based on breast ptosis with ratings on a subjective scale made by experienced clinical observers and more recently, Lee et al. [83] who introduces a novel quantitative measure of breast curvature based on catenary. They compared the length, the area enclosed by the curve, and the curvature measure from the catenary curve to those from manual tracings of the breast contour. Likewise, the same procedure was applied to frontal photographs. This can be related with the difficulty of extracting robust features from lateral views. The question is: “is it important to include this information?” Fitzal et al. [48] stated that lateral and oblique views are difficult to standardise and the solution will be to move into a 3D version of the software.

B. Three-Dimensional Methods

Potential advantages of 3D imaging as a tool for objective cosmetic evaluation include the ability to view the breast from a significant different number of angles, to estimate volume/volume deficit and to plan future surgeries. There are several variations of this tool, from relatively simple volumetric analyses to more sophisticated programmes, which make it possible to perform quantitative measurements and software that makes it possible to simulate the post-operative outcome. Although these benefits are known,
some authors still doubt the ability of current 3D image analysis methods [84].

1) Existing Systems: Several research groups have recently made attempts with 3D approaches (see Table 3). 3D imaging in breast surgery can be used for several clinical purposes. The most practical applications are evaluating breast asymmetries, both congenital and acquired, and factors affecting breast shape in augmentation mammoplasty. Other uses of 3D imaging for clinical purposes include the evaluation of patients who wished to have a reduction mammoplasty and the evaluation of patients undergoing unilateral breast reconstruction to determine the expander and permanent implant size that gives the best symmetry with the contralateral breast.

Galdino et al. [85] investigated the use of clinical 3D imaging to determine quantitative information about the breast, such as volume or projection. They applied this approach to real cases, providing objective data on the breast and surgical mammoplasty (especially augmentation mammoplasty). This helped surgeons better understand the factors that contribute to breast shape and influence surgical outcomes. A couple of limitations were found, highlighted by patients with significant ptosis or suffering from obesity, which may introduce errors into the 3D data, making them unreliable.

Losken and colleagues [86-88] developed an objective technique based on a 3D camera and software to quantify the cosmetic results of BCCT. This software package enables the comparison between the treated and untreated breast by analysing the surface area and volume differences. The camera includes 12 individual digital lenses arranged in 3 planes with a single focal point at the manubrium. Images are captured with the patients’ arms in two different positions: at

---

Table 3. Comparison of the Most Significant 3D Technologies Involved on Breast Cancer Cosmetic Evaluation

<table>
<thead>
<tr>
<th>Type</th>
<th>Authors</th>
<th>Paper</th>
<th>Year</th>
<th>Technology</th>
<th>Type of Vision</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-camera</td>
<td>Losken et al.</td>
<td>[86-88]</td>
<td>2005-2008</td>
<td>Camera with 12 individual digital lenses arranged in 3 planes with a single focal point</td>
<td>Passive</td>
<td>High</td>
</tr>
<tr>
<td>Stereo-photogrammetry</td>
<td>Bert et al.</td>
<td>[89]</td>
<td>2005</td>
<td>3D stereo-photogrammetry and surface registration</td>
<td>Passive</td>
<td>Average</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Isogai et al.</td>
<td>[90]</td>
<td>2006</td>
<td>3D laser scanning and quantitative comparison with Moiré patterns projection</td>
<td>Active</td>
<td>Average</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Farinella et al.</td>
<td>[91, 92]</td>
<td>2006</td>
<td>3D laser scanning combined with anatomical landmarks</td>
<td>Active</td>
<td>High</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Tepper et al.</td>
<td>[93, 94]</td>
<td>2006-2008</td>
<td>3D laser scanning and surface registration</td>
<td>Active</td>
<td>High</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Balaniuk et al.</td>
<td>[95]</td>
<td>2006</td>
<td>3D laser scanning in combination with a generic 3D model of the internal structure of the breast using virtual reality approaches combined with soft tissue modeling</td>
<td>Active</td>
<td>High</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Kovacs et al.</td>
<td>[96-98]</td>
<td>2006-2007</td>
<td>3D laser scanning and comparison with reference measurements obtained nuclear magnetic resonance imaging</td>
<td>Active</td>
<td>High</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Catanato et al.</td>
<td>[99-101]</td>
<td>2008-2009</td>
<td>3D laser scanning and stereo optoelectronic tracking system with breathing artifact correction</td>
<td>Active</td>
<td>High</td>
</tr>
<tr>
<td>Multi-camera</td>
<td>Henselet et al.</td>
<td>[102]</td>
<td>2011</td>
<td>3D model generated with 4 pods and 8 cameras, 2 cameras on each pod and comparison with volume measured with BAT software [48]</td>
<td>Passive</td>
<td>Average</td>
</tr>
<tr>
<td>3D Low Cost RGB-D camera</td>
<td>Oliveira et al.</td>
<td>[103, 104]</td>
<td>2011-2012</td>
<td>Disparity map generation with Kinect device. Automation of prominent points and breast contour, and comparison with results obtained with BCCT.core [47]. Volumetric detection using a low cost solution</td>
<td>Active (Infra-red)</td>
<td>Low</td>
</tr>
<tr>
<td>3D Laser Scanner</td>
<td>Eder et al.</td>
<td>[105]</td>
<td>2012</td>
<td>3D laser scanning by overlapping the mirrored left breast with the right breast. 3D contour definition and comparison with contour obtained with BCCT.core software [47]</td>
<td>Active</td>
<td>Average</td>
</tr>
</tbody>
</table>
their sides and on their hips. The software is then used to determine the level of asymmetry between the breasts by overlapping the 3D meshes from the two breasts (see Fig. 4).

