Patient Reported Outcome and Quality of Life After Delayed Breast Reconstruction - An RCT Comparing Different Reconstructive Methods in Radiated and Non-radiated Patients

Fredrik Brorson,1,2 Anna Elander,1,2 Andri Thorarinsson,3 Emma Hansson1,2

Abstract

This study compares different techniques in radiated and non-radiated patients, respectively, as regards health-related quality of life. The radiated patients were randomized to latissimus dorsi flap or DIEP flap and non-radiated patients to thoracodorsal flap or expander/implant. There was a clear improvement in quality of life in all groups; although, no distinct differences could be seen for different methods.

Background: Health-related quality of life (HRQoL) is one of the core outcomes for breast reconstruction. The aim of this study was to evaluate whether the method of delayed breast reconstruction affects long-term HRQoL. Methods: Participants were divided into 2 arms depending on previous radiotherapy, and subsequently randomized between 2 methods of breast reconstruction: a latissimus dorsi flap or a deep inferior epigastric artery perforator flap in the radiated arm and a thoracodorsal flap and implant or an expander in the non-radiated arm. Validated HRQoL instruments were used: BREAST-Q to evaluate breast specific HRQoL and satisfaction, RAND-36 and EQ-5D to evaluate generic HRQoL, and BDI-21 to measure symptoms of depression and anxiety. Results: During the recruitment period (2009-2015), 233 patients were randomized. After opt-outs and exclusions, the remaining 107 participants comprise the study sample. Postoperative HrQoL was measured on average 7to 8years post-operatively. Response rates varied between 60 and 82 per cent. The BREAST-Q scores were higher after the reconstruction than before for the great majority of domains in both arms; albeit statistically significant only between the 2 methods for physical well-being chest in the radiated arm. Most participants in both arms had minimal or mild depression both before and after the operation. Conclusion: No distinct differences in long-term HrQoL could be seen for different methods There was a clear improvement in HrQoL compared to pre-reconstruction in all groups, but the effect of specific reconstructive methods on scores could not be reliably demonstrated.

Introduction

A breast reconstruction is performed to improve the patient’s quality of life and satisfaction with her breasts. Quality of life, women’s cosmetic satisfaction, normality, self-esteem, emotional well-being, and physical well-being are part of the core outcome set for breast reconstruction.1 Nonetheless, there is no consensus regarding how these outcomes should be evaluated. There are few studies providing high-quality evidence supporting the use of different techniques for breast reconstruction and it is well-known that it is challenging to conduct randomized controlled trials (RCTs) in breast reconstruction, mainly due to patients’ and surgeons’ preferences leading to recruitment difficulties.2 Consequently, there are

Keywords: PROM, DIEP, Latissimus dorsi flap, Expander, Implant, Thoracodorsal flap, Randomized controlled trial
few RCTs comparing quality of life in different breast reconstruction techniques.\textsuperscript{5,6} Quality of life can be defined in different ways and can allude to health, quality of life, and health-related quality of life (HRQoL) and there is no golden standard to measure it.\textsuperscript{6} In practice, HRQoL is measured using generic patient reported outcome measures (PROMs), which can be used in any patient independent of health condition, and disease- or condition-specific PROMs, which measure symptoms of relevance in a certain disease. Generic instruments can be used to compare patients with different conditions to each other but might be too general to capture some of the problems particular patient groups have and therefore a combination of the two are often used.\textsuperscript{7} Several of the most widely used generic instruments have been used previously to evaluate breast reconstruction: the Medical Outcomes Study 36-Item Short Form (SF-36) Health Survey/RAND 36-Item Short Form Health Survey,\textsuperscript{8,9} Health Utilities Index (HUI),\textsuperscript{10} EuroQol Instrument (EQ-5D),\textsuperscript{11} and Patient-Reported Outcomes Measurement Information System (PROMIS).\textsuperscript{12} Moreover, symptom specific instruments have been used previously, such as the Beck Depression Inventory (BDI) and Hospital Anxiety Depression Scale (HADS).\textsuperscript{13} There are 3 validated and reliability tested breast reconstruction specific PROMs\textsuperscript{14}: BREAST-Q, BRECON-31, and EORTC QLQ-BRECON-23, of which BREAST-Q is the most widely used.\textsuperscript{15} They all contain items on aspects of the core outcome set for breast reconstruction: cosmetic satisfaction, normality, self-esteem, emotional well-being, and physical well-being related to the breast(s).

