Surgical determinants of patient-reported outcomes following post-mastectomy reconstruction in women with breast cancer

Running head: Breast reconstruction patient outcomes

R Jeevan[1,2], JP Browne[3,4], C Gulliver-Clarke[5], J Pereira[6,7], CM Caddy[8],

JHP van der Meulen[1,3], DA Cromwell[1,3]

[1] Clinical Effectiveness Unit, The Royal College of Surgeons of England, 35-43 Lincoln's Inn Fields, London, WC2A 3PE

[2] St Helens and Knowsley Teaching Hospitals NHS Trust, Warrington Road, Prescot, Merseyside, L35 5DR

[3] Department of Health Services Research & Policy, Department of Public Health & Policy, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT

[4] Department of Epidemiology and Public Health, University College Cork, Cork, Ireland

[5] Western Sussex Hospitals NHS Foundation Trust, Worthing Hospital, Lyndhurst Road, Worthing, West Sussex, BN11 2DH

[6] James Paget University Hospitals NHS Foundation Trust, Lowestoft Road, Gorleston, Great Yarmouth, Norfolk, NR31 6LA

[7] University of East Anglia, Norwich Research Park, Norwich, Norfolk, NR4 7TJ

[8] Sheffield Teaching Hospitals NHS Foundation Trust, Royal Hallamshire Hospital, 4 Claremont Place, Sheffield, S10 2JF

Correspondence to:

Prof David Cromwell

Clinical Effectiveness Unit, The Royal College of Surgeons of England, 35-43 Lincoln's Inn Fields, London WC2A 3PE

Tel: +44 207 869 6608

Fax: +44 207 869 6644

Email: dcromwell@rcseng.ac.uk

ABSTRACT

Background: This national prospective cohort study compared the patient-reported outcomes of breast cancer patients having post-mastectomy autologous reconstruction to those who had breast implants, in terms of aesthetic appearance, levels of psychological, physical and sexual well-being and overall satisfaction.

Methods: Of 5,063 women who underwent immediate (IR; n=3,349) or delayed (DR; n=1,714) reconstruction between 1 January 2008 and 31 March 2009 in England, 2,923 women who consented were sent validated, procedure-specific 18-month follow up questionnaires. Outcome scale scores ranged from 0 (poor) to 100 (excellent); multiple linear regression was used to adjust scores for patient and treatment characteristics.

Results: 2,289 women (78%) returned completed questionnaires (1,528 IR, 761 DR). For IR, mean overall satisfaction scores for the various techniques ranged from 67 to 85 (median 67 to 93). For DR, mean overall satisfaction scores ranged from 70 to 85 (median 75 to 100). For both groups, similar gradients were observed for the other outcome scales across techniques. Reconstruction using patients' own tissues tended to have higher mean adjusted scores compared with those techniques using implants alone ($p<0.0001$ for aesthetic appearance, psychological well-being, sexual well-being, and satisfaction with outcomes for IR and DR groups).

Conclusion: Women who underwent autologous reconstruction tended to report greater satisfaction than those who had an implant reconstruction. These results can inform patients of the anticipated outcomes of their selected surgery but further research is required to confirm whether autologous reconstruction is superior in general.

Introduction

Women who undergo mastectomy surgery for breast cancer can have their breast mound reconstructed in a variety of ways.[1] All of the options aim to address the aesthetic impact of mastectomy and may also improve patients' physical, emotional and sexual well-being.[2] There has been much debate regarding the best way to assess the effectiveness of different types of reconstructive procedures.[3,4] Internationally, the consensus view has shifted towards the use of patient-reported outcome measures rather than anthropomorphic measurements or measurement scales scored by clinicians.[2] This is in part because previous studies have found little correlation between clinician-reported and patient-reported aesthetic satisfaction following such surgery and partly because of a recognition that the patient should be the ultimate judge of subjective outcomes such as aesthetic appearance.[5,6]

Numerous studies have evaluated patient-reported outcomes following different types of breast reconstruction in populations of between 45 and 2,328 women.[6-15] However, these studies demonstrated methodological weaknesses in one or more of the following areas: inadequate length of follow up; a failure to include immediate reconstruction (IR) and delayed reconstruction (DR) patients, the full range of reconstructive techniques, and all eligible centres; lack of a prospective design; a lack of case-mix adjustment; a lack of validated and surgery-specific scales; and a failure to formally calculate outcome scores for these scales to enable valid comparisons.

