REVIEW

Assessing Cosmetic Results After Breast Conserving Surgery

MARIA JOÃO CARDOSO, MD, PhD,¹* HELDER OLIVEIRA, PhD,² AND JAIME CARDOSO, PhD²

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

“Taking less treating better” has been one of the major improvements of breast cancer surgery in the last four decades. The application of this principle translates into equivalent survival of breast cancer conserving treatment (BCT) when compared to mastectomy, with a better cosmetic outcome. While it is relatively easy to evaluate the oncological results of BCT, the cosmetic outcome is more difficult to measure due to the lack of an effective and consensual procedure. The assessment of cosmetic outcome has been mainly subjective, undertaken by a panel of expert observers or/and by patient self-assessment. Unfortunately, the reproducibility of these methods is low. Objective methods have higher values of reproducibility but still lack the inclusion of several features considered by specialists in BCT to be fundamental for cosmetic outcome. The recent addition of volume caretakers, allowing not only the comparison of results but the improvement of performance.

This scenario suggests that evaluation of cosmetic outcomes should be essential in any institution performing breast cancer treatment, to improve current strategies, and enable the identification of variables that affect the final cosmesis and maybe amenable to improvement [4].

With the development of new oncoplastic techniques in breast conservation, it is even more important to evaluate and compare cosmetic results, thus helping to tailor the spectrum of techniques available to individual cases without compromising oncological, and cosmetic results [5].

In the process of accessing cosmetic outcome there are several important issues to be considered:

1. Which are the factors that have a determinant impact on cosmetic outcome in breast conservation treatment (BCT)?
2. Which parameters or features should be evaluated regarding cosmetic assessment of BCT?
3. How can patients be evaluated?
4. Which scales can be used for this evaluation?
5. Which methods are currently available for evaluation of cosmetic results of BCT?

KEY WORDS: breast cancer; conserving surgery; cosmetic assessment

INTRODUCTION

Breast-conserving therapeutic approaches to breast cancer (usually surgery and radiotherapy) aim to obtain, besides local tumor control, and survival rates equivalent to mastectomy, better cosmetic results [1,2].

While the oncological outcome of breast conservation procedures can be estimated objectively by disease-free survival and overall survival rates, cosmetic outcome has yet no standard of evaluation.

Although a considerable amount of research has been dedicated to breast-conservation techniques, the diversity of available procedures as well as different working practices, contribute to diverse cosmetic results [3].

This scenario suggests that evaluation of cosmetic outcomes should be essential in any institution performing breast cancer treatment, to improve current strategies, and enable the identification of variables that affect the final cosmesis and maybe amenable to improvement [4].

Factors That Have a Determinant Impact on Cosmetic Outcome

To obtain a good cosmetic outcome it is essential to have a clear notion of the factors that can influence results [4]. Analysis of these factors will obviously depend on the method used to evaluate results. If the evaluation method is not highly reproducible, correlation with involved factors will be variable between centers and as a consequence will lack the necessary impact, to be considered universal.

This correlation is also dependent of the timing. When a patient’s evaluation is done after surgery and before radiotherapy, consequences derive only from the surgical procedure. On the other hand, if patient’s evaluation is done after radiotherapy, cumulative effects of both will be present. The same reasoning applies to chemotherapy, hormonal therapy, etc.

In spite of all these setbacks, several papers have been published about the correlation of factors with cosmetic outcome of BCT. Results, however, are not consensual as a consequence of the diversity of methods and timing used to evaluate results.

Historically, factors impacting on cosmetic outcome are classically divided into patient, tumor, and treatment-related factors. Factors most frequently reported as being important to the final cosmetic result are
patient weight [6], breast size [6–8], tumor localization [7,9,10], tumor size [6,11,12], specimen weight or volume [6–8,13–15], placement of incisions [4,6,8,15], chemotherapy [12,16–18], and irradiation [7,12,19].

Today, with the upcoming oncoplastic techniques, new factors related to surgery will arise. Specimen size and weight have been until now probably the most consensual of the interfering factors in the cosmetic result of classic BCT. However with the new therapeutic reduction mammoplasties, larger volumes are removed along with the tumor having a paradoxical effect of a better cosmetic outcome since both breasts are simultaneously submitted to surgery. Also, new factors related to different radiotherapy techniques like partial breast irradiation will probably result in different cosmetic outcomes when compared with all-breast irradiation.