The aim of this study was to evaluate whether the method of delayed breast reconstruction affects long-term HRQoL in recurrence free women treated for breast cancer with unilateral mastectomy. Participants were divided into 2 arms dependent on previous radiotherapy, and subsequently randomized between 2 methods of breast reconstruction: a latissimus dorsi flap (LD) or a deep inferior epigastric artery perforator (DIEP) flap in the radiated arm and a thoracodorsal flap and implant (TD) or expander (EXP) in the non-radiated arm. Validated instruments were used: BREAST-Q to evaluate breast specific HRQoL and satisfaction, RAND-36 and EQ-5D to evaluate generic HRQoL and BDI-21 to measure symptoms of depression and anxiety.

**Patients and Methods**

**Study Design, Protocol, and Ethics**

This study is a clinical randomized prospective trial with 2 arms: non-radiated and radiated patients. The Go Breast Prospective study protocol has been published on ClinicalTrials.Gov (identifier NCT03963427). The Regional Ethical Committee of Gothenburg reviewed and approved the study (043-08). Procedures followed were in accordance with the Helsinki Declaration of 1964, as revised, and the Good Clinical Practice (GCP) guidelines. Participants gave their written informed consent to participate in the study and to publication of the results.

**Participants and Randomization**

Participants, recruitment, and sample size calculations have been described previously.\textsuperscript{16} In brief, all women aged $>18$ years with a unilateral mastectomy defect referred to our department for delayed breast reconstruction were assessed for inclusion. Women who smoked, had a BMI $>30$, or were unable to give informed consent were excluded. Radiated women who previously had had abdominal liposuction or surgery that make a DIEP flap inappropriate and/or were over the age of 60 years were also excluded, as well as non-radiated women with extensive scarring on the thorax. When the study was designed, the department was reluctant to reconstruct patients over the age of 60 and comorbidities with DIEP flaps and therefore they were excluded from randomization. This is no longer departmental practice. If a patient had a recurrence during the follow-up time of the study, she was excluded. In the non-radiated arm, patients were randomized to either a one-stage lateral thoracodorsal flap with a permanent implant (TD)\textsuperscript{17} (a perforator based flap similar to the anterior lateral intercostal artery perforator flap (LICAP),\textsuperscript{18} but without intramuscular dissection of the perforator) or a 2-staged expander reconstruction (EXP).\textsuperscript{19} In the radiated arm, patients were randomized either to a latissimus dorsi-flap (LD) combined with a permanent implant\textsuperscript{20} or to a deep inferior epigastric artery perforator flap (DIEP).\textsuperscript{21} Contralateral procedures for symmetry were performed at the time of the breast reconstruction, if indicated. Information on demographics, patient characteristics, comorbidities, breast cancer stage, breast cancer treatment, and early complications have been published previously.\textsuperscript{16}

**Patient Reported Outcomes**

Patient reported outcomes were measured pre- and post-operatively. The participants were sent an envelope including the questionnaires and a stamped return envelope. The participants received two reminders, after two and five weeks, and a maximum of 5 attempts to reach non-responders by phone were made 3 to 9 months later. The instruments used are described below.

**BREAST-Q** reconstruction measures different aspects of satisfaction with breast reconstruction, with care and health-related quality of life specific to the breast(s). The following domains were analyzed: Satisfaction with breast(s) (16 items), Satisfaction with outcome (7 items), Psychosocial well-being chest (10 items), Sexual well-being (6 items), Physical well-being chest (16 items), and Satisfaction with information (15 items). Each item is rated on a Likert scale and for each domain a raw scale summed scores is calculated and converted to a standardized score between 0 and 100, based on a conversion table created with transformed Rasch logits. A higher score indicates a higher level of patient satisfaction. Normative data have been described for 2 American populations including a total of 1500 women\textsuperscript{22,23} and 1 Australian population including 500 women\textsuperscript{24} (Electronic supplement 1). There are no anchor-based minimal important differences (MIDs) published for BREAST-Q,\textsuperscript{25} but distribution-based MIDs, indicating the lowest change value beyond the measurement error,\textsuperscript{26,27} is 4 for Satisfaction with Breasts, 4 for Psychosocial Well-being, 3 for Physical Well-being, and 4 for Sexual Well-being.\textsuperscript{28} BREAST-Q has been validated\textsuperscript{29,30} and translated to Swedish. It has been extensively used in previous studies on breast reconstruction.\textsuperscript{15} Use of BREAST-Q, authored by Drs. Klassen, Puic and Cano, was made under license from Memorial Sloan Kettering Cancer Center, NY.

