Discussion

The ultimate objective of all cancer strategies is to eliminate deaths from cancer, which in reality is close to impossible. However, by investigating the effectiveness of both the medical care itself and the associated activities of this medical care, cancer prevention and treatment measures can be significantly improved. The calculation of 5-year or 10-year (relative) survival rates plays an important role in measuring effectiveness and thus achieving this objective. However it is difficult to find statistics that enable easy comparisons. To start with, relative survival cancer statistics are not available in Japan at the national level. Some information has been collated from the site-specific registries that are run by academic societies in Japan (Watanabe et al., 1995). However, studies only calculate the observed survival rate and omit to calculate the relative survival rate. Another problem is that the study periods used in the survival calculations differ from study to study and therefore vary across primary sites of the cancer. This causes difficulty in determining survival rates that relate to a particular period.

In this discussion, stomach cancer is used as a representative primary site for discussion purposes because of the stomach cancer leading in cause of death in Japan. The survey found the 5-year observed survival rate for this site to be 62.3% and the relative survival rate to be 68.8%. This observed survival rate is consistent with the 5-year observed survival rate of 57.8% for patients (in whom cancer had been diagnosed from 1979 through 1982) reported by one of the site-specific registries in Japan (The Japanese Research Society for Gastric Cancer, 1995). Another study conducted by the Osaka Medical Center for Cancer and Cardiovascular Diseases, determined this survival rate to be 68.1% (relative) for in-patients (1987-1990), using the hospital cancer registry data (Tanaka, 1997).

The survival rates discussed so far are principally observed survival rates. They are calculated, in many cases, without any regard for the composition of gender, age or earliness of detection rate of the targeted patient group. It is thus difficult to compare survival rates geographically or chronologically between target groups that differ in terms of gender, age or earliness of detection rate. Even if the survival rates are calculated taking into account gender and age, another major problem arises. It is often unclear whether the calculation includes patients who died from causes other than the cancer in question. Furthermore, if the survival rate is calculated considering gender and age, in many cases the number of subjects drops dramatically, making it difficult to obtain a reliable survival rate. The relative survival rate is thus a way of eliminating these comparison problems. (Parkin, 1991).

The results of this study on JACCCs were calculated from data that contained a relatively high percentage of censored cases (9%). It is important to note that the higher rate of censored cases in this study is likely to overestimate survival, especially for patients with a less favorable prognosis.

The Study Group plans to accurately track and tabulate the relative survival rate annually and, as of 2005, they are in the process of defining guidelines for the standardization of data collection, data processing, and publication of survival rates. With these guidelines in mind, the Study Group aims to collect reliable data from participating institutions and monitor cancer survivals in future. The 5-year relative survival rate for these institutions that specialize in cancer treatment will become an index for Japanese cancer treatment.

Acknowledgements

The following institutions and doctors participated in the Survival Study Group of Japanese Association of Clinical Cancer Centers. Hokkaido Cancer Center: Yamashiro K.; Aomori Prefectural Central Hospital: Harada, Y., Murata Y.; Iwate Prefectural Central Hospital: Sasaki T.; Miyagi Cancer Center: Nagai Y.; Yamagata Prefectural Medical Center for Cancer & Life-style Related Diseases: Ikeda E., Kikuchi J.; Tochigi Cancer Center: Tominaga K.; Ibaragi Prefectural Central Hospital: Okazaki N, Itabashi M.; Gunma cancer Center: Fukuda T.; Saitama Cancer Center: Sekine T, Tabei T.; The Cancer Institute Hospital: Nakajima S., Hayashi I.; National Cancer Center: Koshiji M.; Tokyo Metropolitan Komagome Hospital: Ishiwata J., Mori T.; Niigata Cancer Center: Sasaki J.; Aichi Cancer Center: Ohasi K., Fuwa N.; Nagoya Medical Center: Kondo K.; Fukui Medical Center for Geriatric Diseases: Hosokawa O.; Shiga Medical Center for Geriatric Diseases: Nishimoto H.; Osaka Medical Center for Cancer & Cardiovascular Diseases: Kuroda T., Saji F.; Hyogo Medical Center for Geriatric Diseases: Okawa J., Koizumi T.; National Kure Medical Center: Hada Y., Koseki M.; Yamaguchi Grand Medical Center; Shikoku Cancer Center: Tanimizu M., Kawamura S.; Kyushu Cancer Center: Baba H. This work was partly funded by the Grant-in-Aid for Cancer Research (12-1, 16-2) from the Ministry of Health, Labor and Welfare of Japan.

