

Aesthetics in Breast Conserving Therapy: Do Objectively Measured Results Match Patients' Evaluations?

Joerg Heil, MD¹, Julia Dahlkamp¹, Michael Golatta, MD¹, Joachim Rom, MD¹, Christoph Domschke, MD¹, Geraldine Rauch, PhD², Maria Joao Cardoso, MD, PhD³, and Christof Sohn, MD¹

¹University Breast Center Heidelberg, Heidelberg, Germany; ²Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany; ³Porto Faculty of Medicine, Breast Research Group—INESC Porto, Porto, Portugal

ABSTRACT

Background. To analyze the relationship of objective and subjective evaluation tools of breast aesthetics, we compare the results of the BCCT.core (breast cancer conservative treatment.cosmetic results) software, a semiautomated objective symmetry evaluation tool, with those of the Aesthetic Status of the BCTOS (Breast Cancer Treatment Outcome Scale) patient questionnaire.

Materials and Methods. We included 128 patients with one-sided, primary breast cancer, treated conservatively in a prospective, exploratory study in order to assess the inter-rater reliability of the BCCT.core and the agreement between the BCCT.core and the BCTOS preoperatively, shortly and 1 year after surgery. Therefore, we use agreement rates, multiple (mk), and weighted (wk) kappa coefficients as statistical methods. Furthermore, we analyzed patient-, tumor-, and therapy-related variables as possible covariates to explain agreement.

Results. The inter-rater reliability for the semiautomated BCCT.core is very good with agreement rates up to 84% (mk = 0.80). The agreement rates of the BCCT.core and the BCTOS Aesthetic Status range between 35 and 44% subject to the different times of assessment (wk = 0.34 at best). Moreover, the patients judge their aesthetic outcome more positively than the software. None of the considered patient-, tumor-, and therapy-related covariates turned out to explain agreement.

Conclusion. The BCCT.core is a reliable instrument that shows fair agreement with patient's perspective.

There are 2 ways to evaluate aesthetics in breast cancer surgery: subjective evaluation tools (either the patient or a third person such as the surgeon assesses the outcome) and objective ones (try to measure aesthetic results objectively). This use of very different measurement approaches makes it difficult to compare results of various surgical procedures in particular in oncoplastic surgery.¹

Subjective possibilities to describe aesthetic results range from single questions of overall assessment to questionnaires that try to differentiate between different aspects of breast aesthetics such as size, color, shape, position, and shape of the nipple-areolar complex and scars.^{2–5} The Breast Cancer Treatment Outcomes Scale (BCTOS) is a clearly structured and valid questionnaire addressing aesthetic issues.^{5,6} The Aesthetic Status Scale consists of 7 items.

Objective approaches vary considerably in focus and complexity, for example, the evaluation of nipple retraction, breast contour and the nipple asymmetry measurements, the quantitative measurement of breast asymmetry, a so-called “Breast Compliance Evaluation,” and the evaluation of ptosis.^{7–11} The semiautomated BCCT.core (breast cancer conservative treatment.cosmetic results) software combines many of these objective approaches. During its development, an expert panel rating validated this software; additionally, it would be desirable that this objective tool, the BCCT.core software, could reproduce a comprehensive and rationally guided evaluation of the patients themselves.¹²

Before addressing this topic, we assess the inter-rater reliability of the BCCT.core as the software is only semi-automated. Following, we present a first comparison of the

BCCT.core results with a patient's perspective on breast aesthetics. Moreover, we evaluate time of assessment, patient-, tumor-, and therapy-related variables as possible covariates explaining agreement.

METHODS

The Study Instruments: BCTOS and BCCT.core

The BCTOS (Breast Cancer Treatment Outcome Scale) The BCTOS was designed to assess women's subjective evaluation of the aesthetic and functional outcome after breast cancer treatment.⁵ The 22 items lead to the subscales: Aesthetic Status, Functional Status and Breast Sensitivity Status. The Aesthetic Status subscale of the BCTOS consists of 7 items.⁶ These are: breast size, nipple appearance, breast shape, elevation, scar tissue, fit of bra, and clothing. The patient rates these items according to symmetry between treated and untreated breast on a 4-point Likert scale (1 = no difference between treated and untreated breast to 4 = large difference between treated and untreated breast). The rounded mean of all 7 items is defined as the Aesthetic Status.