We undertook a national prospective cohort study that examined patient-reported outcomes after mastectomy and breast reconstruction surgery for women treated in England, Scotland and Wales.[16] The study collected data from women having IR or DR procedures using a set of validated surgery-specific scales that have been used widely to evaluate the outcomes of

breast reconstruction.[2,17] In this paper, we compare the outcomes of six different reconstructive techniques.

Methods

Treatment sites and inclusion criteria

During the 15 month recruitment period from 1 January 2008 to 31 March 2009, all 150 English NHS acute hospital trusts providing acute breast cancer treatment, six NHS acute trusts in Wales and Scotland, and 114 independent hospitals in England participated in the study. Treatment sites were asked to prospectively collect and record clinical data on women aged 16 years and over with a diagnosis of invasive carcinoma of the breast or ductal carcinoma in-situ (DCIS) undergoing mastectomy surgery with IR or primary DR following a previous mastectomy, and to obtain written consent from eligible women to allow them to be sent follow up questionnaires.

Reconstructive procedure types

We compared six different surgical techniques: tissue expander or breast implant alone; pedicled latissimus dorsi flap combined with an expander or implant (LD with implant/expander); LD flap alone (autologous LD); pedicled transfer of a transverse rectus abdominis myocutaneous flap (pedicled TRAM); free tissue transfer of this flap (free TRAM); or a free tissue transfer of a similar flap without the underlying muscle, based on the deep inferior epigastric artery perforator or superficial inferior epigastric artery (DIEP or SIEA). The remaining patients had a flap of skin and fat with or without muscle taken from the upper or lower buttock or inner thigh regions (SGAP, IGAP or TUG), but they were excluded from our analyses due to the extremely small numbers enrolled.

<u>Clinical data collection</u>

A range of data items were recorded by clinicians for each patient. These included details of surgical procedures, patient clinical and sociodemographic characteristics, and consent status. The full dataset is available at https://www.rcseng.ac.uk/standards-and-research/research/clinical-effectiveness-unit/

<u>Patient questionnaires</u>

The post-reconstruction questionnaires addressed patient satisfaction with their breast area appearance (16 items with an additional 2 for implant patients), which we believe to be synonymous with aesthetic outcome from the patient's perspective, along with emotional well-being (10 items), physical well-being (16 items), sexual well-being (6 items), and the outcomes of surgery (7 items) respectively. The scales were derived from the Breast-Q,[2] and were pre-tested with English breast cancer patients by the authors prior to their use to ensure that there were no issues with language, comprehension or acceptability. Copies of the full questionnaires are available at https://www.rcseng.ac.uk/standards-and-research/research/clinical-effectiveness-unit/

<u>Questionnaire administration and collection</u>

Clinicians were asked to obtain consent from any eligible women at the time of surgery to allow follow-up questionnaire administration. The questionnaires were sent to consented patients at their home address 18 months after surgery by a co-ordinating team of researchers that did not include the treating hospitals or clinicians, once the team had confirmed the patient was still alive by cross-checking their details against mortality data held by the National Strategic Tracing Service (NSTS). A prepaid envelope was enclosed to facilitate the return of the completed questionnaire. Questionnaires were marked only with a unique numeric patient identifier. Non-respondents were sent a single reminder letter and an additional copy of the questionnaire at a five week interval.

Statistical analysis and multiple linear regression model

Patients' responses were entered into a database and then transferred to a bespoke Q-Score software package to calculate scores for each Breast-Q scale, with 0 the lowest and 100 the highest possible scores.[2] Patient scores were linked to their associated clinical data using their unique numeric identifier.

Separate linear multiple regression models were developed for the immediate and delayed reconstruction patient groups, and were used to predict the outcome scores for each Breast-Q scale (dependent variables) based on patient characteristics and reconstructive procedure type. The preliminary models included only patients with complete outcome and case-mix data, and were constructed using a backward stepwise process with variables dropped from the models if the strength of their association with an outcome was weak (Wald test for variable inclusion, p-value<0.05). Variables were included in continuous or categorical format depending on their type. The case-mix variables included sociodemographic items (age, a geographically assigned measure of socioeconomic deprivation, and hospital of treatment), patient factors known to affect subsequent morbidity (smoking status, body mass index, diabetes status, American Society of Anaesthesiologists grade, Eastern Cooperative Oncology Group performance status), and tumour characteristics (invasive status and Nottingham Prognostic Index).[18-21] Deprivation in England is measured by the government using seven distinct domains or indicators of poverty (income; employment; education, skills & training; health; crime; barriers to housing and services; living environment), that are weighted before a deprivation score and rank are produced for each geographical area (super output area or SOA, a small local population of a few thousand people).