References

Arimoto H, Kitagawa C, Arai H (1985). Cohort Survival Table. KOSEI-NO-SHIHYO 32:25-30. (http://www.ncc.go.jp/jp/ncca/cohort_table.txt) (in Japanese)

Cutler SJ, Ederer F (1958). Maximum utilization of the life table in analyzing survival. *J Chronic Dis*, 8, 699-712.

Esteve J, Benhamou E, Raymond L (1994). Statistical Methods in Cancer Research, volume IV: Descriptive Epidemiology. IARC Scientific Publications No.128. IACR, Lyon.

Japanese Research Society for Gastric Cancer (ed.) (1985). The General Rules for Gastric Cancer Study (The 11nd Edition), Kanehara & Co., Ltd, Tokyo.

Japanese Society for Cancer of the Colon and Rectum (ed.) (1989). General Rules for Clinical and Pathological Studies on Cancer of the Colon, Rectum and Anus (The 5th Edition), Kanehara & Co., Ltd, Tokyo.

Japanese Society of Obsterics and Gynecology (ed) (1988). The General Rules for Clinical and Pathological Management of

Uterine Cervical Cancer (The 2nd Edition), Kanehara & Co., Ltd, Tokyo.

Kaplan EL & Meier P (1958). Nonparametric estimation from incomplete observations. *J Am Stat Soc*, **53**, 457-81.

Okamoto N (ed.) (2004). Annual report of the study group for "Improvement for Clinical Cancer Centers in Japan". (in Japanese)

Parkin DM, Hakulinen T (1991). Analysis of survival, Cancer Registration Principles and Methods, IACR Sientific Publication No.95, pp159-176, IACR, Lyon.

Tanaka H, Tsukuma H, Okuda S et al (1997). Five-year relative survival rates of patients in whom cancer had been diagnosed at the Osaka Medical Center for Cancer and Cardiovascular Diseases. *Jpn J Cancer Clin*, **43**, 511-518.

The Japanese Breast Cancer Society (ed) (1989). General Rules for Clinical and Pathological Recording of Breast Cancer (The 10th Edition), Kanehara & Co:, Ltd, Tokyo.

The Japanese Lung Cancer Society (ed) (1989). General Rule for Clinical and Pathological Record of Lung Cancer (The 3rd Edition), Kanehara & Co., Ltd, Tokyo.

The Japanese Research Society for Gastric Cancer (1995). Treatment results of gastric cancer patients: An analysis of nationwide database. Cancer Treatment and Survival, pp47-56, Japan Scientific Societies Press, Tokyo.

Watanabe S, Tominaga S, Kakizoe T (1995). Cancer Treatment and Survival. Site-Specific Registries in Japan. Japan Scientific Societies Press, Tokyo.

Jpn J Clin Oncol 2006;36(4)197–206
doi:10.1093/jjco/hyl019

Evidence-Based Risk Factors for Seroma Formation in Breast Surgery

Katsumasa Kuroi[1], Kojiro Shimozuma[2], Tetsuya Taguchi[3], Hirohisa Imai[4], Hiroyasu Yamashiro[5], Shozo Ohsumi[6] and Shinya Saito[7]

[1]Division of Surgery and Breast Oncology, Nyuwakai Oikawa Hospital, Fukuoka, [2]Department of Healthcare and Social Services, University of Marketing and Distribution Sciences, Kobe, [3]Department of Surgical Oncology, Osaka University Graduate School of Medicine, Suita, Osaka, [4]Department of Epidemiology, National Institute of Public Health, Wako, Saitama, [5]Department of Gastroenterological Surgery, Kyoto University Graduate School of Medicine, Kyoto, [6]Department of Surgery, National Hospital Organization Shikoku Cancer Center, Matsuyama and [7]Department of Health Science, Kochi Women's University, Kochi, Japan

