

CHOOSING OBSERVERS FOR EVALUATION OF AESTHETIC RESULTS IN BREAST CANCER CONSERVATIVE TREATMENT

MARIA JOÃO CARDOSO, M.D.,* ANA CRISTINA SANTOS, M.P.H.,‡ JAIME CARDOSO, PH.D. ST.,†
HENRIQUE BARROS, M.D., PH.D.,‡ AND MANUEL CARDOSO DE OLIVEIRA, M.D., PH.D.*

*Surgery B Department, S. João Hospital, Porto Medical School, Porto, Portugal; †Hygiene and Epidemiology Department, Porto Medical School, Porto, Portugal; ‡Telecommunications and Multimedia Unit, INESC, Porto, Portugal

Purpose: The subjective evaluation of aesthetic results in conservative breast cancer treatment has largely been used without questioning the observer's skills. The aim of this study was to evaluate interobserver agreement of the aesthetic results of breast cancer conservative treatment in three groups of observers with different levels of experience.

Methods and Materials: Photographs were taken of 55 women who had undergone conservative unilateral breast cancer treatment and 5 control women with no breast disease. The images were then distributed to 13 observers who were divided into three groups according to their experience in breast cancer treatment: experienced, medium experienced, and inexperienced. They were first asked to distinguish the patients from the controls and for the patients to identify the operated side. Subsequently, they were asked to classify the aesthetic result as excellent, good, fair, or poor. The accuracy in identifying controls, patients, and side of treatment was calculated individually for all observers. The interobserver agreement for the aesthetic result was calculated using observed agreement and multiple κ statistic (κ) in each of the three groups.

Results: Inexperienced observers performed significantly worse than experienced observers in identifying controls, patients, and the side of treatment. Agreement of the aesthetic result was significantly greater in the group of experienced observers ($\kappa = 0.59$) than in the medium experienced ($\kappa = 0.35$) and inexperienced ($\kappa = 0.33$) observers.

Conclusion: Previous experience in breast cancer conservative treatment should be considered a prerequisite for the evaluation of the aesthetic results. © 2005 Elsevier Inc.

Breast cancer, Conservative treatment, Aesthetic result, Observers, Subjective evaluation.

INTRODUCTION

Conservative treatment of breast cancer has a proven value for local disease control, as reported in randomized trials (1, 2). The quality of the aesthetic results remains an additional goal that still has to be formally evaluated (3). The methods of evaluating breast cancer conservative treatment are traditionally considered to be subjective (4–7) or objective (8, 9). For subjective evaluations, observers are usually selected from the local medical staff for practical reasons, and the criteria used for this selection frequently are not described (10–12). It is possible that previous experience in breast cancer conservative treatment will influence the evaluation of the aesthetic results, because this may enable the distinction between the general aesthetic features valued by society and those related to the surgical procedure itself. To address this issue, we asked three groups of observers with

different levels of experience to evaluate the aesthetic results of patients who had undergone unilateral breast cancer conservative treatment to assess the interobserver agreement in each group.

METHODS AND MATERIALS

Photographs were taken of 55 breast cancer patients who had undergone unilateral conservative breast surgery and radiotherapy. The treatment had been completed at least 1 year before the onset of the study. All patients provided written informed consent to participate. A digital camera with a resolution of 4 mega pixels was used, having as background a blue panel of 0, 128, 255. Photographs were taken in four positions with the patient standing on floor marks: face, arms down; face, arms up; left side, arms up; and right side, arms up. Photographs of 5 healthy women were added to the group. The series of 60 subjects was distributed

Reprint requests to: Maria João Cardoso, M.D., Serviço de Cirurgia B, H.S. João, Faculdade de Medicina do Porto, Alameda do Prof. Hernâni Monteiro, Porto 4200-319, Portugal. Tel. and Fax: (+351) 22-5505589; E-mail: mjcard@mail.med.up.pt

Acknowledgments—We thank Prof. Ayres-de-Campos for his help in the preparation of the manuscript and the following for

their participation as observers in this study: Drs. Matos Lima, Pinto-Sousa, António Gouveia, John Preto, Susy Costa, Sandra Ferreira, Eduardo Costa, Francisco Santos, Carla Morgado, Ricardo Seromenho, Susana Silva, and Sandra Sousa.