The Eastern Cooperative Oncology Group (ECOG) functional status categorical scoring system is an American measure used internationally to measure the functional status of patients undergoing cancer treatments. It is also known as the WHO or Zubrod score.

The models were then used to produce adjusted means and confidence intervals for each scale score, by procedure type, which demonstrated the effect size. Finally, the Wald test was undertaken for the type of reconstructive procedure to examine the heterogeneity of outcomes across different types of reconstruction and determine whether or not the differences in means were due to chance alone.

All statistical tests were two-sided and p-values less than 0.05 were considered to represent a statistically significant result. All statistical analyses were undertaken using STATA/MP 14 (Stata Corporation, College Station, Texas, US) and Microsoft Excel software.

Results

Patient population

Within the study cohort, 3,349 patients underwent IR. Of these, 43 did not have their consent status for the follow up study recorded, 35 were deemed incapable of completing a written questionnaire in English and 1,148 were not asked to participate due to difficulties with the recruitment process at some hospitals. Of the remaining 2,123 women, 1,939 gave their consent and 1,528 returned a completed questionnaire. After excluding 144 women for whom complete case-mix data was not available, responses from 1,384 were included in our final analyses. Questionnaires were completed 586 days after surgery, on average, with an interquartile range of 30 days.

There were 1,714 patients who underwent DR. Of these, 9 did not have their consent status for the follow up study recorded, 6 were deemed incapable of completing a written questionnaire in English and 609 were not asked to participate. Of the remaining 1,090 asked to participate, 984 gave their consent and 761 retuned questionnaires. A further 28 patients were excluded due to incomplete case-mix data; this left 733 for our final analyses. The mean length of time between a patient's mastectomy and their subsequent DR

procedure was, on average, 2.8 years (range 0 to 32 years). There were small differences in the mean time to DR by procedure type, varying from 2.4 years (LD flap with implant) to 2.9 (free flap) and 3.0 years (implant/expander-only) respectively. Questionnaires were completed 582 days after surgery, on average, with an interquartile range of 26 days. Table 1 summarises the consent and response rates for the different reconstruction techniques.

Patient characteristics across the surgical groups

Table 2 displays the sociodemographic and clinical characteristics of the women who underwent each type of procedure. Patients undergoing different procedures were generally similar. Those characteristics that were significantly associated with the outcomes under investigation were included in the two separate multiple linear regression models used subsequently to derive final case-mix adjusted scores for the respective patient groups. During the model development process the characteristics significantly associated with outcomes were ethnicity, deprivation level, American Society of Anaesthesiologists (ASA) grade and tumour burden for immediate reconstruction patients, and ethnicity, Eastern Cooperative Oncology Group (ECOG) score and diabetes status for those undergoing delayed reconstruction. The other characteristics examined (including age, radiotherapy, chemotherapy, smoking status and body mass index) were not independently associated with outcomes and were not therefore included in the final models. The IR patients were of similar age and functional status to those undergoing DR, but were less likely to be from a white ethnic group or the most deprived quintile, or to have a high ASA grade or diabetes. IR patients were also much more likely to have a low tumour burden (DCIS or low risk invasive).

Breast-Q scores for each procedure

Figure 1 presents the IR and DR procedure-specific unadjusted mean scores for each of the five Breast-Q scales, along with the associated 95% confidence intervals.

With respect to breast area appearance, in the IR group women who underwent a pedicled TRAM procedure had the highest outcome scores. However, the sample size for this group was relatively small (n=34) and thus the precision associated with this estimate is relatively poor. The other patient groups who underwent a flap reconstruction (with or without implant) also had relatively high scores. The lowest breast area appearance scores were associated with patients who underwent expander or implant-only reconstruction.

In the DR group, free TRAM flaps, DIEP flaps and then the other types of flap-based reconstruction (with or without implant) were associated with the highest breast area appearance scores. Again, the expander or implant-only patient group had the lowest scores.

After adjustment for patient characteristics, there were persistent differences in the outcomes achieved by the procedure groups, with the exception of physical well-being. Table 3 shows the adjusted differences in the mean scores for each scale, using the expander / implant group as the reference category. In both the IR and DR settings, all flap-based procedures resulted in higher scores on all scales other than physical well-being. The only exception to this pattern was for the IR pedicled TRAM group.