Received November 1, 2005; accepted December 27, 2005

Background: Seroma is a common problem in breast surgery. The aim of this systematic review was to identify risk factors for seroma formation.
Methods: Articles published in English were obtained from searches of Medline and additional references were found in the bibliographies of these articles. Risk factors were graded according to the quality and strength of evidence and to the direction of association.
Results: One meta-analysis, 51 randomized controlled trials, 7 prospective studies and 7 retrospective studies were identified. There was no risk factor supported by strong evidence, but there was moderate evidence to support a risk for seroma formation in individuals with heavier body weight, extended radical mastectomy as compared with simple mastectomy, and greater drainage volume in the initial 3 days. On the other hand, the following factors did not have a significant influence on seroma formation: duration of drainage; hormone receptor status; immobilization of the shoulder; intensity of negative suction pressure; lymph node status or lymph node positivity; number of drains; number of removed lymph nodes; previous biopsy; removal of drains on the fifth postoperative day versus when daily drainage volume fell to minimal; stage; type of drainage (closed suction versus static drainage); and use of fibrinolysis inhibitor. In contrast, sentinel lymph node biopsy reduced seroma formation. Evidence was weak, or unproven, for other factors that were commonly cited in the literature.
Conclusions: Although a number of factors have been correlated with seroma formation, strong evidence is still scarce. However, there is evidence showing that sentinel lymph node biopsy reduces seroma formation.

Key words: seroma – risk factor – breast cancer – mastectomy

INTRODUCTION

Ever since mastectomy was first carried out by Halsted in 1882, surgeons have faced several problems such as necrosis of the skin flaps, breakdown of the wound, hematoma, seroma, and infection (1). Among them, seroma, a subcutaneous collection of serous fluid, is a common problem in breast surgery. As it usually resolves within a few weeks, many surgeons view this problem as an unavoidable nuisance rather than a serious complication (1,2). However, excessive accumulation will stretch the skin and cause it to sag, resulting in patient discomfort and sometimes prolongation of hospital stay (3). To prevent seroma formation, it is important to estimate individual risk of seroma formation. In this study, we carried out a systematic review of risk factors for seroma formation.

METHODS

The primary outcome of interest was the incidence of seroma formation after breast surgery in patients with primary breast cancer. To identify published articles on seroma, a computer-assisted MEDLINE search was conducted from 1966 up to July 2005. We also searched the Cochrane Library database for relevant systematic reviews, and additional references were

For reprints and all correspondence: Katsumasa Kuroi, Division of Surgery and Breast Oncology, Nyuwakai Oikawa Hospital, 2-21-16 Hirao, Chuo-ku, Fukuoka 810-0014, Japan. E-mail: kurochan@dd.iij4u.or.jp

found in the bibliographies of these articles. The reference terms 'breast cancer', 'mastectomy', 'breast-conserving surgery', 'seroma', 'lymphocele' and 'lymphocyst' were used as both keyword and subject terms. We included meta-analysis, randomized controlled trials (RCTs), prospective studies, systematic review of RCTs or prospective studies, and retrospective studies if they included at least 200 patients. The search was limited to studies published in English, and unpublished data were not located.

Data were extracted by one reviewer and checked independently by a second. Attempts were made to contact study authors if additional information was required to adequately complete the data extraction form. The direction of association was defined as follows: increase, significant association between a factor and increase of seroma formation; decrease, significant association between a factor and decrease of seroma formation; no association, no significant association between a factor and seroma formation. When multiple studies were available on a factor, consistency, predominance and bidirectionality were considered when all showed the same direction, when there was a mixture of no association and either an increase or a decrease, and when there was a mixture of an increase and a decrease, respectively.

The quality of evidence was ranked as follows according to the 'levels of evidence and grades recommendation' of the Oxford Center for Evidence-based Medicine (4): level 1, systematic review of RCTs, and individual RCT; level 2, systematic review of cohort studies, and individual cohort study including low-quality RCT; level 3, systematic review of case–control studies, and individual case–control study; level 4, case series, and poor quality cohort and case–control studies; level 5, expert opinion without explicit critical appraisal, or based on physiology, bench research or first principles. The strength of evidence was categorized as grade A (strong), consistent level 1 studies; grade B (moderate), consistent level 2 or 3 studies, or extrapolations from level 1 studies; grade C (weak), level 4 studies or extrapolations from level 2 or 3 studies; grade D (unproven), level 5 evidence or troublingly inconsistent or inconclusive studies of any level. When there was no consistency, extrapolations were made either if there was predominance in the direction with at least two study differences or if evidence was based on a study, and troublingly inconsistent was considered if there was bidirectionality. Otherwise, the evidence was regarded as 'inconclusive'.