Parameters of Cosmetic Assessment

To evaluate the cosmetic result of BCT, an observer usually identifies and evaluates color, shape, geometry, irregularity and roughness of the visual appearance of the treated breast, compared to the untreated one. This type of examination is somewhat subjective, and to a great extent depends on the observers chosen for carrying out the analysis.

Ideally, the comparison of cosmetic results of BCT would involve the pre-treatment image of the patient. Unfortunately, the habit of capturing images before initiating treatment is not a rule in the majority of centers dealing with breast cancer. As a consequence all published papers have used the comparison between both breasts, assuming that better results correspond to more similar breasts.

In an era of emerging oncoplastic techniques where both breasts are frequently submitted to surgical procedures, this has become not only a more practical but also a more realistic approach [20].

Considering then that identical breasts are the ultimate cosmetic objective of BCT, asymmetry is the key parameter for analyzing cosmetic results. There are, however, several important features contributing to asymmetry.

Asymmetry in size, is probably the most important contributor to global cosmetic result. This is primarily dependent of the amount of tissue excised, particularly when surgery is unilateral (Fig. 1) [4,8,14,21–24]. Surgery and radiotherapy derived fibrosis can also impact on symmetry without impairing the size of the breast (Fig. 2) by causing upward retraction of the inferior mammary sulcus and the nipple-areolar complex (NAC) [14,25]. Scar visibility and length also influences cosmetic results contributing to asymmetry (Fig. 3) [4,8,21,24]. Other aspects that need to be considered are generally attributed to radiotherapy and include differences in color, both hyperpigmentation of the treated breast, and hypopigmentation of the NAC complex and to a lesser effect, telangiectasias (Fig. 4) [4,8,26,27].

How Can Patients Be Evaluated

Patients are usually evaluated either by direct observation or indirectly through conventional photographs or slides [4,7,14,16,19,21,28,29], special cameras – telecameras [30,31] or more recently through digital images [32,33].

Direct observation of patients is deemed by the vast majority of personnel involved in BCT to be the most complete form of cosmetic evaluation as it allows not only for a global appreciation of results but also for other factors that are not visualized in captured images, such as skin atrophy and oedema of the breast and arm [4,14]. Photographs, in contrast, while not permitting a detailed appreciation of these characteristics, allow for the visualization of the main factors contributing to final cosmetic results such as symmetry of both breasts, scar visibility and length, and color differences [4]. There are also several practical advantages associated with photographs: images can be saved permanently; visualized when necessary; easily analyzed by different observers [29,33]. More recently, digital photography has almost replaced paper prints and slides, making the whole process much easier and less expensive, while maintaining picture quality [32,33].

Eadie et al. compared the cosmetic evaluation of breast conserving...
treatment in 10 patients through direct observation, printed analogical and printed digital photographs and digital images on screen. Although printed analogical photographs were preferred by observers, correlation of results between direct observation and images was independent of the support material used [33]. Circulation of digital images raises specific concerns, even within a given hospital or clinic, and access should be restricted to appropriate professionals, using protected logins and passwords.

Scales of Assessment

The most widespread scale used in published papers since the beginning of conservation breast procedures until today is the Harvard scale, introduced by Jay Harris in 1979. It classifies cosmetic results in four classes: excellent; good; fair; poor (Fig. 5) [19]. Harris also introduced a scoring system to classify results of biopsy procedures (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 added other parameters in an attempt to obtain a more complete and discriminative classification. These include skin oedema and thickening, fibrosis, retraction and telangiectasias [23,34,35]. They use scores for each of the considered parameters and add them up to obtain a final value allocating patients to one of the four classes of the Harvard scale [4,35,36] or to other similar scales some with different categories (usually three or five) [7,37].

Other scores have been developed with the aim of obtaining a more objective classification [23,37] but none of them have until now been widely established as an alternative to the Harvard scale.

These scales were developed for subjective visual evaluation of patients by observers. Using the Harris scale cosmetic outcome of BCT is reported as good or excellent in a high percentage of patients. Therefore, these scales seem to have a limited discriminative power for the evaluation of cosmetic results of BCT. However, due to its widespread use and acceptance even objective methods have adopted them in their reported evaluation [38–40].

Methods of Assessment

Methods used for cosmetic evaluation of BCT can be divided into subjective and objective.