Discussion

<u>Key findings</u>

Women who underwent reconstruction tended to be satisfied with the results of the procedure, with mean overall satisfaction scores of 71.1 (95% CI 69.8 to 72.4) for IR procedures and 79.3 (95% CI 77.7 to 81.0) for DR procedures.

Patients who underwent autologous reconstruction reported higher scores than those women who had an expander or implant alone, in both the IR and DR settings, across four of

the five scales. For immediate procedures, the size of these differences was typically between 1 and 7 points. For delayed procedures, the difference between the scores tended to be between 7 and 15 points.[22] To interpret the magnitude of these differences it is useful to compare to a cohort of patients who underwent mastectomy alone over the same period at the same group of hospitals.[16] These patients recorded a mean score of 56 on the breast area appearance scale, which is similar to the mean unadjusted score recorded for implant-only patients recorded in this study but between 8 and 21 points lower than that recorded for patients undergoing autologous IR procedures. A similar pattern was seen for emotional well-being (1 point difference for implant-only; 5 to 16 point difference for autologous procedures) and sexual well-being (7 point difference for implant-only; 13 to 28 point difference for autologous procedures), while physical well-being scores were similar in those who underwent mastectomy or any form of IR. There is no agreed definition of what constitutes a clinically important difference for the Breast-Q scales used in our study but a difference of approximately 10 points on each scale may be taken to be equivalent to a moderate effect size using Norman's standard formula of one half a standard deviation.[23]

We would caution that these results should not be interpreted as a prescriptive indication in favour of one set of procedures over another. We were not able to collect data on women's perspectives on what might have led them to select one type of reconstruction over another. Consequently, we have no way of taking into account their baseline expectations or the influence of their preferences for the level of surgical insult they were willing to endure. We therefore suggest that the results should only be used to inform women of the outcomes they might expect to achieve with different treatment scenarios and should be employed alongside information about the surgical approach, complication rates and recovery time.

Strengths and limitations

Our study had a number of strengths. First, we used specifically developed and previously validated outcome measures that were distributed centrally at a standardised follow up

interval of 18 months, to ensure that the great majority of women had completed any adjuvant treatments and secondary reconstructive procedures required.[2,17,22,24,25] Second, we examined outcomes for a national population that was recruited prospectively, with explicit written consent obtained prior to inclusion. Third, we had excellent response rates of close to 80%. Fourth, we included women with failed reconstructive procedures within the study to minimise bias. Fifth, to minimise the risk of confounding we undertook robust case-mix adjustment within a multiple linear regression model that included all factors that significantly affected our study outcomes. This adjustment had little effect on the outcome scores of each scale, suggesting that our procedure-specific comparisons were safe and robust.

There were also some weaknesses. While the overall sample of around 1,500 IR and 700 DR patients was large compared to other studies in this field, the size of certain subgroups (e.g. those who had a pedicled TRAM flap) was relatively small, with wide confidence intervals and an increased risk of sampling bias.

Next, not all eligible women were asked to participate in the study and it was not possible to estimate the recruitment rates for specific procedures. It is possible that recruitment was higher for certain procedures and that this has introduced a bias to the estimated outcomes we have recorded. However, there is no obvious reason why recruitment may have been higher for certain procedures and the diffusion of patients in different procedure groups across a very large number of hospitals makes it unlikely that variable recruitment introduced a systematic error to our findings.

There was also the potential for heterogeneity of outcome across the large number of hospitals we included in the study. However, we included the identity of these organisations as a variable in our regression models to adjust for any organisation-level clustering.

Finally, our comparisons do not include recently developed reconstructive techniques such as lipomodelling and acellular dermal matrix and dermal sling procedures that were not widely used during the study period.

Interpretation of our findings

This study is consistent with a growing body of evidence that reconstruction using patients' own tissues is associated with better aesthetic outcomes than reconstruction using only an implant or expander. [6-14] There are a number of explanations for why autologous procedures are associated with superior aesthetic outcomes.