MAIN FEATURES

One meta-analysis (5), 51 RCTs (6–56), 7 prospective studies (57–63) and 7 retrospective studies (64–70) were eligible for formal appraisal and inclusion in this review (Table 1). Considering the quality of the RCTs, all of them except one were graded as level 2, as these were usually underpowered, and the method of random allocation and concealment and sample size justification were not described in detail. If provided, methods of random allocation and concealment were inappropriate, or quality was dubious due to multiplicity of comparison. The meta-analysis was also graded as level 2, as this was based on low-quality RCTs. Moreover, studies included were heterogeneous, and parameter estimates were not provided in several. Therefore, we did not use formal meta-analytic techniques, but provided summarized evidence by detailed systematic review of the best evidence available on seroma formation. For this, risk factors for seroma formation were subdivided into four categories: patients and tumor characteristics, surgical factors, postoperative management and nonsurgical modalities.

Patients and Tumor Characteristics

In this category, age, anemia, body mass or obesity index, body weight, breast size, diabetes mellitus, grade, histological type, hormone receptor status, hypertension, nodal status or positivity of lymph nodes (LNs), number of positive LNs, pathological tumor size, side, smoking, specimen size, specimen weight, stage, tumor location and tumor size were assessed.

Among them, two studies had found a positive association between body weight and seroma formation (31,58), and one of them had also found that hypertension is associated with an increase of seroma formation (58). In contrast, as for hormone receptor status (17,51), nodal status or positivity of LNs (17,27,58,59), and stage (17,57), studies consistently showed no association with seroma formation. Similarly, no individual study found a significant association with other factors such as presence of anemia (64) or diabetes mellitus (64), smoking (64), breast size (64), grade (51), histological type (68), pathological tumor size (51), side (51), specimen weight or size (69) and tumor location (58). On the other hand, existing evidence was inconclusive for age (17,27,31,32,51,57,58), body mass index or obesity index (30,31), number of positive LNs (13,16,51,61) and tumor size (51,58).

Surgical Factors

This category was further subdivided into extent of mastectomy, wound drainage, surgical devices, suture flap fixation and miscellaneous.

Extent of Mastectomy

With respect to the extent of mastectomy, two studies have demonstrated that extended radical mastectomy increases seroma formation as compared with simple mastectomy (64,65). In contrast, one study has indicated that immediate reconstruction following MRM decreases seroma formation (69). However, no association was found between preservation or removal of the pectoral fascia and seroma formation (49), and association was inconclusive when radical mastectomy was compared with modified radical mastectomy (MRM) (64,65) and was bidirectional among six studies comparing MRM and breast-conserving surgery (16,31,33,51,57,59).

With respect to axillary dissection, four studies have consistently indicated that the number of removed LNs does not influence seroma formation (17,31,51,59). Similarly, one study has demonstrated that the extent of axillary dissection does not affect seroma formation (68). On the other hand, an