The BCCT.core (Breast Cancer Conservative Treatment. Cosmetic Result) Software The BCCT.core software summarizes all objective symmetry measurements available in a single software.^{12–14} The software carries out a semiautomatic analysis of torso photographs in an

anterior–posterior view. Predetermined points are designated by the examiner, followed by automated software calculations of different relations/asymmetries measures including breast volume, skin color, and scars (see Fig. 1). An algorithm combines all the measurement results in an overall score of aesthetic outcome. The results of the software analysis have been validated by an expert panel rating.^{12,14} The results are displayed on a 4-point scale (1 = excellent, 2 = good, 3 = fair, 4 = poor).

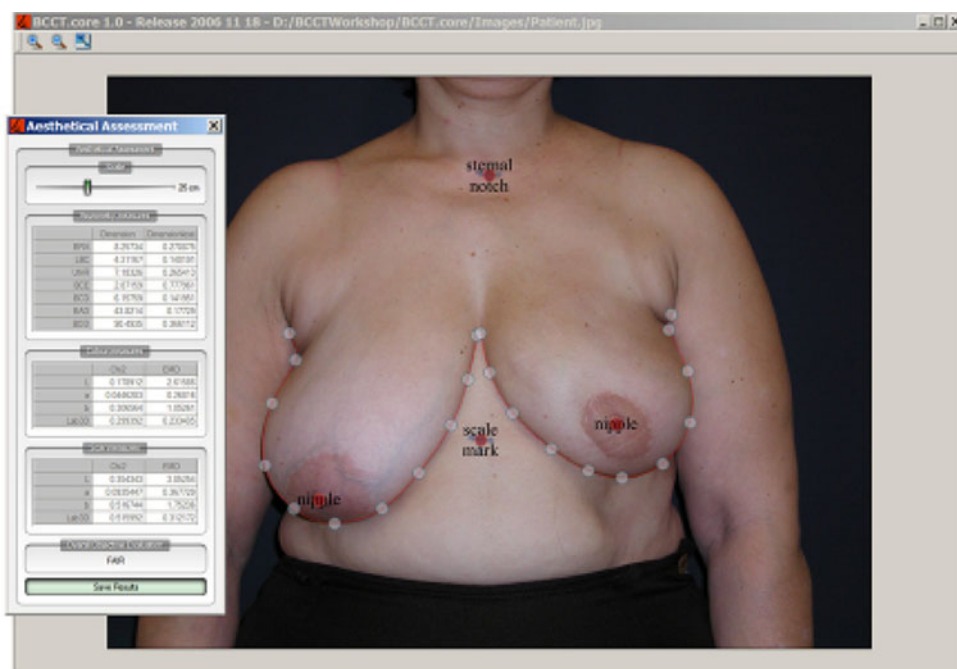
Influencing Factors on Agreement

The knowledge of influencing patient-, tumor-, and therapy-related factors on agreement could at best promote the understanding of aesthetic results evaluated by patients and the software. Moreover, it could guide further development and modification of objective methods to measure aesthetic results. We documented age, weight, height, breast size, tumor size, localization of scars, weight of the specimen, and number of re-excisions as possibly relevant covariates guided by the existing literature and our clinical experience.^{12,15–17}

Course of the Study and Description of the Study Population

This analysis is designed as an explorative, single-institution, prospective cohort study to evaluate aesthetic and functional outcome parameters after breast conserving

FIG. 1 Screenshot of BCCT.core software¹²



therapy. The project was approved by the ethics commission of the University of Heidelberg Medical School. All patients gave their written informed consent to participate.

We included patients who met the following criteria: date of surgery between October 1, 2007 and September 30, 2008; breast conserving therapy in primary, unilateral, histologically proven breast cancer; and existing photographs and BCTOS questionnaires before, shortly after, and 1 year after surgery.

Of 276 preoperatively screened patients, 250 patients gave informed consent to participate in the study. Shortly after surgery, 46 of these patients dropped out, 25 because of secondary mastectomy, and 21 for individual reasons (patients dismissed from hospital without being photographically documented or withdrew their informed consent). Of these 204 patients, 128 (62%) attended the 1-year follow-up control.

At mean the population was 59 years old (standard deviation [SD] = 9 years, 95% confidence interval [95% CI] = 40–74 years) with a body mass index of 26 (SD = 4, 95% CI = 18–37), the majority of patients were postmenopausal, had no family history of breast or ovarian cancer, a breast size of cup A/B, a tumor size less than 2 cm, a segmental resection of breast tissue, and a sentinel node biopsy. Most of the patients were diagnosed with infiltrating ductal carcinoma (see Table 1).