An implant or expander on its own can replace the volume that has been lost during mastectomy but cannot fully reconstitute the breast mound with respect to its shape or position on the chest wall. They are also static devices that will not adjust or change automatically with a patient's body habitus. Implants generally do not produce natural ptosis of the breast as they adhere to the chest wall and cannot be positioned to extend below the inframammary fold, unlike native breast tissue or a flap. An implant can usually only be placed safely under mastectomy skin if covered by an additional layer, whether muscle or alternatively an acellular dermal matrix or dermal flap or sling. If not using a flap, the standard approach is to place an expander under the pectoralis major and stretch the muscle out, as otherwise the muscle is inadequate in terms of coverage and leaves the lower pole of the implant exposed. Using a sub-muscular implant or expander placement without breast tissue to provide additional cover means that the final breast mound shape achieved is difficult to control and predict. Although this may be less of an issue with bilateral reconstruction, our study included only those women with unilateral reconstruction whose contralateral breast was preserved.

Finally, implants lead the body to form a capsule of scar tissue around them. This capsular tissue contracts in a significant proportion of patients, more commonly following

12

radiotherapy. An implant with a contracted capsule sits proud on the chest wall and may cause pain in addition to distortion. These implants may need to be removed or exchanged with capsule management (excision or release) at more frequent intervals for those women who are worst affected; a lifetime of additional procedures for some.

In contrast, using fat with or without muscle and skin from the back, abdomen, buttock or thigh (a flap) has a number of advantages. First, because they consist of the patient's own tissues they grow and shrink with the patient, and the contralateral breast, as their weight changes. Second, as flap constituents are similar to the breast tissue excised, they are better able to mimic the contralateral breast's natural shape, ptosis and movement. Third, infection risk is much lower as the tissues have an intact or restored blood supply and can respond to pathogens in a normal manner. Fourth, there are no capsular issues, and the reconstruction, if primarily successful, is definitive and lasts for life in most cases without the need for replacement.

The principal disadvantage of flap-based reconstruction is the need for a flap donor site that will inevitably be left with a scar, and sometimes with a contour defect, muscle weakness or bulge, sensory changes or another type of longstanding or permanent morbidity. Another disadvantage is that these procedures are usually longer in duration and may expose patients to a higher risk of distant and systemic complications.

If an implant is used in combination with a flap (e.g. in the latissimus dorsi with implant group) the flap provides many of the advantages mentioned above, and ameliorates some of the aesthetic limitations of implant-only reconstruction by providing soft tissue coverage and a more natural shape, while allowing definitive implant placement without the need for a temporising expander. This is reflected in the high outcomes scores associated with such procedures, which in this study were comparable to autologous flap reconstruction.

However, the patient is left with both a degree of donor site morbidity and the need for capsule management and implant replacement over time.

Implications for clinical practice

Clinicians should ensure that women are provided with appropriate and adequate information about all breast reconstruction options as part of the decision-making process, including our principal finding that autologous procedures are associated with better aesthetic outcomes.

Our case-mix findings suggest that although there may be some selection of reconstructive procedure type by indication, the populations undergoing different procedures are broadly comparable. This suggests that many women who undergo implant-based reconstruction may also be suitable for flap-based reconstruction. While not all women would choose autologous procedures due to their greater complexity, longer duration, donor site morbidity and higher overall complication rates, they should be informed that implant-based procedures are not without risks, particularly implant failure,[16] and have now been shown to be associated with lower patient-reported outcome scores.

Conclusion

Women who accessed autologous techniques tended to report greater satisfaction with their reconstruction than those who had an expander or implant. We recommend that clinicians act on our findings by providing all patients with appropriate written and verbal information regarding the full range of reconstructive options, including those not available at their treating hospital, and their associated patient-reported outcomes.

References

1. Jeevan R, Mennie J, Mohanna P-N, O'Donoghue JM, Rainsbury RM, Cromwell DA. National trends and regional variation in immediate breast reconstruction rates between April 2000 and March 2014. Br J Surg. 2016; 103(9):1147-56.

2. Pusic A, Klassen A, Scott A, Klok J, Cordeiro P, Cano S. Development of a new patient reported outcome measure for breast surgery: The BREAST-Q. Plast Reconstr Surg. 2009; 124:345–353.

3. Pusic AL, Chen CM, Cano S, et al. Measuring quality of life in cosmetic and reconstructive breast surgery: a systematic review of patient-reported outcomes instruments. Plast Reconstr Surg. 2007; 120(4):823-37; discussion 838-9. Review.

4. Winters ZE, Benson JR, Pusic AL. A systematic review of the clinical evidence to guide treatment recommendations in breast reconstruction based on patient- reported outcome measures and health-related quality of life. Ann Surg. 2010; 252(6):929-42.

5. Beesley H, Ullmer H, Holcombe C, Salmon P. How patients evaluate breast reconstruction after mastectomy, and why their evaluation often differs from that of their clinicians. J Plast Reconstr Aesthet Surg. 2012; 65(8):1064-71.