Bert et al. [89] described a 3D stereo-photogrammetry system which captured 3D surface anatomy data, making it possible to perform a volumetric assessment of the breasts. In order to reproduce the position of the patient in fractionated radiotherapy, the patient surface model is registered to a previously recorded reference surface. This is performed using surface registration, where the system determines the rigid-body transformation that minimises the distance between the treated and the reference surface models in a region-of-interest (ROI). The implemented system is accurate and sufficiently stable to be used in clinics. Errors were computed by comparing the surface model with the computed tomography (CT) geometry, and were in the order of submillimeters. Alignment deviation and gating-signal tests were of the same magnitude.

A quantitative analysis using 3D imaging was proposed by Isogai et al. [90], who used laser light scanning to generate 3D images, which were then subjected to Moiré patterns for a quantitative comparison of the reconstructed breast with the normal breast. The evaluation is based on breast symmetry, volume and shape after breast reconstruction using the 3D laser light scanner system. Captured images of the normal breast were mirror-reversed and overlapped on images of the reconstructed breast.

Farinella et al. [91, 92] used 3D laser scanning combined with anatomical landmarks identified by surgeons and developed the Breast Shape Analyzer 0.1 (BSA 0.1) software, which provided useful objective quantitative measurements to surgeons. This technique uses only well-defined anatomical points (Fig. 5(a)), which have been identified and selected by surgeons.

A simple sequence of geometric operations is performed to divide the breast surface into four anatomic subunits, according to the clinically derived breast meridian and equator lines. Using this breast shape partition, useful measurements can be extrapolated on a 3D model data set. The models used in this work were collected following the most common criteria for cosmetic evaluation and reconstruction (such as age, menopausal status, ptosis classification). The acquisition was made using a commercial laser scanner, applied on volunteers sitting on a chair with their back at 45 degrees (Fig. 5(b)). Each volunteer was scanned three times: facing the camera and rotating the chair at 45 degrees to the left and to the right. This system has an important limitation related to some of the patients’ uncontrollable physiological movements. During the acquisition, the breathing and imperceptible body torsions influenced the data, and thus the authors decided not to merge the different scans. The authors suggest that this can be overcome by adopting a faster but less accurate shape-from-video technique.

Tepper et al. [93, 94] used 3D imaging to enable an objective volumetric analysis of breasts prior to breast reconstruction. Surgeons are then capable of visualising the size, shape, contour and symmetry of the breast with 3D breast models, as well as obtaining quantitative breast measure-
ments and volumetric calculations. This application represents a significant advance from traditional approaches to aesthetic and reconstructive breast surgery.

A 3D imaging tool was developed by Balaniuk et al. [95], using virtual reality approaches combined with soft tissue modelling methods which allow surgical simulations to predict cosmetic outcomes and options for breast augmentation and reconstructive procedures. They discuss the adequacy of the different medical procedures to the patient’s choice, and state that it is very important, for patients and for surgeons, to have a tool to visualise the potential outcomes of the surgery and make decisions on their surgical options. In (Fig. 6), we can see an example of the model created with this application for a breast with and without an implant.

Fig. (6). Application example from Balaniuk et al. [95]. The breast with and without an implant.