**Beck’s Depression Inventory (BDI-2)** comprises 21 items on the patient’s symptoms of depression during the last week. Each item is
rared on a 4 point-scale (0-3), where 0 indicates minimal symptoms and 3 maximum symptoms, and a total score is calculated. A total score of 0-13 indicates minimal depression, 14-19 mild depression, 20-28 moderate depression, and 29-63 severe depression. BDI-21 has previously been used to evaluate breast reconstruction. BDI-21 was purchased from Pearson Clinical.

RAND 36-Item Short Form Health Survey, a public domain form of the Medical Outcome Study Short Form 36- Item Health Survey (MOS SF-36), is a generic health-related quality of life questionnaire. RAND 36 was developed as part of the Medical Outcomes Study (MOS). It comprises eight subdomains: physical functioning (10 items), role limitations caused by physical health problems (4 items), role limitations caused by emotional problems (3 items), social functioning (2 items), emotional well-being (5 items), energy/fatigue (4 items), pain (2 items), and general health perceptions (5 items). The patients score their health during the last four weeks on Likert scales. Items are transformed into a percentage of the highest possible score with each item utilizing 100 as the most positive result. Individual constituent items are subsequently averaged to give a score for each subdomain. RAND-36 has been translated to Swedish and validated for Sweden and has previously been used to evaluate breast reconstruction (eg). There are age-matched normative Swedish values for RAND-36 (Electronic supplement 1). The instrument was used with permission from the RAND corporation.

EuroQol-5 dimensions (EQ-5D) was developed to enable comparison of different diseases by giving both descriptive data and an overall index of health. The questionnaire has 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, where the patient rates his/her health on a 3-level scale. A global score is calculated, where 1 indicates "perfect health" and 0 "death". EQ-5D-3L also comprises a visual analogue scale (VAS) where the patient marks his/her current health state, from 0 ("worst imaginable") to 100 ("best imaginable"). EQ-5D has been validated for Swedish and for breast reconstruction. There are age-matched Swedish normative values for EQ-5D (Electronic supplement 1). EQ-5D was used with permission from EuroQol Research Foundation, Rotterdam, the Netherlands.

Statistical Analyses
Descriptive data were given as both median and interquartile ranges (IQRs) and ranges, as well as means and standard deviations (SD), and frequencies, when applicable. Missing values were treated as described in the manual for BREAST-Q, that is if less than half of the items were missing for a domain, missing data was replaced with the mean of the answered items and if more than half of the items were missing the domain was excluded. QScore was used to calculate domain scores. For the other instruments, incomplete questionnaires were excluded. Changes in scores for the instruments were calculated for each participant, presented as delta (δ) (net change in score between pre- and postoperative score), and differences between the groups were analyzed with the non-parametric Mann-Whitney U-test for unpaired samples (BREAST-Q scores, RAND 36 scores, and EQ5D VAS), and Pearson’s Chi-squared test.
(\(\chi^2\)) for categorical variables (dichotomized EQ 5D dimensions). Statistical analyses were not performed for BDI-21 as many of the questionnaires had to be excluded due to missing values. Statistical comparison was not made between the values of the groups and the norm population as we do not have access to the data sets, only the mean and SDs, for the norms. Statistical analyses were preformed using SPSS version 25 (IBM, Armonk, NY).

**Results**

**Participants, Recruitment, Randomization, and Intervention**

During the recruitment period (2009-2015), we identified 684 potential participants at the referral stage and a total of 405 were invited after assessment at the consultation. Two-hundred thirty-three accepted and were randomized between the methods of reconstruction (Figure 1). One-hundred ninety-one participants were operated with the allocated method (Figure 1). After opt-outs and exclusions (Figure 1), the remaining 107 participants comprise the present study sample (Table 1). None of the participants developed a recurrence during the follow-up time. Postoperative HrQoL was measured on average 7 to 8 years post-operatively (Table 1). Response rated varied between 60 and 82 per cent in the different groups (Figure 1).

