

Clinical Outcomes in Reduction Mammoplasty: A Systematic Review and Meta-analysis of Published Studies

ELENIE B. CHADBOURNE, MD; SHUHUAN ZHANG, MS; MATTHEW J. GORDON, BA; ERIC Y. RO, BA; SUSAN D. ROSS, MD, FRCPC; PAUL L. SCHNUR, MD; AND PETRA R. SCHNEIDER-REDDEN, MD

This systematic review and meta-analysis were undertaken to determine whether reduction mammoplasty improves measurable outcomes in women with breast hypertrophy. A systematic review of the literature in 5 languages from 1985 until March 1999 was performed, and data were compared for meta-analysis. Eligible studies were both experimental and observational and involved women with preoperative physical and/or psychosocial signs and symptoms who underwent reduction mammoplasty for breast hypertrophy. Outcomes assessed were postoperative physical signs and symptoms such as shoulder pain, shoulder (bra strap) grooving, and quality-of-life domains, such as physical and psychological functioning, and were expressed primarily as risk differences (RDs). Twenty-nine studies of 4173 patients met all eligibility criteria. Reduction mammoplasty was associated with a statistically significant improvement in physical signs and symptoms involving shoulder pain (RD, 0.71 [95% confidence interval (CI), 0.62-0.80]); shoulder grooving (RD, 0.69 [95% CI, 0.60-0.78]); upper/lower back pain (RD, 0.59 [95% CI, 0.48-0.70]); neck pain (RD, 0.50 [95% CI, 0.37-0.64]); in-

tertrigo (RD, 0.44 [95% CI, 0.34-0.54]); breast pain (RD, 0.36 [95% CI, 0.17-0.55]); headache (RD, 0.28 [95% CI, 0.11-0.46]); and pain/numbness in the hands (RD, 0.11 [95% CI, 0.04-0.18]). The quality-of-life parameter of physical functioning was also statistically significant (RD, 0.58 [95% CI, 0.44-0.71]), while psychological functioning was not significant (RD, 0.46 [95% CI, 0.00-1.00]). The evidence suggests that women undergoing reduction mammoplasty for breast hypertrophy have significant postoperative improvement in preoperative signs and symptoms, quality of life, or both.

Mayo Clin Proc. 2001;76:503-510

BH = breast hypertrophy; CI = confidence interval; ES = experimental studies; FEM = fixed effects model; k = number of studies; n = number of patients; NHS = National Health Service; nRCT = nonrandomized controlled trial; OS = observational studies; QOL = quality of life; RCT = randomized controlled trial; RD = risk difference; REM = random effects model; RM = reduction mammoplasty; t = number of patient groups; UCS = uncontrolled case series

Recognition of breast hypertrophy (BH) as a morbid condition has increased in recent decades. This may be due to several factors, such as improved access to medical care, increased patient awareness, and acceptance by the medical profession of unique issues pertaining to women's health. Reduction mammoplasty (RM) is a common plastic surgical procedure and is one of the most frequently performed operations on the female breast. Puosson and Verchere¹ first described a surgical procedure to reduce breast size, but it was not adopted in practice until the 1920s.² Today RM is a routine procedure for which several techniques have been described. Research in the field of RM has focused primarily on the various types of surgical techniques³⁻⁷ rather than on outcomes related to

these techniques. Only recently have studies considered the physical and psychosocial implications of RM for BH.^{1,8-16} Most of these studies have been retrospective in design, using surveys or chart reviews as information sources. One prospective study by Gonzalez et al¹⁷ compared the physical symptoms of 39 women undergoing RM with those of a control group of women with "small breasts." After surgery, their physical symptoms were statistically equivalent, or less, than the levels in the control group.

For editorial comment, see page 459.

Along with the increase in the frequency of RM has come controversy regarding reimbursement of this procedure by third-party payers. Although the distinction between "reconstructive" vs "cosmetic" surgery has been defined,¹⁸ the determination of what is medically necessary, and thus reconstructive, is less clear. One of the most common methods that third-party payers apply to determine the medical necessity for RM is to require the removal of at least 350 g of breast tissue per breast. Some insurance companies have also required that patients be no more than 10% above their ideal body weight for reim-

From MetaWorks Inc, Medford, Mass (E.B.C., S.Z., M.J.G., E.Y.R., S.D.R.); Division of Plastic and Reconstructive Surgery, Mayo Clinic, Scottsdale, Ariz (P.L.S.); and American Society of Plastic Surgeons, Arlington Heights, Ill (P.R.S.-R.).

Supported in part by the American Society of Plastic Surgeons.