Received May 11, 2004. Accepted for publication Jun 21, 2004.

individually to 13 observers, who had been divided into three groups according to their level of experience in breast cancer treatment: 4 experienced observers (surgeons with previous experience in breast cancer conservative treatment); 4 medium experienced observers (general surgery residents with at least 3 years of experience and having observed breast cancer conservative treatment surgery); and 5 inexperienced observers (individuals with degrees in nonmedical areas and no previous contact with the disease). None of the observers knew, or was in anyway connected to, the subjects studied. For each case, the observers were asked to classify it as a breast cancer or control case and, in the former, to identify the side that had undergone surgery. Subsequently, they were asked to evaluate the aesthetic result, classifying each in one of four categories according to the Harris classification (4): excellent, treated breast nearly identical to untreated breast; good, treated breast slightly different from untreated breast; fair, treated breast clearly different from untreated breast but not seriously distorted; and poor, treated breast seriously distorted.

The accuracy was calculated for each of the 13 observers regarding their capacity to identify controls, treated cases, and side of treatment. The observed agreement and multiple κ statistic were used to calculate the interobserver agreement in each of the three groups of observers.

RESULTS

In the identification of the treated cases, the accuracy was >95% in all experienced observers, in all but one of the medium experienced observers, and in only one of the inexperienced group (Table 1). The proportion varied between 96.7% and 98.3% in the experienced group, between 83.3% and 100% in the medium experienced group, and between 83.3% and 96.7% in the inexperienced group. Similar results were obtained for the identification of the treated side (Table 1).

The overall observed agreement and multiple κ value in the evaluation of the aesthetic results by all 13 observers was 0.51 and 0.33, respectively. The observed agreement and κ value were significantly greater for experienced observers (0.71 and 0.59, respectively) than for the medium experienced (0.52 and 0.35, respectively) and inexperienced (0.51 and 0.33, respectively) observers (Table 2).

DISCUSSION

The subjective evaluation of the aesthetic results in breast cancer patients after conservative treatment is usually reported as performed globally by individuals with varying levels of expertise (13–16). Very few studies have compared the agreement of the aesthetic results in groups of observers with different levels of experience (3, 17).

Christie *et al.* (3), in a study of 47 conservative breast cancer patients, evaluated the interobserver agreement in the aesthetic assessment as excellent, good, fair, and poor among a panel of 5 observers (2 trained and 3 untrained). In the group of more experienced observers, the percentage of cases reaching absolute agreement was greater (49% vs. 19% in the inexperienced group).

Table 1. Accuracy

Observers	Accuracy (%)	
	Patient/control	Treated side
Experienced (surgeons)		
1	98.3 (92.1–99.9)	100 (94.6–100.0)
2	98.3 (92.1–99.9)	100 (94.6–100.0)
3	98.3 (92.1–99.9)	100 (94.6–100.0)
4	96.7 (89.4–99.4)	100 (94.6–100.0)
Medium experienced (general surgery residents)		
5	100 (95.1–100.0)	100 (94.6–100.0)
6	83.3 (72.3–91.2)	80.8 (67.8–90.2)
7	98.3 (92.1–99.9)	100 (94.6–100.0)
8	96.7 (89.4–99.4)	100 (94.6–100.0)
Inexperienced (from other areas)		
9	86.7 (76.2–93.6)	93.8 (83.9–98.4)
10	83.3 (72.3–91.2)	95.6 (86.1–99.2)
11	86.7 (76.2–93.6)	93.6 (83.6–98.4)
12	93.3 (82.5–96.9)	98.1 (91.0–99.9)
13	96.7 (89.4–99.4)	96.4 (91.4–99.9)

Data in parentheses are 95% confidence intervals.

Pezner *et al.* (17) analyzed the cosmetic evaluation by 44 observers using 14 projected color photographs of breast cancer patients after conservative treatment. Eight observers were considered experienced because they had performed >20 breast cancer conservative treatments, and their answers were analyzed independently. Although a formal evaluation of agreement was not performed, the authors noted that this group of experts made fewer errors in identifying the operated side and reached a consensus more often than the other observers.