6. Lindegren A, Halle M, Docherty Skogh AC, Edsander-Nord A. Postmastectomy breast reconstruction in the irradiated breast: a comparative study of DIEP and latissimus dorsi flap outcome. Plast Reconstr Surg. 2012; 130(1):10-8.

7. Liu C, Zhuang Y, Momeni A, et al. Quality of life and patient satisfaction after microsurgical abdominal flap versus staged expander/implant breast reconstruction: a critical study of unilateral immediate breast reconstruction using patient-reported outcomes instrument BREAST-Q. Breast Cancer Res Treat. 2014; 146(1):117-26.

8. Yueh JH, Slavin SA, Adesiyun T, et al. Patient satisfaction in postmastectomy breast reconstruction: a comparative evaluation of DIEP, TRAM, latissimus flap, and implant techniques. Plast Reconstr Surg. 2010; 125(6):1585-95.

9. Hu ES, Pusic AL, Waljee JF, et al. Patient-reported aesthetic satisfaction with breast reconstruction during the long-term survivorship period. Plast Reconstr Surg. 2009; 124(1):1-8.

10. Tønseth KA, Hokland BM, Tindholdt TT, Abyholm FE, Stavem K. Quality of life, patient satisfaction and cosmetic outcome after breast reconstruction using DIEP flap or expandable breast implant. J Plast Reconstr Aesthet Surg. 2008; 61(10):1188-94.

11. Eltahir Y, Werners LL, Dreise MM, Zeijlmans van Emmichoven IA, Werker PM, de Bock GH. Which breast is the best? Successful autologous or alloplastic breast reconstruction: patient-reported quality-of-life outcomes. Plast Reconstr Surg. 2015; 135(1):43-50.

12. Alderman AK, Wilkins EG, Lowery JC, Kim M, Davis JA. Determinants of patient satisfaction in postmastectomy breast reconstruction. Plast Reconstr Surg. 2000; 106(4):769-76.

13. Weichman KE, Hamill JB, Kim HM, et al. Understanding the recovery phase of breast reconstructions: patient-reported outcomes correlated to the type and timing of reconstruction. J Plast Reconstr Aesthet Surg. 2015; 68:1370–1378.

14. Atisha DM, Cox CE, Rushing CN, Samsa GP, Pusic AL, Abernethy AP. A national snapshot of satisfaction with breast cancer procedures. Ann Surg Oncol 2014; 21:10e1.

15. Spear SL, Newman MK, Bedford MS, Schwartz KA, Cohen M, Schwartz JS. A retrospective analysis of outcomes using three common methods for immediate breast reconstruction. Plast Reconstr Surg. 2008;122(2):340-7.

16. Jeevan R, Cromwell DA, Browne JP, et al. Findings of a national comparative audit of mastectomy and breast reconstruction surgery in England. J Plast Reconstr Aesthet Surg. 2014; 67(10):1333–44.

17. Cohen WA, Mundy LR, Ballard TN, et al. The BREAST-Q in surgical research: A review of the literature 2009-2015. J Plast Reconstr Aesthet Surg. 2016; 69(2):149-62.

18. Noble M, McLennan D, Wilkinson K, et al. The English indices of deprivation 2007: summary. London: HMSO, 2008.

19. Dripps RD. New classification of physical status. Anesthesiology 1963;24:111.

20. Oken MM, Creech RH, Tormey DC, et al. "Toxicity and response criteria of the Eastern Cooperative Oncology Group". Am. J. Clin. Oncol (1982); 5(6):649–55

21. Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. British Journal of Cancer. 1982; 45(3):361-6.

22. Cano SJ, Klassen AF, Scott A, et al. Interpreting clinical differences in BREAST-Q scores: minimal important difference. Plast Reconstr Surg. 2014; 134:173e–175e.

23. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003; 41(5):582-92.

24. Cano SJ, Klassen AF, Scott AM, et al. The BREAST-Q: further validation in independent clinical samples. Plast Reconstr Surg. 2012; 129:293–302.