Address reprint requests and correspondence to Susan D. Ross, MD, MetaWorks Inc, 10 President's Landing, Medford, MA 02155 (e-mail: SDR@Metawork.com).

bursement regardless of the patient's physical or psychosocial complaints resulting from the BH.¹² Frequently, the patient's breast size in relation to her body proportion plays a significant role in physical symptoms, and therefore, body weight and breast tissue resection weight may not be adequate indicators of improved outcomes after RM. Schnur et al¹⁹ evaluated the relationship between body surface area and breast resection weight as possible indicators for medical necessity. Seitchik²⁰ was unsuccessful at creating a formula incorporating body weight, specimen weight, and height to determine medical necessity.

Efforts to validate objectively the medical necessity for RM have been difficult given that physical and psychosocial morbidity in BH can be very subjective. The physical signs and symptoms noted by patients with BH include headache, shoulder pain, upper and lower back pain, intertrigo, and shoulder (bra strap) grooving. Psychosocial effects include difficulty in participating in sports, difficulty finding clothes to fit, and poor self-image. The present study is a systematic literature review and meta-analysis of all published observational and experimental studies reporting discrete clinical outcomes after RM to determine whether RM is associated with measurable improvement.

METHODS

A study protocol was designed prospectively to define the study objectives, search parameters, study eligibility criteria, key data elements for extraction, and analytical methods to be implemented.

Literature Search

A MEDLARS search was conducted using the following proximity expressions: *breast adj reduc*: or *reduc: adj mammoplasty* or *breast adj hypertrophy*. Also used were *macromastia* (textword); *hypertrophy* with subheading *surgery* combined with *mammoplasty* (MeSH and textword). All years from 1985 through March 26, 1999, were searched with the inclusion of the following languages: English, Italian, Spanish, German, and French. A CD-ROM version of *Current Contents* was also searched from 1994 through March 26, 1999, using the following terms: *macromastia* and/or *reduction mammoplasty* and/or *breast reduction* and/or *breast hypertrophy*. The computer-based search was supplemented by manual searches of bibliographies of all retrieved studies and review articles to detect other potentially acceptable studies. Abstracts and/or full articles were obtained for each citation considered possibly relevant.

Study Selection

Two levels of study screening were used. The first level involved the rejection of all abstracts that contained any

one of the following exclusion criteria: (1) case reports, abstracts only, letters, comments, and reviews; (2) studies investigating surgical procedures in which clinical outcomes of interest were not reported; (3) mixed surgical procedures (ie, augmentation and reduction); (4) virginal breast hypertrophy; (5) gigantomastia secondary to pregnancy; (6) RM in the contralateral breast in the setting of breast cancer; or (7) mixed gender studies. At the second level of screening, full articles of all accepted abstracts were then retrieved and screened for the presence of all the following inclusion criteria: (1) minimum of 10 patients per study; (2) experimental study designs (randomized controlled trials [RCTs], nonrandomized controlled trials [nRCTs], uncontrolled case series [UCS]) and observational studies (historical cohorts and cross-sectional surveys); (3) diagnosis of BH or macromastia, either unilateral or bilateral; (4) RM as the surgical procedure of interest; and (5) clinical outcomes of interest.

Data Extraction

Studies were blinded as to authorship, journal, and funding source, and data extraction was performed by 2 reviewers. The key data elements that were sought for each study were categorized as (1) study characteristics, (2) patient characteristics, (3) treatment characteristics, and (4) clinical outcomes. Clinical outcomes were captured as pre- and postoperative values according to signs/symptoms, coincident nonsurgical treatment modalities, and quality of life (QOL). In addition, studies were assigned a level of evidence score.²¹

Statistical Methods

The primary objective of the analysis was to examine the clinical efficacy outcomes of RM. Descriptive statistics were used to summarize all study, patient, treatment, and outcome characteristics by applying means weighted for study size. For patient and treatment summary tables, the number of patients enrolled or analyzed for efficacy were used as denominators (see footnotes of summary tables). For outcome summary tables, the number of patients analyzed for efficacy was used as the denominator. When the number of patients analyzed for efficacy was missing, the number of patients enrolled was substituted for the missing value. A conservative nominal level of $P=.01$ was used to establish statistical significance for all pre- and postoperative comparisons.

For meta-analyses, the efficacy data from multiple studies were combined by using 2 different approaches: (1) fixed effects model (FEM)²² and (2) random effects model (REM).²³ Although both FEM and REM were used in the meta-analyses, only the results from the REM are reported here, since no important differences were observed.

Based on the amount of data available and on clinical relevance, 8 efficacy and 2 QOL outcomes were analyzed with use of meta-analytic techniques. The efficacy outcomes were improvement in or resolution of headache, neck pain, shoulder pain, upper/lower back pain, shoulder grooving, pain/numbness in hands, breast pain, and intertrigo. The QOL outcomes were physical and psychological functioning.