The results of our study suggest that observers with

Table 2. Interobserver agreement

Observers	Observed agreement (%)	Multiple κ value
All 13 observers	51	0.33
Experienced (surgeons)	71	0.59
Medium experienced (general surgery residents)	52	0.35
Inexperienced (from other areas)	51	0.33

experience in breast cancer conservative treatment obtain greater levels of agreement compared with individuals without this experience. The latter have difficulties in distinguishing treated from nontreated breasts and in identifying the treated side correctly. It could be that basic female characteristics such as age, weight, and body fitness are excessively valued by this group instead of the symmetry of the breasts and acceptability of the surgical result, the issues that are really at stake in this

evaluation. Probably as a result of the lack of consistency in the assessment of all these parameters, the interobserver agreement was low.

The choice of a homogeneous group of observers with experience in breast cancer conservative treatment needs to be a prerequisite for the aesthetic evaluation of this form of treatment and will provide better interobserver agreement than a mixed group involving clinicians with different levels of expertise.

REFERENCES

1. Fisher B, Anderson S, Bryant J, *et al.* Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med* 2002;347:1233–1241.
2. Veronesi U, Cascinelli N, Mariani L, *et al.* Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *N Engl J Med* 2002;347:1227–1232.
3. Christie D, O'Brien M, Christie J, *et al.* A comparison of methods of cosmetic assessment in breast conservation treatment. *Breast* 1996;5:358–367.
4. Harris JR, Levene MB, Svensson G, *et al.* Analysis of cosmetic results following primary radiation therapy for stages I and II carcinoma of the breast. *Int J Radiat Oncol Biol Phys* 1979;5:257–261.
5. Beadle GF, Come S, Henderson IC, *et al.* The effect of adjuvant chemotherapy on the cosmetic results after primary radiation treatment for early stage breast cancer. *Int J Radiat Oncol Biol Phys* 1984;10:2131–2137.
6. Liljegren G, Holmberg L, Westman G, for the Uppsala-Orebro Breast Cancer Study Group. The cosmetic outcome in early breast cancer treated with sector resection with or without radiotherapy. *Eur J Cancer* 1993;29A:2083–2089.
7. Cetintas S, Ozkan L, Kurt M, *et al.* Factors influencing cosmetic results after breast conserving management (Turkish experience). *Breast* 2002;11:72–80.
8. Pezner RD, Patterson MP, Hill LR, *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;11:575–578.
9. Krishnan L, Stanton AL, Collins CA, *et al.* Form or function? Part 2. Objective cosmetic and functional correlates of quality of life in women treated with breast-conserving surgical procedures and radiotherapy. *Cancer* 2001;91:2282–2287.
10. Wazer DE, Kramer B, Schmid C, *et al.* Factors determining outcome in patients treated with interstitial implantation as a radiation boost for breast conservation therapy. *Int J Radiat Oncol Biol Phys* 1997;39:381–393.
11. Fowble BL, Solin LJ, Schultz DJ, *et al.* Ten year results of conservative surgery and irradiation for stage I and II breast cancer. *Int J Radiat Oncol Biol Phys* 1991;21:269–277.
12. Krishnan L, Jewell WR, Mansfield CM, *et al.* Cosmetic results in early breast cancer treated with lumpectomy, peri-operative interstitial irradiation, and external beam radiation. *Int J Radiat Oncol Biol Phys* 1988;15:205–211.
13. Al-Ghazal SK, Blamey RW, Stewart J, *et al.* The cosmetic outcome in early breast cancer treated with breast conservation. *Eur J Surg Oncol* 1999;25:566–570.
14. Sacchini V, Luini A, Tana S, *et al.* Quantitative and qualitative cosmetic evaluation after conservative treatment for breast cancer. *Eur J Cancer* 1991;27:1395–1400.
15. Sneeuw KC, Aaronson NK, Yarnold JR, *et al.* Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 1. Comparison of patients' ratings, observers' ratings, and objective assessments. *Radiother Oncol* 1992;25:153–159.
16. Vrieling C, Collette L, Bartelink E. *et al.*, for the EORTC Radiotherapy and Breast Cancer Cooperative Groups. Validation of the methods of cosmetic assessment after breast-conserving therapy in the EORTC "boost versus no boost" trial. *Int J Radiat Oncol Biol Phys* 1999;45:667–676.
17. Pezner RD, Lipsett JA, Vora NL, *et al.* Limited usefulness of observer-based cosmesis scales employed to evaluate patients treated conservatively for breast cancer. *Int J Radiat Oncol Biol Phys* 1985;11:1117–1119.