25. Cano S, Klassen A, Scott A, Pusic A. A closer look at the BREAST-Q. Clin Plast Surg. 2013; 40:287–296.

Table 1 – Patient recruitment and response rates associated with the different reconstructive procedure groups, in both immediate and delayed reconstruction settings

	Expander / Implant		LD with implant or expander		Autologous LD		Pedicled TRAM		Free TRAM		DIEP / SIEA		Overall
Immediate reconstruction													
Number of patients in study	1242		729		850		82		92		354		3349
Number (%) of patients asked to participate	759	(61.1)	453	(62.1)	594	(69.9)	52	(63.4)	60	(65.2)	222	(62.7)	2140
Number (%) of patients who gave their consent	681	(89.7)	414	(91.4)	520	(87.5)	41	(78.8)	57	(95.0)	204	(91.9)	1917
Number (%) of patients who returned a questionnaire	522	(76.7)	330	(79.7)	430	(82.7)	37	(90.2)	44	(77.2)	165	(80.9)	1528
Number of patients who returned a questionnaire and had complete clinical data	475		284		402		34		36		153		1384
Delayed reconstruction													
Number of patients in study	280		434		356		90		191		363		1714
Number (%) of patients asked to participate	169	(60.4)	253	(58.3)	251	(70.5)	58	(64.4)	115	(60.2)	248	(68.3)	1094
Number (%) of patients who gave their consent	149	(88.2)	233	(92.1)	216	(86.1)	49	(84.5)	104	(90.4)	222	(89.5)	973
Number (%) of patients who returned a questionnaire	109	(73.2)	183	(78.5)	166	(76.9)	35	(71.4)	85	(81.7)	183	(82.4)	761
Number of patients who returned a questionnaire and had complete clinical data	104		176		160		35		83		175		733

Table 2 – Patient characteristics by reconstructive procedure type

		Expander/ Implant	LD with implant or expander	Autologous LD	Pedicled TRAM	Free TRAM	DIEP/ SIEA	Total
Immediate reconstruction								
Number of responders		475	284	402	34	36	153	1,384
Age (years)	Median	51	51	53	52	51.5	51	51
	IQR	15	14	13	12	14.5	12	14
Ethnicity / %	White	85.1	84.5	83.3	97.1	88.9	81.1	84.4
	Non-white	3.4	4.6	4.0	0.0	0.0	9.2	4.3
	Not known / not stated	11.6	10.9	12.7	2.9	11.1	9.8	11.3
Index of multiple deprivation (IMD) Quintile / %	1 (least deprived)	28.8	34.5	28.1	32.4	36.1	33.3	30.6
	2	20.0	25.4	26.4	17.7	33.3	23.5	23.6
	3	22.1	16.9	21.9	8.8	22.2	21.6	20.6
	4	14.1	13.7	12.7	29.4	2.8	12.4	13.5
	5 (most deprived)	15.0	9.5	11.0	11.8	5.6	9.2	11.7
ASA grade / %	1	69.1	76.4	73.1	70.6	72.2	69.9	72.0
	2	28.0	23.2	26.1	29.4	25.0	30.1	26.7
	3 or 4	3.0	0.4	0.8	0.0	2.8	0.0	1.4

ECOG Status / %	0	90.9	93.3	92.0	91.2	94.4	96.1	92.4
	1	7.0	6.4	7.5	8.8	5.6	3.3	6.6
	2 plus	2.1	0.4	0.5	0.0	0.0	0.7	1.0
Diabetes status / %	Diabetic	1.7	1.1	1.5	2.9	2.8	1.3	1.5
Tumour burden / %	DCIS	26.5	27.8	32.3	41.2	36.1	26.8	29.1
	Low NPI	36.2	37.7	30.1	23.5	27.8	26.8	33.2
	Intermediate NPI	24.0	22.9	21.4	26.5	16.7	28.8	23.4
	High NPI	13.3	11.6	16.2	8.8	19.4	17.7	14.3

Delayed reconstruction

Number of responders		104	176	160	35	83	175	733
Age (years)	Median	55	50.5	52	52	50	51	51
	IQR	11.5	14	14	14	11	13	14
Ethnicity / %	White	92.3	90.9	89.4	88.6	79.5	91.4	89.5
	Non-white	1.9	2.3	3.1	0.0	6.0	2.9	2.9
	Not known / not stated	5.8	6.8	7.5	11.4	14.5	5.7	7.6
Index of multiple deprivation (IMD) Quintile /	1 (least deprived)	24.5	23.2	21.2	34.4	35.0	30.2	26.5
	2	28.4	22.0	23.8	18.8	21.3	25.0	23.8
	3	18.6	20.8	19.9	12.5	17.5	19.2	19.2