**Breast Specific Quality of Life: BREAST-Q**

The BREAST-Q scores were higher after the reconstruction than before for the great majority domains in both arms (Table 2, Electronic supplement 2). In the non-radiated arm, the increase in score was highest for the domains psychosocial well-being (13 and 10 points in the groups) and sexual well-being (13 and 10 points). The greatest difference in change in scores between the two groups were seen for satisfaction with breasts (10 vs. 5 points) and physical well-being (6 vs. -4 points) where the satisfaction increased the most in the EXP group. There was a statistically significant difference between the 2 methods for physical well-being chest in the radiated arm. In the radiated arm, the increase in score was highest for the domain psychosocial well-being (20 and 19 points) in both groups and for satisfaction with breast and sexual well-being in the DIEP-group. The greatest difference in change in scores between the 2 groups were seen for satisfaction with breast (26 vs. 4 points) and sexual well-being (18 vs. 4 points), albeit only the difference for sexual well-being was statistically significant (Table 2, Electronic supplement 2). The scores were on the same level as the normative scores (Electronic supplement 1).

**Symptoms of Depression: Beck’s Depression Inventory (BDI-21)**

The great majority of participants in both arms had minimal or mild depression both before and after the operation (Table 3). Statistical analyses were not performed as many of the questionnaires had to be excluded due to missing values.

**General Quality of Life: RAND-36 and EQ-5D**

As measured with RAND-36, small differences could be detected between the preoperative and postoperative scores in both arms, for most dimension; although, no statistically significant differences were seen (Table 4, Electronic supplement 2). In the non-radiated arm, the median difference in scores before and after the reconstruction varied between -5 and 8. There was a statistically significant difference between the groups for Social functioning and Mental health, where the EXP group scored higher (P-value 0.05 and 0.0017, respectively). In the radiated arm, the median difference in scores before and after the reconstruction varied between -13 and 17. There were no statistically significant differences in RAND-36 between the groups in the radiated arm (Table 4, Electronic supplement 2).
Please cite this article as: Fredrik Brorson et al. Patient Reported Outcome and Quality of Life After Delayed Breast Reconstruction - An RCT Comparing Different Reconstructive Methods in Radiated and Non-radiated Patients, Clinical Breast Cancer, https://doi.org/10.1016/j.clbc.2022.09.004

Most patients did not report any problems pre- or post-operatively as measured with EQ-5D (Table 5). In the non-radiated arm, the EXP group reported more positive changes than the TD group. There was a decrease in reported problems for both groups in Mobility, and an increase in reported problems was seen in both groups for Pain/discomfort, but none of those changes reached statistical significance. The greatest difference between groups was seen in Anxiety/depression where the proportion of the EXP group reporting problems clearly decreased while the TD group showed a slight increase; although, this did not reach statistical significance (P-value .25). There was no statistically significant difference between the groups in VAS scale changes, although the EXP group showed a numerical increase while the TD group remained stationary.

In the radiated arm, the LD group reported greater improvement in Usual activities and worsening in Pain, as opposed to the improvement found in the DIEP group. Both groups reported an improvement for Anxiety/depression, but none of these changes reached statistical significance. Both groups reported an increase in EQ5D VAS with a statistically significant greater improvement in the DIEP group (P-value .039,Table 5).