Statistical Methods

We analyze the inter-rater reliability for the BCCT.core software by calculating agreement rates and multiple kappa statistics (mk) for 4 different semiprofessional raters, that is, medical students working as research assistants in this project. Similarly, we use agreement rates and weighted kappa statistics to quantify the agreement between the Aesthetic Status of the BCTOS with the BCCT.core software result for all 3 times of assessment.^{18,19} Moreover, we perform the same analysis for the 7 single items of the BCTOS defining the Aesthetic Status. To interpret the kappa coefficients, we use the following definition: equal to 0 indicates poor agreement; 0.01–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; 0.81–0.99 almost perfect agreement; and 1.00 perfect agreement.²⁰

Additionally, we investigate whether the agreement group (no difference between BCTOS and BCCT.core result at the 1-year follow-up) differs from the nonagreement group with respect to patient-, tumor-, and therapy-related covariates. Depending on the scale level, we compare the groups by a 2-sample *t* test or a nonparametric Mann–Whitney test, respectively.

TABLE 1 Description of the study cohort (*n* = 128)

Variable	No.	Percentage (%)
Breast size = bra size		
Cup A	16	12.7
Cup B	56	44.4
Cup C	46	34.9
Cup D–G	10	7.9
Tumor histology		
Ductal carcinoma in situ	22	17.5
Invasive ductal carcinoma	82	65.1
Invasive lobular carcinoma	12	9.5
Others	10	7.9
Localization of tumor		
Upper outer quadrant	80	63.5
Lower outer quadrant	20	15.9
Upper inner quadrant	18	14.3
Others	10	6.4
Pathological T stage		
Tis	22	17.5
pT1a–c	86	68.2
pT2–3	20	14.3
Pathological N stage		
pN0	92	73
pN+	26	19.1
No lymphadenectomy performed	10	7.9
Method of resection of breast cancer		
Segmental resection	100	79.4
Quadrant resection	28	20.6
Performance of an intraoperative second-look resection	45	26.3
Lymphadenectomy		
Sentinel node biopsy	96	76.2
Classical axillary lymphadenectomy	18	14.3
None	14	9.5

RESULTS

Inter-rater Reliability of the BCCT.core Software Results

Independent of the time of assessment, there is a very good agreement between 4 semiprofessional users of the semiautomated software. Complete agreement (no difference between any of the raters) occurs in about 80% at all time points as shown in Table 2. The multiple kappa coefficients ranging from 0.75 before surgery to 0.82 at 1 year after surgery confirm these results.

TABLE 2 Agreement of 4 raters using the BCCT.core software

Time of assessment	Complete agreement (no difference between any of the raters)	Multiple kappa coefficients
Before surgery	101 (79%)	0.75
Shortly after surgery	107 (84%)	0.81
1 year after surgery	105 (82%)	0.82

TABLE 3 Agreement of the BCCT.core software result and BCTOS Aesthetic Status depending on time of assessment

Time of assessment	Complete agreement	Weighted kappa coefficient
Before surgery (<i>n</i> = 18)	41 (35%)	0.04
Shortly after surgery (<i>n</i> = 104)	46 (44%)	0.34
1 year after surgery (<i>n</i> = 116)	52 (44%)	0.19
Over all times (<i>n</i> = 330)	139 (42%)	0.21

Agreement of BCCT.core Software Results and BCTOS Aesthetic Status

The absolute rate of agreement between the BCCT.core and BCTOS results is between 35 and 44% for all times of assessment corresponding to weighted kappa coefficients of 0.04 to 0.34 stating a slight to fair agreement (see Table 3).

Table 4 displays the corresponding agreement tables for the BCTOS Aesthetic Status and the BCCT.core software scores. We observe that patients' perspective on aesthetic outcome is generally more positive than the scoring of the software for all times of assessment. For example, the majority of patients rate the Aesthetic Status preoperatively as "excellent" whereas the software evaluation only states a "good." Analyzing all 3 times of assessment in detail, a general downward trend in the scores can be observed.

With respect to the 7 single items of the BCTOS the best overall consensus rate is 57% (*wk* = 0.28) for breast elevation, followed by breast shape and breast size at the 1-year after surgery follow-up. Scars, fit of bra, and clothing did not seem to be closely related to BCCT.core results.