For each efficacy and QOL outcome available, group averages for pre- and postoperative measurements were computed, and the differences expressed as a weighted risk difference (RD) with 95% confidence interval (CI).

Following the primary data analyses, sensitivity analyses were performed by examining the correlation between the study-specific outcomes of interest and study-specific covariates. These analyses were supplemented with use of general linear models. These analyses examined any heterogeneity found in the meta-analyses and examined the influence of potential covariates that were identified in the protocol: study size, study location, study design (observational or experimental), mean patient age in each study, mean patient height in each study, and mean patient weight in each study.

All calculations were performed with use of SAS software version 6.12. All summary results are expressed as means with ranges or SDs.

RESULTS

In the section that follows, study counts are designated by "k," patient groups by "t," and patient counts by "n."

Study Characteristics

The initial search yielded 530 citations. Of these, 399 abstracts were rejected during the first screening. The second level of screening of full articles resulted in further exclusion of 102 articles, with the most common reasons for rejection being (1) outcomes of interest not available; (2) abstracts, letters, comments, and reviews; and (3) language not specified. Twenty-nine studies satisfied all criteria for eligibility.^{1,8,9,12,14-17,20,24-43}

Study characteristics for the 29 included studies are summarized in Table 1. The studies were conducted in North America (k=15), Europe (k=12), and elsewhere (Taiwan [k=1] and Qatar [k=1]). Reporting languages were English (k=27), German (k=1), and French (k=1), and publication dates extended from 1986 to 1998.

Studies were categorized by study design as either observational studies (OS) or experimental studies (ES). The results that follow are subgrouped according to these categories so as not to obscure any discrepancies that might be present on the basis of study design. There were 18 OS^{8,9,12,15,16,20,24,26,29,32-39,43} consisting of cross-sectional

Table 1. Study Characteristics*

Characteristics	Observational studies [†]	Experimental studies [‡]
No. of studies	18	11
Location		
North America	12	3
Europe	5	7
Other	1	1
Language		
English	18	9
French	0	1
German	0	1
Study design		
Cross-sectional	17	0
Historical cohort	1	0
Uncontrolled case series	0	8
Nonrandomized controlled trial	0	3
Level of evidence ²¹		
I	0	0
II	0	0
III	18	11
IV	0	0
V	0	0
Preoperative information obtained [§]		
Before surgery	4	10
After surgery	13	0
Mean No. of charts reviewed (k)	304 (10)	NA
Mean No. of patients to receive survey (k)	219 (17)	NA
Mean No. of respondents (k)	160 (18)	NA
Mean ± SD time for survey sent postoperatively (mo) (k)	45±29 (12)	NA
Mean ± SD follow-up time (mo) (k)	NA	12±16 (8)

*(k) = number of studies; NA = not available.

[†]References 8, 9, 12, 15, 16, 20, 24, 26, 29, 32-39, and 43.

[‡]References 1, 14, 17, 25, 27, 28, 30, 31, and 40-42.

[§]Two studies (1 observational, 1 experimental) did not report when preoperative information was obtained.

(k=17) and historical (k=1) designs, and all were reported in English. Preoperative signs/symptoms and QOL information were obtained either prior to or after RM and typically in the form of surveys. Four OS^{8,12,35,43} obtained preoperative information prior to surgery, 13 recorded preoperative information on postoperative surveys, and 1 was not reported. The mean ± SD number of months that sur-

Table 2. Patient Characteristics*

Characteristics	Observational studies	Experimental studies
No. of patients enrolled (k)	3605 (18)	568 (11)
No. of patients analyzed for efficacy (k)	2639 (17)	371 (10)
No. of patients analyzed for safety (k)	2332 (14)	355 (8)
No. of patients analyzed for quality of life (k)	2462 (16)	291 (7)
Mean age (y)	35.8	35.4
Range (k)	11.0-86.0 (13)	14.0-72.0 (9)
Mean height (cm)	160.6	161.8
Range (k)	142.5-189.0 (5)	147.0-178.0 (2)
Mean weight (kg)	73.9	71.6
Range (k)	43.2-134.1 (7)	49.0-100.0 (2)
Mean body mass index (kg/m ²)	27.5	29.6
Range (k)	26.4-28.0 (2)	27.6-30.0 (2)
Diagnosis of breast hypertrophy† (mean %)	82.9	94.8
Range (k)	47.2-100.0 (10)	74.0-100.0 (5)
(n)	(1654)	(250)
Breast asymmetry† (mean %)	15.1	19.5
Range (k)	2.6-28.0 (3)	14.0-25.0 (2)
(n)	(128)	(11)

* (k) = number of studies; (n) = number of patients.