Table 1. Summary of risk factors of seroma formation

Author, year	Level of evidence	Study type	Sample size	Type of mastectomy	Intervention and seroma formation (%)		Factors and direction of association*
Morris, 1973 (6)	2	RCT	53	Rad	Suction drainage versus Static drainage 29 versus 12 (NS)	→	Type of drainage
Say, 1974 (64)	3	Retrospective	1551	Rad	–	↑	Older age, heavier body weight, surgeon (resident), multiple blood transfusions, unstented dressing, without skin graft, diagonal skin incision (versus vertical)
						→	Anemia, breast size, operation time, drain type (Penrose versus catheter drains), pressure dressing, radiation
				Ext, Rad, MRM, simple		↑	Type of surgery (Ext > Simple)
						→	Type of surgery (Rad versus MRM, Simple, Ext versus Rad)
Bourke, 1976 (7)	2	RCT	51	Simple	Suction drainage versus Corrugated drainage 8 versus 7 (NS)	→	Type of drainage
Britton, 1979 (8)	2	RCT	46	Simple	High vacuum evacuated bottle type versus Low vacuum bellows type as suction drainage unit (NS)	→	Type of drainage
Flew, 1979 (9)	2	RCT	64	Rad	Immobilization using a triangular bandage for 7 days versus Exercise from second POD 7 versus 20 (NS)	→	Immobilization
Aitken, 1984 (65)	3	Retrospective	204	Rad, MRM, Simple with suture flap fixation	–	↑	Type of mastectomy (Rad > MRM, Simple)
Corneillie, 1984 (66)	3	Retrospective	606	Rad, Mod, Simple, BCS	–	→	LN status
Tejler, 1985 (57)	2	Prospective	385	MRM, BCS	–	↑	Age, drainage volume in the 24 hours before removal (>50ml/day), total drainage volume, surgeon
						→	Stage, biopsy, intensity of negative suction pressure, duration of drainage, type of mastectomy
Cameron, 1988 (10)	2	RCT	40	BCS	Suction drain versus No drain 45 versus 4 ($P < 0.04$)	↑	No drainage
Dawson, 1989 (11)	2	RCT	100	MRM	Immobilization using a sling for 5days vs Exercise from first POD 43 versus 59 (NS)	→	Immobilization
Jansen, 1990 (12)	2	RCT	144	MRM, BCS	Shoulder exercise from first POD versus From eighth POD (NS)	→	Timing of shoulder movement
Petrek, 1990 (13)	2	RCT	57	MRM, BCS	Arm mobilization from second POD versus From fifth POD (NS)	↑	Number of positive LNs
						→	Timing of shoulder movement
Inwang, 1991 (14)	2	RCT	84	BCS	Drain removal at fifth POD versus When < 20 ml/2 days (Usually 10th to 14 th POD) 49 versus 28 (NS)	→	Timing of drain removal, total drainage volume
Vinton, 1991 (67)	3	Retrospective	560	MRM, BCS	–	↑	Type of mastectomy (MRM > BCS), drainage volume in the 24 hours before removal, age in MRM
						→	Smoking, obesity, age in BCS
Chilson, 1992 (68)	3	Retrospective	351	MRM with or without suture flap fixation	–	→	Age, histological type, tumor size, axillary LN involvement, extent of LN dissection, skin graft
						↓	Suture flap fixation
Parikh, 1992 (15)	2	RCT	100	MRM	Drain removal at third POD versus At sixth POD (NS)	→	Timing of drain removal
Petrek, 1992 ((16)	2	RCT	65	MRM, BCS	Single drain versus Multiple drains (NS)	→	Number of drain, biopsy, number of involved LNs, type of mastectomy
Somers, 1992 (17)	2	RCT	227	BCS	Suction drain versus No drain 89.1 versus 73.1 ($P = 0.008$)	↑	No drainage
						→	Stage, number of removed LNs, positivity of LNs, blood loss, type of anesthesia (local versus general), ER or PgR status, age