Explaining Covariates of Agreement between BCCT.core Result and BCTOS Aesthetic Status

To assess the influence of patient-, tumor-, and therapy-related variables on agreement, we compared the agreement group (no difference between BCTOS and BCCT.core result) and the nonagreement group (with a difference between BCTOS and BCCT.core result) with

TABLE 4 Agreement between BCCT.core result and BCTOS Aesthetic Status according to time of assessment

Time	BCCT.core result ^a	BCTOS Aesthetic Status ^b			
		1	2	3	4
Before surgery	1	25	5	0	0
	2	64	16	1	0
	3	5	2	0	0
	4	0	0	0	0
Shortly after surgery	1	1	6	1	0
	2	27	40	4	3
	3	6	10	5	0
	4	0	1	0	0
1 year after surgery	1	4	1	0	0
	2	20	37	6	3
	3	7	27	11	0
	4	1	2	0	0
Overall	1	30	12	1	0
	2	101	93	11	6
	3	18	39	16	0
	4	1	3	0	0

^a 1 = very good ... 4 = poor aesthetic result

^b 1 = no difference ... 4 = large difference between treated and untreated breast

respect to age, weight, height, breast size, tumor size, localization of scars, weight of the specimen, and number of re-excisions, but could not identify a significant influence on agreement of aesthetic results in 1 of these variables.

DISCUSSION

The BCCT.core software is an objective instrument to measure aesthetic outcome after breast conserving therapy. We think that it is the most comprehensive approach to measure breast aesthetics, as the software integrates a lot of different aspects and dimensions as discussed in the literature.⁷⁻¹¹ Although it is only semiautomated, the inter-rater reliability is very good, even if used by semiprofessional raters.^{12,21} However, the software results show only slight to fair agreement with the patients' perspective of aesthetics, measured by the BCTOS Aesthetic Status. The lack of agreement between the 2 instruments can be explained by several considerations.

First of all, a subjective evaluation of aesthetics is always critical to judge, as the subjective impression can be influenced by many different, mostly unknown, factors, which are therefore difficult to control. The intra-patient reliability of the BCTOS in a defined period of time with respect to an intervention such as a breast conserving surgery has not been tested yet. Hence, it might be possible

that the patient's perspective on aesthetics is not well reproducible when the BCTOS questionnaire is repeatedly answered.

The analysis shows that, in general, patients evaluate their individual aesthetic results more positively than the software did. The fact that former studies have been fairly consistent in noting that patients tend to evaluate the aesthetic outcomes more positively than health care providers could explain that our agreement rates are even lower than those of the validation study using an expert panel as "golden standard."^{16,22}

Nonetheless, both instruments, the BCTOS and the BCCT.core, address similar aspects that are relevant to an aesthetical perception of the female breast. Those dimensions may be summarized by asymmetry of size and shape, color differences, and appearance of scars.

We therefore recommend using and reporting both subjective and objective instruments to measure breast aesthetics in breast surgery even if one knows about the differences with respect to the results. The importance of patient-reported aesthetic outcome measures is obvious, but probably influenced by many partly unknown factors, not only by surgical techniques and clinical situation. That is why it seems to be important to improve and validate objective methods to compare aesthetic results of different surgical and radiotherapeutical procedures in comparable clinical situations with less confounding variables.

In conclusion, our analysis shows that objective measurements and patients' perspectives evaluate similar dimensions differently; even when applying 2 comprehensive, rationally guided, and multifactorial instruments. Therefore, it seems to be necessary to apply both approaches in order to gain a more comprehensive knowledge of breast aesthetics.