Discussion
PROM is considered one of the central outcome measures in the evaluation of breast reconstruction. In this study, a combination of breast-specific and generic instruments was used. This study is one of few prospective, randomized studies comparing the outcome of different methods for delayed breast reconstruction in radiated and non-radiated patients. It is illustrative of the methodological difficulties inherent in randomized surgical trials.
Table 4  RAND-36 Scores in the Two Study Arms

<table>
<thead>
<tr>
<th>RAND-36 Dimension</th>
<th>Preoperative Median (IQR)</th>
<th>Postoperative Median (IQR)</th>
<th>δ</th>
<th>Preoperative Median (IQR)</th>
<th>Postoperative Median (IQR)</th>
<th>δ</th>
<th>Difference Between Groups (Mann Whitney U-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EXP n=45</td>
<td>TD n=34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d EXP vs. d TD (δ)</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>90 (5)</td>
<td>92.5 (10)</td>
<td>0 (10)</td>
<td>95(5)</td>
<td>95 (7.5)</td>
<td>0 (5)</td>
<td>0.56</td>
</tr>
<tr>
<td>Role physical</td>
<td>100 (13)</td>
<td>100 (13)</td>
<td>0 (13)</td>
<td>100 (13)</td>
<td>100 (25)</td>
<td>0 (13)</td>
<td>0.25</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>74 (14)</td>
<td>74 (15)</td>
<td>0 (16)</td>
<td>75 (20)</td>
<td>74 (20)</td>
<td>0 (13)</td>
<td>0.99</td>
</tr>
<tr>
<td>General Health</td>
<td>60 (2.5)</td>
<td>57 (3)</td>
<td>-5 (8.8)</td>
<td>57 (6)</td>
<td>60 (2.5)</td>
<td>0 (5)</td>
<td>0.099</td>
</tr>
<tr>
<td>Vitality</td>
<td>55 (5)</td>
<td>55 (5)</td>
<td>0 (7.5)</td>
<td>55 (7.5)</td>
<td>50 (6.3)</td>
<td>0 (7.5)</td>
<td>0.78</td>
</tr>
<tr>
<td>Social functioning</td>
<td>100 (13)</td>
<td>100 (13)</td>
<td>0 (6.2)</td>
<td>100 (6.3)</td>
<td>100 (13)</td>
<td>0 (6.3)</td>
<td>0.05</td>
</tr>
<tr>
<td>Role emotional</td>
<td>100 (12)</td>
<td>100 (12)</td>
<td>0 (17)</td>
<td>100 (17)</td>
<td>100 (17)</td>
<td>0 (0)</td>
<td>0.12</td>
</tr>
<tr>
<td>Mental health</td>
<td>64 (4)</td>
<td>68 (4)</td>
<td>8 (6)</td>
<td>66 (2)</td>
<td>64 (4)</td>
<td>0 (4)</td>
<td>0.0017</td>
</tr>
</tbody>
</table>

|                   | DIEP n = 14              | LD n = 12                  |    |                          |                             |   | DIEP vs LD (δ)                                  |
| Physical functioning | 95 (6.2)                  | 95 (13)                    | 0 (6.2) | 90 (7.5)                | 80 (18)                     | -10 (7.5) | 0.22                                           |
| Role physical     | 100 (13)                 | 100 (0)                    | 0 (0) | 75 (38)                 | 100 (50)                   | 0 (13) | 0.17                                           |
| Bodily pain       | 74 (12)                  | 73 (22)                    | 0 (15) | 64 (6)                  | 51 (23)                    | -10 (41) | 0.17                                           |
| General Health    | 60 (3.2)                 | 56 (8.2)                   | -5 (7.5) | 62 (5)                  | 59 (6.5)                   | 0 (13) | 0.15                                           |
| Vitality          | 50 (6.2)                 | 50 (5)                     | -5 (10) | 60 (7.5)                | 55 (5)                     | 0 (7.5) | 0.91                                           |
| Social functioning| 100 (16)                 | 100 (6.2)                  | 0 (13) | 88 (17)                 | 75 (28)                    | -13 (16) | 0.12                                           |
| Role emotional    | 100 (17)                 | 100 (0)                    | 0 (8.2) | 67 (50)                 | 67 (34)                    | 17 (51) | 0.85                                           |
| Mental health     | 48 (5)                   | 64 (4)                     | 4 (8) | 68 (8)                  | 66 (6)                     | 4 (6) | 0.94                                           |

δ: Delta – difference between pre- and postoperative values
DIEP: Deep inferior epigastric artery perforator flap
EXP: Expander
IQR: Interquartile range
LD: Latissimus dorsi flap
TD: Thoracodorsal flap