In the work carried out by Kovacs et al. [96-98], it is stated that the 3D scanner represents a simple and promising method to measure the breast volume. In their work, they present a research study on the optimisation of the imaging technology for the mammary region with a 3D laser scanner, which is used to evaluate the method’s precision and accuracy, and to allow optimum data reproducibility. The standardisation of the imaging and scanning system is decisive in order to guarantee that the method is reproducible. They have tested the most favourable imaging technology on dummy models for scanner-related factors, such as the scanner position in comparison with the torso and the number of scanners and single shots. This application showed the correct settings for 3D imaging of the breast region with a laser scanner with an acceptable degree of accuracy and reproducibility.

Catanuto et al. [99, 100] present a set of parameters to unambiguously estimate the shape of the natural and of the reconstructed breast, using an optoelectronic tracking system in seven female volunteers. This study allows a real-time breathing artifact correction [101] and a surface patch fusion with no intervention from the operator. With these parameters, it is possible to describe several anatomical geometrical properties, such as: distances (linear measurements between relevant anatomical landmarks); surface measurements (total surface area, anatomical subunits area); angles (divergence between the nipples from the patient’s point of view, inclination of the breast mound on the chest wall); curvature (curved surface properties); symmetry (symmetry between the two sides of the chest wall); and IMF shape (natural characterisation and reconstructed breast shape). With this technique, the most interesting result obtained was a graphic depiction of the curvature of the thoracic surface. While the first, the "divergence angle", gives a perspective on the point of view of the patient's own breasts (Fig. 7(a)), the latter introduces a representation of the breast based on a colour map which expresses flat regions and curvature (Fig. 7(b)). The authors conclude that the colour based curvature representation can replace volume terminology, as curvature can be expressed either with colours or its degree by coefficients. The breast surface was segmented into four quadrants using reproducible landmarks. The proposed methodology relies on geometric planes (bilateral symmetry plane, meridian plane and equatorial plane) that make it possible to calculate significant clinical angles.

Fig. (7). Application example from Catanuto et al. [99]. a) Example of divergence angles on a 3D model; b) Example of colour based map. Colour scale associates values to a curvature colour scale.

A 3D multiple stereo camera system was introduced by Henseler et al. [102] to objectively assess breast shape. The system is composed of 4 pods and 8 cameras, 2 cameras on each pod. Several dummy patient breast models and breast surfaces were captured. Then a 3D model was constructed and the volume was measured with BAT software [48]. The researchers stated that the obtained results were very satisfactory, on 3D data obtained from the dummy models. In the real breast model, all measurements of total breast volume were consistently higher compared to the results obtained with stereophotogrammetry (see Fig. 8).

Fig. (8). Application example from Henseler et al. [102]. Segmented breast on 3D model, breast segment with chest wall.
Moreover, the arbitrary identification of the chest wall during the immersion of the breast was the cause of the poor reproducibility of this method. This highlights the technical difficulties of the method. The measurements of volume were more reproducible with the 3D imaging method. Nevertheless, the true volumes of the breasts remained uncertain due to undefined breast boundaries.

Eder et al. [105] conducted a study based on a 3D evaluation protocol to analyse breast symmetry according to 3D breast contour differences between the left and the right breasts using surface imaging. In this study, researchers objectively compared breast symmetry using a 3D scanner by overlaying the mirrored left breast with the right breast and determining the mean 3D contour difference between the two breast surfaces. Three observers analysed the evaluation protocol using two dummy models. They also evaluated the potential of this approach in clinical applications comparing it with BCCT.core [47], on 23 breast reconstruction patients. The authors concluded that this approach could assist surgeons in the pre-operative planning and optimisation of breast corrections after reconstruction and breast conservative approaches.

In 2010 Tepper et al. [106] provided an overview of 3D breast photography, with an emphasis on its potential role to establish a standardised system for breast analysis, by introducing a new concept entitled “mammometrics”. In this concept, 3D-based breast measurements can be used to help guide operative planning and objectively analyse surgical results. In their work, the researchers validated the use of 3D breast photography in various clinical arenas, including autologous breast reconstruction, prosthetic breast reconstruction, reduction mammoplasty, and augmentation mammoplasty. They state that 3D imaging technology has the ability not only to obtain well-established breast measurements accurately, but also to find new measurements that were not previously possible with conventional tools, such as total breast volume, volumetric distribution and breast projection.

Recently the Kinect (Microsoft Corp. Redmond, WA) [103, 104] was introduced as promising low cost and easy to use equipment for BCCT cosmetic evaluation. This tool can not only facilitate automation, but also provide volumetric information (see Fig. 9).