† Values derived using the number of patients enrolled.

veys were sent postoperatively was 45 ± 29 months, and the survey period ranged from 6 to 92 months. One study³³ was not included in this calculation because only a range was provided for the number of months, not a mean. Other study characteristics such as average number of charts reviewed, the number of patients receiving the survey, the number of respondents, and the number of studies reporting these characteristics are displayed in Table 1.

For the 11 ES,^{1,14,17,25,27,28,30,31,40-42} 8 were UCS and 3 were nRCTs. English (k=9), French (k=1), and German (k=1) were the reporting languages in this group. Postoperative follow-up, at which time surveys were sent, was noted in 8 studies and occurred from 1 to 42 months after surgery with a mean \pm SD follow-up time of 12 ± 16

months. Ten of the 11 ES captured preoperative information before surgery, while 1 did not record timing of collection.

Patient Characteristics

The study set included 4173 patients who were distributed between the OS (n=3605) and the ES (n=568). Demographic features of patients in these studies are displayed in Table 2. The number analyzed for efficacy was approximately 25% lower than the number enrolled, due to unexplained loss of patients to follow-up for the specific outcomes assessed. The mean age in each of these study designs was similar among the 2 study designs: OS, 35.8 years, and ES, 35.4 years. Height, weight, and body mass index (BMI) were also similar. Breast asymmetry was noted in 3 OS (mean 15.1% of patients per study) and 2 ES (mean 19.5% of patients per study).

Treatment Characteristics

Treatment characteristics are displayed in Tables 3 and 4. The surgical specialty performing RM in almost all studies was plastic surgery (k=25), and 1 study reported a general surgeon. Three studies did not report the surgical specialty. Various operative techniques were used. The most frequent operative technique was the inferior pedicle in the OS (k=8) and the Strombeck in the ES (k=2). A hospital surgical setting was reported in 2261 OS patients and 518 ES patients, and an ambulatory setting was reported for 189 OS patients. None of the ES patients had surgery performed in an ambulatory setting. The numbers of patients with a surgical setting not reported were 1155 for OS and 50 for ES.

The mean weight of total breast tissue resected was similar in both designs (OS, 1429.4 g, and ES, 1567.2 g) with 93.3% of the OS patients and 94.1% of the ES patients undergoing bilateral RM (Table 4).

Outcomes

Table 5 displays both pre- and postoperative results for signs and symptoms. The mean sternal notch-to-nipple distance was an additional outcome reported in 2 ES,^{17,27} which reported a mean \pm SD distance of 27.4 ± 34.5 cm preoperatively and 21.2 ± 5.8 cm postoperatively.^{12,14,35,40}

The QOL domains of social functioning, overall life satisfaction/well-being, and perception of health status were documented infrequently preoperatively and were even less so postoperatively. Four studies used various QOL scales in addition to the 5 domains. No specific information can be obtained from these 4 studies, given that each one used a different type of scale. In general, however, all scales reflected overall improvement.

Table 3. Treatment Characteristics*

Treatment	Observational studies	Experimental studies
Operative technique		
Free nipple-graft	3	0
Inferior pedicle	8	1
Skoog	4	0
Strombeck	4	2
Other		
Amputation	2	0
Central parenchymal pedicle	1	1
Free nipple-areola	1	0
Inferior lateral	1	0
Lejour vertical mammoplasty	0	1
McKissock	2	0
Modified McKissock	0	1
Regnault	1	0
S approach	0	1
Superior	1	0
Vertical pedicle	2	0
Wise pattern	1	1

*The number of studies reporting the operative technique used is greater than the total number of studies because some studies reported use of more than 1 technique.

Meta-analytic Results

The outcomes with sufficient data that were available for meta-analysis were postoperative signs/symptoms and 2 of 5 QOL categories. The outcomes were expressed as RDs and their respective 95% CIs, comparing pre- and postoperative data. Sensitivity analysis comparing results by study design did not reveal a significant difference, except for shoulder grooving. Therefore, meta-analysis was performed after pooling of results from the 2 categories of study designs.

All REM and FEM meta-analytic results were similar; therefore, only the REM is reported here. Denominators of numbers analyzed for efficacy were used for all RD calculations and were the same for pre- and postoperative values except in 1 study³⁵ in which 166 patients were enrolled, 128 completed the preoperative survey, and 58 completed the postoperative survey.