Subjective methods include patient self-assessment, assessment by an isolated observer, or a panel of observers. Objective methods involve several different types of quantifications.

For both groups of methods evaluation can be made carried out on the patient or by means of photographs (prints, slides, or digital images) (Table I).
The same author aroused interest over objective methods of evaluation by introducing the first measure for evaluating breast retraction of the treated breast, the Breast Retraction Assessment (BRA). Using a marked acrylic sheet over patients thorax, he calculated upward and inward retraction of the treated versus non-treated breast. Higher BRA values corresponded to less favorable cosmetic results. BRA was subsequently correlated with tumor size, chemotherapy and radiation fields [14]. Van Limbergen et al., using an identical methodology applied to patient photographs, come up with two new asymmetry measurements, in addition to the BRA: the Lower Breast Contour (LBC) and the Upward Nipple Retraction (UNR). A strong correlation was found between the obtained values and a subjective classification undertaken by observers [7]. The same line of thought was followed by several other authors in more recent work [4,27,29,31,42].

Noguchi introduced a sum of objective and subjective evaluations. The objective part was undertaken with a Moire’s topographic camera, sorting out 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 [30]. A similar reasoning was followed by others using both types of methods and adding the scores to obtain a final result [21,23]. The rationale for this solution resides in the fact that, for the majority of authors, neither method is successful in translating the complete evaluation of cosmetic results. When a poor cosmetics is consequent to a very visible misplaced scar, the correlation between asymmetry measurements and observer classification may be lost [7].

Semi-automatic 2D. New approaches have emerged in the last five years arising from two different European research groups. In both, work started by constructing a database of patient images, evaluated subjectively by a panel of experts, to obtain a classification, which served as a standard to compare the results of the objective methods.

Fitza et al. [56] described a breast symmetry index (BSI), to evaluate the cosmetic outcome of breast conservation. A software called Breast Analyzing Tool (BAT) was developed to measure differences between left and right breast size from patients digital pictures. The sum of all differences results in the BSI, reported as percent difference (%d) or as a difference factor (df). Both units represent differences in size between breasts. It was possible to demonstrate BAT accuracy compared with experts and additionally that BSI is reproducible when used by experts and non-experts.

The BSI index measured from frontal pictures correlates significantly with subjective evaluation by experts differentiating well between good and fair cosmetic outcome. However, no correlation was found between the BSI index and patient’s opinion. The BSI also did not differentiate well between excellent and good results, as well as between fair and poor results. The BSI thus can be used to replace subjective analyses with the advantages of increasing reproducibility and fastening evaluation time. It improves on the methods described by Pezner et al. [14] and van Limbergen et al. [7] as all quadrants are assessed in addition to nipple retraction. The BSI is also able to differentiate well between good and poor cosmetics with the benefit of requiring only one frontal picture of the patient (Fig. 6).
Cardoso et al. [57] presented another approach to automated analysis of BCT images, categorizing results as excellent, good, fair, and poor. To accomplish this objective a concise representation of a BCT image is first obtained based on asymmetry, color, and scar visibility features [58]. Asymmetry between breasts is evaluated using a large set of indices, some of them introduced for the first time. To extract color features of each breast, a histogram analysis is carried out, followed by an evaluation of dissimilarity. Scar visibility is translated into local color dissimilarity, by comparing corresponding breast sectors. A correct four class classification rate of around 70% is obtained when comparing with a consensus expert panel of evaluation [54].

A software tool was developed to incorporate the aforementioned algorithms: the Breast Cancer Conservative Treatment Cosmetic Results (BCCT.core). It introduces user independent evaluation of results preceded by automatic localization of fiducial points (nipples, breast contour, and sternum jugular notch) on digital photographs. All measures are of individual cosmetic characteristics are automatically reported and this is converted onto an overall classification of cosmetic results using the four class scale (Fig. 7). Comprehensive reports can be generated and the results stored in a database to facilitate trend and statistical analyses.

These tools, while innovative and reproducible, have several points that need to be addressed. Both only make use of a face-view of the patients. Side or oblique views can provide useful information but are difficult to standardize, so the solution may move to the 3D representation of the images. Cardoso et al. [59] compared the performance of the BCCT.core software using face-views only with that of four-views per case, and again using a consensus panel evaluation as the standard, obtained identical results. It is possible that as the evaluation of the BCCT.core software is based on more than just asymmetry measurements, it combines enough information to evaluate cosmetic outcome accurately based only on face-view images.