%	4	15.7	19.1	14.6	15.6	17.5	12.2	15.6
	5 (most deprived)	12.8	14.9	20.5	18.8	8.8	13.4	14.9
ASA grade / %	1	66.4	59.1	70.0	77.1	51.8	68.6	64.8
	2	28.9	37.5	28.8	22.9	44.6	30.9	32.9
	3 or 4	4.8	3.4	1.3	0.0	3.6	0.6	2.3
ECOG Status / %	0	88.5	90.3	91.3	91.4	92.8	95.4	91.8
	1	7.7	9.7	7.5	8.6	7.2	4.6	7.4
	2 plus	3.9	0.0	1.3	0.0	0.0	0.0	0.8
Diabetes status / %	Diabetic	1.9	2.3	3.8	2.9	1.2	3.4	2.7
Tumour burden / %	DCIS	15.8	8.5	19.2	6.9	8.2	11.7	12.6
	Low NPI	37.6	25.5	19.9	31.0	30.1	29.2	31.4
	Intermediate NPI	26.7	38.2	29.8	31.0	30.1	29.2	31.4
	High NPI	19.8	27.9	31.1	31.0	37.0	32.5	29.6

Table 3 – Adjusted mean difference in Breast-Q scores by type of reconstruction, with expander / implant as reference, for each immediate and delayed reconstruction procedure type

<u>Immediate reconstruction patients</u>

Reconstruction type	Breast area appearance		Emotional well being		Physical well being		Sexual well being		Satisfaction with outcome	
	Mean diff	95% CI	Mean diff	95% CI	Mean diff	95% CI	Mean diff	95% CI	Mean diff	95% CI
Expander/implant	Reference									
LD with implant or expander	8.2	5.3 to 11.1	7.7	4.6 to 10.8	0.7	-1.5 to 2.9	5.8	2.0 to 9.6	6.2	2.3 to 10.1
Autologous LD	8.6	5.4 to 11.8	4.9	1.8 to 8.0	-1.4	-3.2 to 0.4	4.3	0.2 to 8.4	3.0	-0.6 to 6.6
Pedicled TRAM	21.2	11.0 to 31.4	14.1	6.8 to 21.4	4.6	0.4 to 8.8	21.1	7.2 to 35.0	17.5	10.7 to 24.2
Free TRAM	5.2	0.2 to 10.2	2.6	-2.6 to 7.8	-1.3	-4.6 to 1.9	5.1	-2.1 to 12.4	3.5	-1.3 to 8.4
DIEP/SIEA	9.3	5.6 to 13.1	5.9	2.2 to 9.6	1.4	-0.5 to 3.4	7.1	2.3 to 11.8	4.3	-0.8 to 9.3
Wald test of difference – p-value	$p<0.0001$		$p<0.0001$		$p=0.0074$		$p=0.0030$		$p=0.0001$	

Table 3 (continued)

<u>Delayed reconstruction patients</u>

Reconstruction type	Breast area appearance		Emotional well being		Physical well being		Sexual well being		Satisfaction with outcome	
	Mean diff	95% CI	Mean diff	95% CI	Mean diff	95% CI	Mean diff	95% CI	Mean diff	95% CI
Expander/implant	Reference									
LD with implant or expander	9.5	4.7 to 14.3	8.2	2.6 to 13.8	-1.4	-6.4 to 3.6	7.1	-1.9 to 16.1	7.9	2.1 to 13.6
Autologous LD	11.7	6.1 to 17.2	7.1	0.7 to 13.6	-0.4	-4.3 to 3.6	4.6	-5.2 to 14.4	8.0	1.0 to 15.0
Pedicled TRAM	9.4	2.6 to 16.2	6.2	-2.6 to 15.0	-1.7	-8.7 to 5.3	5.7	-7.8 to 19.1	2.8	-6.3 to 11.8
Free TRAM	19.2	12.9 to 25.5	10.8	3.3 to 18.4	-0.1	-4.9 to 4.6	15.3	3.6 to 27.1	14.9	8.1 to 21.8
DIEP/SIEA	13.8	8.2 to 19.5	12.1	5.5 to 18.7	4.0	-0.2 to 8.2	12.0	1.0 to 23.0	13.9	7.5 to 20.2
Wald test of difference – p-value		$p<0.0001$		$p=0.0054$		$p=0.0084$		$p=0.0211$		$p<0.0001$

Figure 1 – Forest plot of unadjusted mean scores for each Breast-Q scale following immediate and delayed reconstruction, with associated 95% confidence intervals