Signs/Symptoms

The preoperative measures per study were considered as the comparison group for all postoperative results. The 3 nRCTs consisted of control groups with female volunteers without BH who did not undergo surgery.^{14,17,31} These were not incorporated into any analyses since only these 3 studies had such controls. One of these nRCTs¹⁷ showed improvement in signs/symptoms after surgery, while signs/symptoms in the control group remained unchanged. The second

Table 4. Treatment Characteristics*

Characteristics	Observational studies	Experimental studies
Surgical setting†		
Hospital (% of patients)	73.6	100.0
Range	0.0-100.0	100.0-100.0
(k)	(14)	(10)
(n)	(2261)	(518)
Ambulatory (% of patients)	4.7	0.0
Range	0.0-64.9	0.0-0.0
(k)	(14)	(10)
(n)	(189)	(0)
Mean tissue resection weight‡ (g)		
Right	840.1	639.3
Range	120.0-4200.0	145.0-2005.0
(k)	(7)	(6)
Left	839.6	643.9
Range	120.0-4200.0	15.0-2150.0
(k)	(7)	(6)
Total	1429.4	1567.2
Range	100.0-8132.0	250.0-4800.0
(k)	(7)	(6)
Type of mammoplasty†		
Unilateral (% of patients)	2.8	2.8
Range	0.0-7.1	0.0-14.0
(k)	(5)	(5)
(n)	(44)	(7)
Bilateral (% of patients)	93.3	94.1
Range	87.5-100.0	84.6-100.0
(k)	(5)	(5)
(n)	(1286)	(161)
Surgical specialty‡		
Plastic surgery	17	8
General surgery	0	1

*(k) = number of studies; (n) = number of patients.

†Values derived using the number of patients enrolled.

‡Three studies did not report a specialty.

nRCT³¹ described minimal symptoms and upper/lower back pain, which showed improvement after surgery in the study group and no change in the control group. The third nRCT did not report any findings in the control group but did reveal an improvement in the specific scale used to measure QOL in the group undergoing RM. The RDs (95% CIs) for the 8 signs/symptoms experienced by the patient populations in all studies are summarized in Table 6. All 8 outcomes were statistically significant for the reduction or elimination postoperatively of signs/symptoms experienced preoperatively.

Quality of Life

Investigators reported QOL outcomes using various descriptive terms in the studies evaluated. Four studies^{12,14,35,40}

Table 5. Frequency of Pre- and Postoperative Signs and Symptoms*

Signs/symptoms	Preoperative	Postoperative
Headache	33.7	8.8
Range	2.0-88.7	0.0-27.3
(k)	(7)	(7)
(n/N)	(481/1427)	(125/1427)
Neck pain	65.4	9.7
Range	30.9-94.0	2.2-27.3
(k)	(11)	(11)
(n/N)	(1035/1582)	(153/1582)
Shoulder pain	79.7	6.0
Range	47.5-98.5	1.5-24.2
(k)	(12)	(12)
(n/N)	(1547/1829)	(109/1829)
Upper/lower back pain	72.9	13.0
Range	45.7-91.8	2.7-33.3
(k)	(11)	(11)
(n/N)	(841/1153)	(150/1153)
Shoulder grooving	84.1	7.6
Range	20.0-95.5	0.0-24.2
(k)	(12)	(12)
(n/N)	(1546/1838)	(139/1838)
Pain/numbness in hands	18.6	6.5
Range	11.1-22.6	2.7-9.1
(k)	(4)	(4)
(n/N)	(174/934)	(61/934)
Breast pain	50.1	13.0
Range	6.0-90.2	1.3-45.5
(k)	(8)	(8)
(n/N)	(683/1364)	(178/1364)
Intertrigo	50.3	4.4
Range	22.0-77.4	1.2-17.3
(k)	(10)	(10)
(n/N)	(761/1513)	(66/1513)

*Frequency is reported as the mean percentage of patients. Values were derived using the number of patients analyzed for efficacy. (k) = number of studies; (n/N) = number of patients with this sign or symptom/number of patients assessed both pre- and postoperatively for this sign or symptom.

also administered specific QOL scales to assess more rigorously this outcome. Types of scales used included the Body Dysmorphic Disorder Examination Self-report, Short-Form 36 Questionnaire, General Health Questionnaire, and Rosenberg's Self-esteem scale. To analyze QOL, 5 domains of interest⁴⁴ were identified prior to data extraction, and results were captured according to these domains. The 5 domains included physical functioning, psychological functioning, social functioning, overall life satisfaction/well-being, and perceptions of health status.⁷ Due to a paucity of data, only physical and psychological functioning were meta-analyzed.

Physical functioning was the most frequently reported QOL domain with 15 studies* meta-analyzed reporting both preoperative and postoperative results. If a study reported more than 1 response in these domains, the response with the greatest change from preoperative to postoperative, ie, the best response, was used in calculation of the RD. The RD and 95% CI for physical functioning based on the 15 studies referenced above was 0.58 (0.44-0.71) and for psychological functioning based on 4 studies^{16,24,27,34} was 0.46 (0.00-1.00).