The BAT and the BCCT.core software have recently been compared on the same set of cases. Results show a similar performance on low quality images and a superior performance BCCT.core software on higher quality images, perhaps due to the inclusion of color and scar features [60]. One of the obvious limitations of these methods is the incapacity to evaluate breast volume.

**The 3D addition.** 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 deficit and to plan future surgeries. There are several variations of this tool, from relatively simple volumetric analyses to more sophisticated programmes that allow quantitative measurements and simulation of the post-operative outcome. Although these benefits are known, some authors still doubt the ability of current 3D methods [61].

Several research groups have recently made attempts with 3D approaches. Losken and colleagues developed an objective technique based on a 3D camera and software to quantify the cosmetic results of BCT [62–64]. This software package enables the comparison between the treated and untreated breast by analyzing the surface area and volume differences. The camera includes 12 individual digital lenses arranged in three planes with a single focal point at the manubrium. Images are captured with the patients’ arms in two different positions; along the sides and on the hips. The software is then used to determine the level of asymmetry between the breasts by overlapping the 3D meshes from the two breasts.

Eder et al. [65] conducted a study based on a 3D evaluation protocol to analyze 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 breasts surfaces. Three observers analyzed the evaluation protocol precision using two dummy models. They also evaluated the potential of this approach in clinical application by comparing it with BCCT.core on 23 breast reconstruction patients. The authors concluded that this approach could assist surgeons in the pre-operative planning and optimization of breast corrections after reconstruction and breast conservative approaches.

Recently the Kinect (Microsoft Corp., Redmond, WA) was introduced as promising low cost and easy to use equipment for BCT cosmetic evaluation [66,67]. This tool can not only facilitate automation, but also provide volumetric information. With this solution authors improved automation of prominent points and of the breast contour comparatively to those obtained with the original BCCT.core. Furthermore, other promising results were achieved, including the detection of volumetric differences of the breasts using the disparity map generated from the Kinect. It was shown that depth-map images facilitate the automation of BCCT.core and thus this software seems a 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 BCT results.

The main drawback of 3D techniques is the demand for specialized 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 cosmetic result for a more informed
choice of treatment, nor are they suitable, as isolated tools for the automatic evaluation of the cosmetic outcome after the surgery [68–71].

CONCLUSIONS

There is currently an increased interest regarding the cosmetic results in BCT due to the addition of oncoplastic techniques to the classic breast conserving techniques.

Additionally, innovations in radiotherapy with the application of partial breast irradiation have also a strong focus on cosmetic results which need to be compared with those obtained by whole breast irradiation.

There is therefore a demand for an accurate, discriminative and simple method for the evaluation of cosmetic outcome of BCT.

The fifth edition of the European Organisation for Research and Treatment of Cancer (EORTC) manual for clinical research in breast cancer published in 2005 supported the concept of adding qualitative and quantitative assessment of cosmetic outcome [72]. A subjective evaluation by a panel of at least five observers, classifying results according to the Harris scale [19] and an objective evaluation using the measurements of asymmetry described by Pezner et al. [14], with the addition of the Turesson classification for skin damage graded as the area of telangiectasias and skin necrosis [18,73]. These recommendations although aimed at including all important aspects for evaluating cosmetic results of BCT are virtually impracticable for routine practice in centers with a medium/large volume of patients. In 2012, trying to fill this gap, experts in cosmetic evaluation of BCT produced a recommendation paper on how to evaluate results, but still no standard method was defined [74].

Computer analysis of digital photographs seems to be the solution that fits most of the needed requirements. Ideally, all kinds of photographs should be amenable to evaluation, but it is possible that some minimum standards for picture quality (definition, backlight, background) will be demanded if one is to expect discriminative power in the evaluation of cosmetic results.

Recent initiatives in computerized analysis of BCT images have provided promising results and, it is possible that further developments, like the inclusion of volumetric information using low cost solutions, will lead to a generalizable method of objective evaluation that could be used by all those interested in achieving good cosmetic outcomes.

REFERENCES

27. Sneeuw KC, Aaronsen NK, Yarnold JR, et al.: Cosmetic and functional outcomes of breast conserving treatment for early stage...