Previous studies indicate that generic quality of life instruments might not be sensitive enough to capture post-operative change. The present study largely corroborates this, as EQ-5D-3L (Table 5) and RAND-36 (Table 4, Electronic supplement 2) failed to detect any statistically significant differences in most dimensions. Nor did we detect a discernible pattern of long term changes in general HrQoL as we found positive as well as negative changes. This can be seen as regression to the mean or at least as an uncertain causal relationship with the method of reconstruction. Most participants in the present study were middle aged at inclusion and 7 to 9 years passed between the pre- and postoperative measurements. Therefore, it is reasonable to assume that other life changes and natural aging influence scores and that the effect of the breast reconstruction itself cannot be isolated. Therefore, the lower scores do not exclude a beneficial effect of breast reconstruction. In summary, generic quality of life measurements do not give specific enough information for comparison of methods of reconstructive breast surgery. The number of questionnaires used in this study could have affected the response rate and for future studies it can be discussed if generic quality of life measurements should be excluded, to decrease the risk for questionnaire fatigue in the participants.

Regarding the breast specific instrument, BREAST-Q, there were few statistically significant differences between the methods of the 2 arms, except for physical well-being chest in the non-radiated group and sexual well-being in the radiated group. Nonetheless, this does not exclude that there are clinically relevant differences between the groups and that 1 method could be superior to the other. Indeed, the change in BREAST-Q score was larger in the EXP-group than in the TD-group and in the DIEP-group than in the LD-group (Table 2, Electronic supplement 2), which could indicate that the methods give slightly higher patient satisfaction. The difference between pre- and post-operative BREAST-Q scores were larger than the published distribution-based MIDs for all domains, except for physical well-being chest in the LD-group, indicating that the changes were beyond the measurements error. However, to fully interpret the clinical relevance of differences in scores we need knowledge about the clinically relevant MIDs of the instrument, that is anchor based MIDs, and how repeated HrQoL measurements in breast reconstruction patients are affected by other factors, such as response shift, that is the person adapts to his/her situation with time. Another way to interpret PROM scores is to compare them with normative values. We present age-matched data from 2 samples in a Swedish...
population survey using EQ5D and RAND 36 [44] and 3 recently published studies using BREAST-Q [13, 14, 20] (Electronic supplement 1). However, the latter come from American and Australian populations and there seem to be cultural differences in how women score BREAST-Q, which might make the comparison misleading. Moreover, none of the studies report scores or women of comparable age to the postoperative PROM measurements in the present study.

It can be argued that combined overall improvements in generic as well as in breast specific quality of life may be taken as a form of indirect evidence that 1 method of reconstruction is superior; for example, if 1 reconstructive method shows improvement in all 3 PROM scales whilst another shows improvement in just 1 or 2 scales. However, we would still lack reliable ways to confirm causal links if correlations were to be found, and there is currently no certain way of transforming EQ-5D, RAND-36 and BREAST-Q scores for a total composite score, even if there are reports on mapping algorithms for SF/RAND-36 and EQ5D comparisons [43, 44]. The same dilemma occurs for reliably assigning weights for meta-analysis comparing different scales. This suggests that scales should be analyzed individually, and to fully appreciate the meaning of our findings, knowledge about MID, factors that affect breast reconstruction patient PROMs, and normative data are essential.

The study was designed as a superiority study. However, the sample size was not calculated based on PROMs and therefore the failure to demonstrate a statically significant difference between the groups does not imply that the reconstructive techniques are equivalent in terms of patient satisfaction. It could merely be a case of insufficient sample size and power. Differing proportions of non-responders at follow-up may affect results, especially in the radiated arm. Moreover, it is unclear how PROMs should be weighed in relation to costs and to other outcomes, such as frequencies of complications and unplanned operations. To enable a balanced collective evaluation of which type of breast reconstruction a patient should be offered, we would have to know both what constitutes a clinically relevant difference in PROM scores and what size of change in score would be relevant to make a reconstructive technique superior enough, in terms of PROMs, to warrant a higher complication and...
re-operation rate or a higher total long-term cost. Moreover, to draw any conclusions on equivalence of treatment, an equivalent trial design and clinically relevant equivalence margins are necessary. To enable RCTs in reconstructive surgery, with PROMs as the primary end point, more studies are needed regarding how PROMs should be weighed against other evidence.