**Fig. (9).** Patient photograph and generated disparity map.

With this solution authors improved automation of prominent points and the breast contour [103], comparatively to those obtained with the original BCCT.core [47]. Furthermore, other promising results were achieved, including the detection of volumetric differences of the breasts using the disparity map generated from the Kinect [104]. It was shown that depth-map images facilitate the automation of BCCT.core and thus this software remains low cost and an easy to use solution. Obtained results also showed an excellent performance and robustness for a wide variety of patients. Authors believe that this kind of systems could be a feasible solution to use in the future for 3D assessments of BCCT results.

2) **Summary:** All research work presented was particularly related to BCCT and more recently other types of breast cancer loco-regional treatment, and also with plastic surgery. Some of the approaches can only be applied to a specific procedure because the characteristics to consider in each situation may be different and the precision level can also differ according to each situation.

Currently, the approaches are mainly based on 3D laser scanners using active and passive lighting. However, these methods do not fill completely physicians’ expectations. The most recent works [102, 105], are based on methodologies introduced previously, and compared the performance with results obtained with 2D tools [47], [48], which continue to be used as a reference. Over the last few years unfortunately there was not a technological evolution of these methodologies, foreseeing a radical change, which may well pass through the use of simpler approaches based on low-cost systems. The main drawback of 3D techniques is the demand for specialised hardware, software and personnel. The high cost and the difficulty of using these methods on a daily basis prevent their widespread use in the near future. Additionally, almost all currently used techniques based on 3D models do not try to predict the aesthetic result for a more informed choice of treatment, nor are they suitable for the automatic evaluation of the aesthetic result after the surgery [85, 97, 99]. Other issue is related with the long acquisition time needed to obtain the 3D model, which can result in discomfort for the patient. During the period of acquisition any movement made by the patient can result in unreliable 3D models. Apart from the 3D scanners based systems, researchers are still using methodologies based on stereo vision [107] or multi-camera systems [108]. However, 3D reconstruction of breasts is a very difficult task since it is featureless.

Another important issue is related with automation of the software. All solutions need total or partial input from the user to obtain features that are used to achieve the expected result. The automatic identification of fiducial points can be difficult to localise, like the manubrium and sternal notch, especially with poor lighting and among overweight and obese patients. This makes a fair comparison between different researchers and institutions difficult. Moreover, reproducibility of the results is very difficult to obtain. The comparison between different methodologies is also very hard to achieve, since databases used in the different studies are not shared.

It would be desirable to have completely automated software, low cost and easy to operate hardware and the ability to compare results with public databases

**VIII. PUBLIC DATABASES**

The demand for more capable and accessible databases is welcome and necessary. Public datasets provide common
standards to measure and compare accuracies of proposed approaches. Databases should be available, preferentially over the Internet, and continuous user support is also indispensable. The importance of having public datasets is recognized, both to evaluate the performance of existing systems and to encourage the development of new algorithms. One of the trends is related to the database size. In this context, a larger data set will give more reliable results. Another trend is that databases have to become more diversified, with respect to patient characteristics, or surgical interventions. Annotation of the data, ground truth, is also mandatory in order to give an independent comparison of the results obtained from different methodologies. In particular, BCCT authors are not accustomed to provide access to the data used in their works, in contrast to radiology researchers [109-111]. This fact makes it difficult to compare results, metrics and the different methodologies employed.

The only exception is the database provide by the Breast Research Group1. This dataset is composed of data acquired from three different institutions in Portugal, comprising photographs from 120 patients. Data acquired was collected from patients treated with conservative breast surgery, with or without auxiliary surgery and radiotherapy, with treatment completed at least one year before acquisition. All patients signed an informed consent to participate. Breast images were acquired employing a 4 Megapixel digital camera. Photographs were taken in four different positions with the patient standing on floor marks: facing, arms down; facing, arms up; operated side, arms up; contralateral side, arms up. Fig. 10 presents a typical set of images for a patient. A mark was made on the skin at the suprasternal notch and at the midline 25 cm below the first mark. These two marks create a correspondence between pixels measured on the digital photograph and the length in centimetres on the patient.