Sensitivity Analyses

Sensitivity analyses were performed to examine the effect of study level and patient level variables on the outcomes. The following covariates identified in the protocol were included based on availability of data, sufficient spread of results, and clinical or methodologic rationale: study location, study design, number of patients enrolled, number of patients analyzed for efficacy, mean patient age, height, and weight. There was only 1 outcome, shoulder grooving, that was significantly affected by any of these covariates: study design and patient age. There was a significant study design effect ($P < .01$ for analysis of variance test) for the RD for pre- and postoperative shoulder grooving, with a higher RD in the OS group than in the ES group. Meta-analysis undertaken separately for OS and ES showed a higher overall RD of 0.77 (95% CI, 0.68-0.85) for OS vs 0.35 (95% CI, 0.15-0.56) for ES. A possible explanation for the greater RD in the OS vs ES may be that OS reporting this outcome included older patients with several more years "at risk" for this structural defect to have developed. Second, the Pearson correlation coefficient and general linear model analysis showed that a significant positive relationship ($P < .01$ for correlation test) exists between the RD of shoulder grooving and mean age. The RD of shoulder grooving increased as the mean patient age of each study increased. These correlations are from study level covariates and should be interpreted with caution, as they may not represent the same relationship as correlation among individual patient data. None of the other covariates was significant in these sensitivity tests at $P = .01$ level, and therefore they were not considered to be confounding the meta-analysis results.

DISCUSSION

Meta-analysis is most valid when the studies pooled are of high quality regarding issues of randomization and blinding, and the individual studies are similar in terms of patient population, treatments, and outcomes. Hence, placebo-controlled, double-blind, randomized trials are ideal

*References 1, 9, 15-17, 24, 26, 30, 32, 36, 37, 39-41, 43.

for meta-analysis because they represent the “gold standard” for establishing efficacy of an intervention. In the surgical setting, such studies are much more difficult to realize than in the medical setting and therefore are much less common in the literature. Most of the studies in this systematic literature review were observational and retrospective in design. These represent the best available evidence on the subject of RM and physical and psychosocial outcomes. This evidence is sufficient to answer the primary question: Does RM improve measurable, discrete patient outcomes? The answer is yes. Meta-analysis reveals that improvement occurred in all signs and symptoms examined. This result is supported by other studies with other outcomes not included in this review. A study by Losee et al⁴⁵ in macromastia and bulimia demonstrated dramatic improvement in symptoms of bulimia in 4 of 5 patients after RM. Another study by Starley et al⁴¹ included in this review also reported on respiratory changes after RM. There were statistically significant improvements in peak expiratory and inspiratory flow rates and change in forced expiratory volume in 1 second after RM. Letterman and Schurter⁴⁶ have also studied the effects of BH on the skeletal system and showed improvements after RM.

Despite these results, the debate has continued regarding whether RM is cosmetic or medically necessary. Klassen et al³⁵ assessed whether RM should be rationed by the National Health Service (NHS) in England. They concluded, “The study provided empirical evidence that supports the inclusion of breast reduction surgery in NHS purchasing contracts.” A US court has ruled that RM is medically necessary.^{47,48} The findings of this systematic literature review and meta-analysis are concordant with these other assessments.

From a methodologic perspective, this review has several key strengths. First is the comprehensive electronic and manual search performed for potentially eligible clinical studies that reported on clinical outcomes after RM for BH. It is unlikely this search missed important published data on this subject. The rigorous methods used for screening and selecting studies, blinding published articles prior to data extraction, and using dual extraction by 2 investigators allowed for reduction of bias and error as much as possible. The fact that this review was carried out by scientists who were not plastic surgeons may have reduced chances for bias in selecting studies, extracting data, and interpreting results.

Limitations of the review include recall bias, given that preoperative information was obtained several months after the surgical procedure in many studies. In addition, the proportion of patients without follow-up results was high (25%) and could bias results in favor of the intervention if patients with bad outcomes were less likely to provide

Table 6. Risk Difference in Reduction Mammoplasty

Signs/symptoms	No. of studies	RD (95% CI)*
Clinical		
Headache	7	0.28 (0.11-0.46)
Neck pain	11	0.50 (0.37-0.64)
Shoulder pain	12	0.71 (0.62-0.80)
Upper/lower back pain	11	0.59 (0.48-0.70)
Shoulder grooving	12	0.69 (0.60-0.78)
Pain/numbness in hands	4	0.11 (0.04-0.18)
Breast pain	8	0.36 (0.17-0.55)
Intertrigo	10	0.44 (0.34-0.54)
Quality of life		
Physical functioning	15	0.58 (0.44-0.71)
Psychological functioning	4	0.46 (0.00-1.00)

*CI = confidence interval; RD = risk difference.

information after surgery. It may have been of interest to capture the level of patient satisfaction, which appeared to be consistently high, although this information was not specifically extracted because it was judged to reflect more an aesthetic outcome, as opposed to a clinical one. Finally, one of the main challenges of this review was to categorize subjective patient information into workable, consistent formats. Although the outcomes formats chosen in this review are somewhat arbitrary, the consistent direction and magnitude of effect on diverse signs and symptoms suggest that the benefit being measured is likely to be real.