REFERENCES

1. Rew DA. Towards a scientific basis for oncoplastic breast surgery. *Eur J Surg Oncol*. 2003;29:105–6.
2. Arenas M, Sabater S, Hernandez V, Henriquez I, Ameijide A, Anglada L, et al. Cosmetic outcome of breast conservative treatment for early stage breast cancer. *Clin Transl Oncol*. 2006;8:334–8.
3. Vrieling C, Collette L, Fourquet A, Hoogenraad WJ, Horiot JC, Jager JJ, et al. The influence of the boost in breast-conserving therapy on cosmetic outcome in the EORTC "boost versus no boost" trial. EORTC Radiotherapy and Breast Cancer Cooperative Groups. European Organization for Research and Treatment of Cancer. *Int J Radiat Oncol Biol Phys*. 1999;45:677–85.
4. Sneeuw KC, Aaronson NK, Yarnold JR, Broderick M, Regan J, Ross G, et al. Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 2. Relationship with psychosocial functioning. *Radiother Oncol*. 1992;25:160–6.
5. Stanton AL, Krishnan L, Collins CA, Form or function? Part 1. Subjective cosmetic and functional correlates of quality of life in women treated with breast-conserving surgical procedures and radiotherapy. *Cancer*. 2001;91:2273–81.
6. Heil J, Holl S, Golatta M, Rauch G, Rom J, Marmé F, et al. Aesthetic and functional results after breast conserving surgery as correlates of quality of life measured by a German version of the Breast Cancer Treatment Outcome Scale (BCTOS). *Breast*. 2010 [Epub ahead of print].
7. Pezner RD, Patterson MP, Hill LR, Vora N, Desai KR, Archambeau JO, et al. Breast retraction assessment: an objective evaluation of cosmetic results of patients treated conservatively for breast cancer. *Int J Radiat Oncol Biol Phys*. 1985;11:575–8.
8. Van Limbergen E, van der Schueren E, Van Tongelen K. Cosmetic evaluation of breast conserving treatment for mammary cancer. 1. Proposal of a quantitative scoring system. *Radiother Oncol*. 1989;16:159–67.
9. Sacchini V, Luini A, Tana S, Lozza L, Galimberti V, Merson M, et al. Quantitative and qualitative cosmetic evaluation after conservative treatment for breast cancer. *Eur J Cancer*. 1991;27:1395–400.
10. Tsoukas LI, Fentiman IS. Breast compliance: a new method for evaluation of cosmetic outcome after conservative treatment of early breast cancer. *Breast Cancer Res Treat*. 1990;15:185–90.
11. Kim MS, Reece GP, Beahm EK, Miller MJ, Neely Atkinson E, et al. Objective assessment of aesthetic outcomes of breast cancer treatment: measuring ptosis from clinical photographs. *Comput Biol Med*. 2007;37:49–59.
12. Cardoso MJ, Cardoso J, Amaral N, Azevedo I, Barreau L, Bernardo M, et al. Turning subjective into objective: the BCCT.core software for evaluation of cosmetic results in breast cancer conservative treatment. *Breast*. 2007;16:456–61.
13. Cardoso JS, Pinto da Costa JF, Cardoso MJ. Modelling ordinal relations with SVMs: an application to objective aesthetic evaluation of breast cancer conservative treatment. *Neural Netw*. 2005;18:808–17.
14. Cardoso MJ, Cardoso J, Santos AC, Barros H, Cardoso de Oliveira M. Interobserver agreement and consensus over the esthetic evaluation of conservative treatment for breast cancer. *Breast*. 2006;15:52–7.
15. Clarke D, Martinez A, Cox RS. Analysis of cosmetic results and complications in patients with stage I and II breast cancer treated by biopsy and irradiation. *Int J Radiat Oncol Biol Phys*. 1983;9:1807–13.
16. Al-Ghazal SK, Fallowfield L, Blamey RW. Patient evaluation of cosmetic outcome after conserving surgery for treatment of primary breast cancer. *Eur J Surg Oncol*. 1999;25:344–6.
17. Taylor ME, Perez CA, Halverson KJ, Kuske RR, Philpott GW, Garcia DM, et al. Factors influencing cosmetic results after conservation therapy for breast cancer. *Int J Radiat Oncol Biol Phys*. 1995;31:753–64.
18. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas*. 1973;33:613–9.
19. Cicchetti D, Allison T. A new procedure for assessing reliability of scoring eeg sleep recordings. *Am J EEG Technol*. 1971;11:101–9.
20. Seigel DG, Podgor MJ, Remaley NA. Acceptable values of kappa for comparison of two groups. *Am J Epidemiol*. 1992;135:571–8.
21. Cardoso MJ, Cardoso JS, Wild T, Krois W, Fitzal F. Comparing two objective methods for the aesthetic evaluation of breast cancer conservative treatment. *Breast Cancer Res Treat*. 2009;116:149–52.
22. Al-Ghazal SK, Blamey RW. Cosmetic assessment of breast-conserving surgery for primary breast cancer. *Breast*. 1999;8:162–8.