This database has two different kinds of annotation: subjective evaluation of the aesthetic result and the definition of breast contour and fiducial points (nipple position, suprasternal notch position and the 25 cm mark). The subjective evaluation used the opinion of several experts with experience in BCCT [66]. A set of 24 clinicians working in 13 different countries were selected, based on their experience in BCCT (number of cases seen per year and/or participation in published work on evaluation of aesthetic results). They individually evaluated a series of 120x4 photographs taken from 120 patients. They were asked to evaluate overall aesthetic results, classifying each case using the Harris scale [25]: excellent, good, fair and poor. In order to obtain a consensus among observers, the Delphi process was used [112, 113]. In the end, the consensus on aesthetic results was obtained in 113 of the 120 evaluated patients, equivalent to 94% of the cases. The distribution of cases is presented in Table 4. With respect to breast contour and fiducial points annotation, 8 independent users manually marked these features, in frontal patient views, using the BCCT.core software tool [47] (see Fig. 11). Both patient data and ground truth annotation files can be accessed at the website of the Breast Research Group2.

### Table 4. Distribution of the patients over the four classes.

<table>
<thead>
<tr>
<th>Class</th>
<th>#Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>14</td>
</tr>
<tr>
<td>Good</td>
<td>64</td>
</tr>
<tr>
<td>Fair</td>
<td>24</td>
</tr>
<tr>
<td>Poor</td>
<td>11</td>
</tr>
</tbody>
</table>

---

1 http://medicalresearch.inescporto.pt/breastresearch/

2 http://medicalresearch.inescporto.pt/breastresearch/
IX. CONCLUSION

The widespread method to evaluate aesthetic results in breast cancer treatment consisted of observer evaluations, known as subjective assessments. Harris et al. [25] introduced an overall cosmetic score: excellent, good, fair and poor. In this method, the aesthetic of the breast patient is evaluated directly or using a photograph. However, the global aesthetical results are the combination of several features, estimated and subjectively combined by the observers by visual inspection. Subjective evaluation is poorly reproducible, even when performed by specialists, which significantly limits comparison between centres dedicated to breast cancer.

Initial objective methods, which appear to overcome the setbacks of subjective methodologies, consisted of comparing the two breasts with simple measurements marked directly on the patients or in photographs. Trying to overcome the sense that objective asymmetry measurements were insufficient, other groups proposed the sum of the individual scores of subjective and objective indices. However, adding subjective evaluation may result in important setbacks related to the complete automation of the approaches, fundamental for high reproducibility. There remains a need to extract volumetric information in order to improve the overall cosmetic evaluation. Furthermore, the solutions should remain low cost.

Potentially, a more accurate and objective tool to predict surgical outcomes to guide the patient and surgeon in the decision-making or planning process is using 3D imaging and surgical simulations. A simulation model also allows patients to visualise the possible outcomes of different treatment options. It is generally accepted that 3D imaging has great potential in a clinical environment, although there are factors that may influence its use in the near future. The high cost of the equipment and the need for specialised people to operate are undesirable circumstances. Consequently, the search for low cost and easy to use equipment is highly desirable.

Two important conclusions are drawn from the concepts addressed previously. The first one is that cosmosis is an important objective of breast conserving treatment and it should be assessed along with oncological outcome. The second one is that it is essential to have easy and reproducible methods of assessment that are consistent worldwide. The immediate consequence of these two assumptions is that subjective methods must be replaced by objective ones to obtain the desired reproducibility and easy manipulation.

A good method of evaluation should be consistent and simple to use with all kinds of images. For example, standards for photograph quality (definition, backlight, background) are required if one is to expect discriminative power in evaluation of aesthetic results. Are we there yet? We do not think so but the way is now open to the development of these new techniques. With the upcoming new breast oncoplastic surgeries there will be an even greater demand for evaluation of cosmetic results. Radiotherapy also with the use of the partial breast irradiation techniques must now pave the way for a comparison between the new cosmetic results and the ones obtained with an entire irradiation of the breast.

In breast-conserving surgery, there is evidence that approximately 30% of women receive a suboptimal or poor aesthetic outcome; however there is currently no standardised method of identifying these women. We agree that a new generation of tools will see considerable growth in the near future; these tools, tailored to the individual patient, should also target the prediction (rather than just the evaluation of the treatment) of the aesthetic outcome of breast conserving surgery. The combination of 3D photography and routinely acquired radiological images (i.e. mammography, ultrasound and MRI, when available), together with information about the tumour (size, location, shape etc.) will enable evaluation of alternative therapeutic strategies and the consequences of the available options. This will aid communication with patients and empower patients to take an active role in a shared decision making process.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

The first author would like to thank Fundação para a Ciência e Tecnologia (FCT) - Portugal for the financial support for the PhD grant with reference SFRH/BD/43772/2008. This work is financed by the ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT within project PTDC/SAU-ENB/114951/2009. The work leading to these results has received funding from the European Community’s Seventh Framework Programme under grant agreement no FP7-600948.

REFERENCES