Table 1. *Continued*

Author, year	Level of evidence	Study type	Sample size	Type of mastectomy	Intervention and seroma formation (%)		Factors and direction of association*
Terrell, 1992 (18)	2	RCT	84	MRM	Axillary drainage versus Axillary and pectoral drainage 13 versus 19 (NS)	→	Number of drain
Coveney, 1993 (19)	2	RCT	39	Mastectomy	Suture flap fixation versus Conventional skin closure 25 versus 85 ($P < 0.001$)	↓	Suture flap fixation
Udén, 1993 (20)	2	RCT	68	MRM	Use of fibrin glue versus None 64 versus 53 (NS)	→	Use of fibrin glue, surgeon
Wyman, 1993 (21)	2	RCT	40	MRM	Use of laser scalpel versus Scalpel 50 versus 55 (NS)	→	Use of laser scalpel
Oertli, 1994 (22)	2	RCT	160	MRM, BCS	Use of tranexamic acid versus None 27 versus 37 (NS)	→	Use of tranexamic acid
Whitfield, 1994 (23)	2	RCT	50	Mastectomy	Suction drainage versus Siphon drainage 20 versus 19 (NS)	→	Type of drainage
Kumar, 1995 (58)	2	Prospective	64	Rad	–	↑	Age (older), body weight (heavier), hypertension, total drainage volume (greater)
						→	Tumor size, tumor location, nodal status, blood transfusion, diabetes mellitus
Forouchi, 1995 (24)	2	RCT	78	MRM	Neoadjuvant therapy versus Immediate surgery 20 versus 33 (NS)	→	Neoadjuvant therapy
Medl, 1995 (59)	2	Prospective	80	MRM, BCS	Use of fibrin glue versus None	→	Use of fibrin glue, type of mastectomy, number of removed LNs, LN status
Yii, 1995 (60)	2	Prospective	100	MRM, Simple, BCS, Ex+Ax,Ex	Short stay (drain removal after 48 h and discharge) versus Long stay (drain removal discharge when drainage was considered acceptable)	↑	Drainage volume in the 24 hours before removal (>50 ml/day)
						→	Timing of drain removal (removal on 2 POD versus longer)
van Heurn, 1995 (25)	2	RCT	78	BCS	Low (150 g/cm^2) versus High (700 g/cm^2) vacuum drainage 30 versus 42 (NS)	→	Intensity of negative suction pressure
Vaxman, 1995 (26)	2	RCT	40	MRM, BCS	Use of fibrin glue versus None 20 versus 5 ($P < 0.05$)	↑	Use of fibrin glue
Browse, 1996 (27)	2	RCT	67	BCS	Immobilization using a collar and cuff for 10 days versus Free 31 versus 43 (NS)	→	Immobilization, nodal status, age
Kerin, 1996 (28)	2	RCT	50	Mastectomy	Use of argon diathermy versus Scalpel, scissors and electrocautery in dissection of flaps and bleeding control 12 versus 17 (NS)	→	Use of argon diathermy
Ackroyd, 1997 (29)	2	RCT	120	MRM, BCS	Removal of drain at fifth POD versus When <30 ml/day 29 versus 25 (NS)	↑	Total drainage volume in the 5 day removal group (greater)
						→	Timing of drain removal, drainage volume in the 24 hours before removal, total drainage volume in the <30ml/day group
Bonnema, 1997 (30)	2	RCT	141	MRM, BCS	Low (15.3 kPa) versus High (95.9 kPa) negative suction pressure (NS)	↑	BMI (greater)
						→	Intensity of negative suction pressure, total drainage volume of the initial 5 POD
Burak, 1997 (31)	2	RCT	101	MRM, BCS	Use of bovine thrombin versus None (with suture flap fixation) 37 versus 40 (NS)	↑	Type of mastectomy (BCS > MRM), total drainage volume during initial 3 days (larger), age (older), weight (heavier) (multivariate analysis)
						→	Use of bovine thrombin, duration of drainage, obesity index, number of removed LNs (multivariate analysis)