In the present study we have analyzed only participants who are recurrence free survivors, with their original method of reconstruction preserved, to identify differences in outcomes between breast reconstructive methods. We have chosen to compare the difference in change (Δ) between groups to minimize the number of analytic statistical tests. Excluding participants post-randomization is sometimes referred to as a modified intention to treat analysis, and it has potential weaknesses. Nonetheless, it may be the only pragmatic way of acquiring data in surgical trials. A strict intention to treat analysis would likely not provide meaningful answers to the research questions due to confounders and attrition, and the sample size would become even smaller.

The problem in RCTs concerning methods for breast reconstruction of surgeons’ and patients’ preferences for certain methods has been discussed previously, making recruitment and randomization and unbiased results a challenge. A fundamental prerequisite to include patients in a trial is a solid uncertainty about which reconstructive method is more beneficial. In the case of reconstructive breast surgery, this uncertainty becomes relevant on several levels. In the present study, roughly 2 thirds of potential participants identified at the referral stage were invited to participate after consultation and about half accepted, raising concerns about potential inclusion and selection bias. The recruiting and operating surgeons must not prefer a certain method (theoretical equipoise), based on personal experiences, neither in terms of a perceived knowledge, nor in terms practical skills or costume. Such a preference could lead to selection bias, as only patients who fit a particular surgeon preconception might be recruited to the study, and to a biased result, as patients asked about inclusion receive biased information. Moreover, the patients included in the study must not have pre-formed ideas about the different methods (principle of indifference). This has proved a barrier in previous studies and in an era were the importance of patient choice and involvement in the decision process is considered crucial. Patient often have pre-formed conceptions of different methods, based on information from, for example, other patients, patient organizations, and the media. The fact that more patients who were randomized to a TD flap in the non-radiated arm and to a LD flap in the radiated arm opted out of the study could bear evidence that the principle of indifference was not met. Since the primary outcome of the present study is PROMs and patient satisfaction, a lack of uncertainty about which reconstructive method is more beneficial could be inimical, especially if a patient perceives that she is not offered the best available treatment. In addition, if one treatment option is believed to be superior, the participants receiving the other treatment might be more prone to breaches of protocol (to opt out and choose the superior treatment), potentially creating a misleadingly high score for the other treatment if analyzed per protocol. In other words, attrition bias would make it more likely for more satisfied participants to remain in the per protocol group. Thus, it is possible that the impact of reconstructive method on HrQoL is greater than the present results imply.

Conclusion
No distinct differences in long-term HrQoL could be seen for different methods in recurrence free women treated with unilateral mastectomy randomized to different delayed breast reconstruction techniques. There was a clear improvement in HrQoL 7 to 8 years post-reconstruction compared to pre-reconstruction in all groups, but the effect of specific reconstructive methods on scores could not be reliably demonstrated. Selection, inclusion, and attrition biases further complicate the interpretation of RCTs in breast reconstruction and may limit generalizability of findings. More studies regarding interpretation of HrQoL instruments are needed.

Clinical Practice Points
There are few studies providing high-quality evidence supporting the use of different techniques for breast reconstruction. In this study, the patients’ health related quality of life was increased in all groups. The BREAST-Q scores were higher after the reconstruction than before for the great majority domains in both arms; albeit, statistically significant only for physical well-being chest between the 2 methods in the non-radiated arm. Most participants in both arms had minimal or mild depression both before and after the operation. The study highlights the difficulty in conducting randomized controlled trials in preference sensitive procedures, such as breast reconstruction. Selection, inclusion, and attrition biases complicate the interpretation of the results.

Acknowledgment
We would like to thank all the women who participated in the study. The study was funded by grants from The Swedish Cancer Society [21 0279 SCIA], the federal government under the ALF agreement [ALFGBG-724171 and ALFGBG-965161], The Percy Falk Foundation (Stockholm, Sweden) for research into prostate cancer and breast cancer, and The Swedish Breast Cancer Association. The sources of funding had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Disclosure
The authors have no conflict of interest.

Supplementary materials
Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.clbc.2022.09.004.

References


Erikson F, Nordhund AA. Health and health related Quality of life as measured by the EQ-5D and the SF-36 in south-east Sweden: results from two population surveys 2002/1. 2002.