We conclude that this systematic literature review and meta-analysis demonstrates improved clinical outcomes in women with BH who undergo RM.

REFERENCES

- Berg A, Stark B, Malec E. Reduction mammoplasty: a way helping females with neck, shoulder and back pain symptoms. *Eur J Plast Surg.* 1994;17:84-86.
- Thorek M. Possibilities in the reconstruction of the human form 1922. *Aesthetic Plast Surg.* 1989;13:55-58.
- Courtiss EH, Goldwyn RM. Reduction mammoplasty by the inferior pedicle technique: an alternative to free nipple and areola grafting for severe macromastia or extreme ptosis. *Plast Reconstr Surg.* 1977;59:500-507.
- Hester TR Jr, Bostwick J III, Miller L, Cunningham SJ. Breast reduction utilizing the maximally vascularized central breast pedicle. *Plast Reconstr Surg.* 1985;76:890-900.
- Cardoso AD, Cardoso AD, Pessanha MC, Peralta JM. Three dermal pedicles for nipple-areola complex movement in reduction of gigantomastia. *Ann Plast Surg.* 1984;12:419-427.
- Regnault P. Reduction mammoplasty by the “B” technique. *Plast Reconstr Surg.* 1974;53:19-24.
- Ramselaar JM. Precision in breast reduction. *Plast Reconstr Surg.* 1988;82:631-643.
- Boschert MT, Barone CM, Puckett CL. Outcome analysis of reduction mammoplasty. *Plast Reconstr Surg.* 1996;98:451-454.
- Davis GM, Ringler SL, Short K, Sherrick D, Bengtson BP. Reduction mammoplasty: long-term efficacy, morbidity, and patient satisfaction. *Plast Reconstr Surg.* 1995;96:1106-1110.

10. Goin MK, Goin JM, Gianini MH. The psychic consequences of a reduction mammoplasty. *Plast Reconstr Surg.* 1977;59:530-534.
11. Goin MK. Psychological reactions to surgery of the breast. *Clin Plast Surg.* 1982;9:347-354.
12. Glatt BS, Sarwer DB, O'Hara DE, Hamori C, Bucky LP, LaRossa D. A retrospective study of changes in physical symptoms and body image after reduction mammoplasty. *Plast Reconstr Surg.* 1999;103:76-82.
13. Harris DL. Self-consciousness of disproportionate breast size: a primary psychological reaction to abnormal appearance. *Br J Plast Surg.* 1983;36:191-195.
14. Hollyman JA, Lacey JH, Whitfield PJ, Wilson JS. Surgery for the psyche: a longitudinal study of women undergoing reduction mammoplasty. *Br J Plast Surg.* 1986;39:222-224.
15. Rapisis T, Zehring RD, Downey DL. Long-term functional results after reduction mammoplasty. *Ann Plast Surg.* 1995;34:113-116.
16. Schnur PL, Schnur DP, Petty PM, Hanson TJ, Weaver AL. Reduction mammoplasty: an outcome study. *Plast Reconstr Surg.* 1997;100:875-883.
17. Gonzalez F, Walton RL, Shafer B, Matory WE Jr, Borah GL. Reduction mammoplasty improves symptoms of macromastia. *Plast Reconstr Surg.* 1993;91:1270-1276.
18. Position paper: reduction mammoplasty. Prepared by the Socioeconomic Committee. Approved by the American Society of Plastic Surgeons Board of Directors, June 1994. Available at: www.plasticsurgery.org/profinfo/pospap/red.htm. Accessibility verified March 23, 2001.
19. Schnur PL, Hoehn JG, Ilstrup DM, Cahoy MJ, Chu CP. Reduction mammoplasty: cosmetic or reconstructive procedure? *Ann Plast Surg.* 1991;27:232-237.
20. Seitchik MW. Reduction mammoplasty: criteria for insurance coverage. *Plast Reconstr Surg.* 1995;95:1029-1032.
21. Cook DJ, Guyatt GH, Laupacis A, Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents [published correction appears in *Chest.* 1994;105:647]. *Chest.* 1992;102(4, suppl):305S-311S.
22. Woolf B. On estimating the relation between blood group and disease. *Ann Hum Genet.* 1955;19:252-253.
23. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7:177-188.
24. Atterhem H, Holmner S, Janson PE. Reduction mammoplasty: symptoms, complications, and late results: a retrospective study on 242 patients. *Scand J Plast Reconstr Surg Hand Surg.* 1998;32: 281-286.
25. Bellmann KP, Hinze O. Reduction plastic surgery of the female breast [in German]. *Zentralbl Chir.* 1986;111:929-934.
26. Bruhlmann Y, Tschopp H. Breast reduction improves symptoms of macromastia and has a long-lasting effect. *Ann Plast Surg.* 1998;41:240-245.
27. Chen TH, Wei FC. Evolution of the vertical reduction mammoplasty: the S approach. *Aesthetic Plast Surg.* 1997;21:97-104.
28. Cole RP, Shakespeare V. Measuring patient-based outcomes in a plastic surgery service: breast reduction surgical patients [letter]. *Br J Plast Surg.* 1998;51:79-80.
29. Dabbah A, Lehman JA Jr, Parker MG, Tantri D, Wagner DS. Reduction mammoplasty: an outcome analysis. *Ann Plast Surg.* 1995;35:337-341.
30. Evans GR, Ryan JJ. Reduction mammoplasty for the teenage patient: a critical analysis. *Aesthetic Plast Surg.* 1994;18:291-297.
31. Fouquet B, Goupille P, Rouif M, Rouif V, Mateu J, Valat JP. Breast hypertrophy and dorsolumbar spine: prognostic influences of lumbar lordosis: preliminary results [in French]. *Rev Rhum Mal Osteo-artic.* 1991;58:453-457.
32. Heddens CJ. Postoperative survey of reduction mammoplasty patients. *Plast Surg Nurs.* 1993;13:148-155.
33. Kallen R, Broome A, Muhlow A, Forsby N. Reduction mammoplasty: results of preoperative mammography and patient inquiry. *Scand J Plast Reconstr Surg.* 1986;20:303-305.
34. Kinell I, Beausang-Linder M, Ohlsen L. The effect on the preoperative symptoms and the late results of Skoog's reduction mammoplasty: a follow-up study on 149 patients. *Scand J Plast Reconstr Surg Hand Surg.* 1990;24:61-65.
35. Klassen A, Fitzpatrick R, Jenkinson C, Goodacre T. Should breast reduction surgery be rationed? a comparison of the health status of patients before and after treatment: postal questionnaire survey. *BMJ.* 1996;313:454-457.
36. Makki AS, Ghanem AA. Long-term results and patient satisfaction with reduction mammoplasty. *Ann Plast Surg.* 1998;41:370-377.
37. McMahan JD, Wolfe JA, Cromer BA, Ruberg RL. Lasting success in teenage reduction mammoplasty. *Ann Plast Surg.* 1995;35:227-231.
38. Miller AP, Zacher JB, Berggren RB, Falcone RE, Monk J. Breast reduction for symptomatic macromastia: can objective predictors for operative success be identified? *Plast Reconstr Surg.* 1995;95: 77-83.
39. Serletti JM, Reading G, Caldwell E, Wray RC. Long-term patient satisfaction following reduction mammoplasty. *Ann Plast Surg.* 1992;28:363-365.
40. Shakespeare V, Cole RP. Measuring patient-based outcomes in a plastic surgery service: breast reduction surgical patients. *Br J Plast Surg.* 1997;50:242-248.
41. Starley IF, Bryden DC, Tagari S, Mohammed P, Jones BP. An investigation into changes in lung function and the subjective medical benefits from breast reduction surgery. *Br J Plast Surg.* 1998; 51:531-534.
42. Velanovich V. Should general surgeons perform specialty procedures? an outcome experience with reduction mammoplasty. *Am Surg.* 1996;62:156-158.
43. Wallace WH, Thompson WO, Smith RA, Barraza KR, Davidson SF, Thompson JT. Reduction mammoplasty using the inferior pedicle technique. *Ann Plast Surg.* 1998;40:235-240.
44. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials.* 3rd ed. St Louis, Mo: Mosby; 1996.
45. Losee JE, Serletti JM, Kreipe RE, Caldwell EH. Reduction mammoplasty in patients with bulimia nervosa. *Ann Plast Surg.* 1997; 39:443-446.
46. Letterman G, Schurter M. The effects of mammary hypertrophy on the skeletal system. *Ann Plast Surg.* 1980;5:425-431.
47. Stott M. Court care favors reduction mammoplasty patient. *Plast Surg News.* March 1997.
48. *Bancroft v Tecumseh Products.* US District Court, ED Mich, case No. 95-40466, order dated December 20, 1996.