Schultz, 1997 (32)	2	RCT	163	MRM	Shoulder exercise from first POD versus From seventh POD	↑	Early shoulder movement, age (older), operation time (longer)
						→	Surgeon (staff = resident)
Abe, 1998 (33)	2	RCT	116	Rad, MRM, BCS	Shoulder movement from first POD versus From seventh POD	↑	Early shoulder movement
						→	Type of mastectomy (MRM = BCS)
Chen, 1998 (34)	2	RCT	40	MRM, BCS	Use of pressure garment for 14 days versus None 0 versus 4.8 (NS)	→	Use of pressure garment
Gilly, 1998 (35)	2	RCT	108	MRM, BCS	Use of fibrin glue versus None 2 versus 1.7 (NS)	→	Use of fibrin glue
Porter, 1998 (36)	2	RCT	80	MRM, simple	Use of electrocautery versus Scalpel in dissection of flaps, fascia (bleeding control with cautery was allowed) 38 versus 13 ($P = 0.01$)	↑	Use of electrocautery, type of mastectomy (MRM > BCS), Use of multiple holes type drain (versus multiple channels type drain), lower blood-loss
Zavotsky, 1998 (37)	2	RCT	46	BCS	Suction drain versus No drain 64 versus 0	↑	No drainage
Chen, 1999 (38)		RCT	344	MRM	Shoulder movement from third POD versus From sixth POD versus After drain removal NS	→	Timing of shoulder movement
Kopelman, 1999 (39)	2	RCT	90	Simple, BCS	Removal of drain at third POD versus When < 35 ml/day 21 versus 4 ($P < 0.02$)	↑	Drain removal at third POD, total drainage volume (larger) in the group of 3 day removal
O'Hea, 1999 (40)	2	RCT	135	MRM, BCS	Use of compression dressing versus Conventional dressing ($P < 0.01$)	↑	Use of compression dressing
Dinsmore, 2000 (41)	2	RCT	27	MRM	Use of fibrin glue versus None 43 versus 23 (NS)	→	Use of fibrin glue
Woodworth, 2000 (69)	3	Retrospective	252	MRM, BCS	–	↑	Surgeon, type of mastectomy (MRM, Bp > MRM+IR), neoadjuvant chemotherapy, duration of drainage (longer)
						→	Age, BMI, specimen weight, specimen size, tumor size, number of removed LNs, number of positive LN, blood loss, biopsy
Berger, 2001 (42)	2	RCT	60	MRM, BCS	Use of fibrin glue versus None 39 versus 42 (NS)	→	Use of fibrin glue
Gupta, 2001 (43)	2	RCT	121	MRM	Removal of drain at fifth POD versus At eighth POD (Drains were removed when 30 ml/ two days>) 48 versus 28 ($P = 0.026$)	↑	Removal of drain at fifth POD
Moore, 2001 (44)	2	RCT	79	MRM, BCS	Use of fibrin glue (4, 8, or 16 ml) versus None 16 versus 21 versus 29 versus 29 (NS)	→	Use of fibrin glue
Purshotham, 2002 (45)	2	RCT	375	Mastectomy, BCS	No drainage with suture flap fixation versus Drainage without suture flap fixation 61 versus 55 (NS) in MRM, 47 versus 51 (NS) in BCS	→	No drainage with suture flap fixation versus drainage without suture flap fixation
Schuijtvlot, 2002 (61)	2	Prospective	97	BCS without drainage	Suture flap fixation (buttress suture) without drainage versus Conventional surgery	→	Number of positive LNs
						↓	Suture flap fixation
Talbot, 2002 (62)	2	Prospective	90	Mastectomy, BCS	Prolonged suction drainage versus Short drainage versus No drain	↑	No drainage
						→	Timing of drain removal (removal on 2 POD versus when deceased to 50 ml/day)
Galatius, 2003 (63)	2	Prospective	59	MRM	Use of ultrasonic scalpel versus scissors and electrocautery in flap dissection	→	Surgical device (ultrasound scalpel versus scissors and electrocautery)
Gonzalez, 2003 (70)	3	Retrospective	359	MRM, BCS	–	↑	Type of mastectomy (MRM > BCS)
						→	Age, presence or number of positive LNs, no. of removed LNs, tumor size, body weight, neoadjuvant chemotherapy
Langer, 2003 (46)	2	RCT	55	MRM, Simple, ALND	Fibrin glue versus None 4 versus 3 (NS)	→	Use of fibrin glue versus not using
Puttawibul , 2003 (47)	2	RCT	60	MRM	Axillary drainage versus Axillary and pectoral drainage 20 versus 37 (NS)	→	Number of drain

Table 1. *Continued*

Author, year	Level of evidence	Study type	Sample size	Type of mastectomy	Intervention and seroma formation (%)		Factors and direction of association*
Ulsoy, 2003 (48)	2	RCT	54	MRM	Use of fibrin glue versus None 18 versus 11 (NS)	→	Use of fibrin glue
Dalberg, 2004 (49)	2	RCT	250	MRM	Removal versus Preservation of pectoral fascia 39.8 versus 31 (NS) Removal of drain at first POD versus When <40 ml/day 48.5 versus 22.2 ($P < 0.001$)	↑ →	Removal of drain at first POD Removal or preservation of pectoral fascia
Jain, 2004 (50)	2	RCT	116	MRM, BCS	Drainage versus No drainage (with or without fibrin glue) 38 versus 26 (NS) Use of fibrin glue versus without fibrin glue in no drainage group 35 versus 41 (NS)	→	No drainage (with or without fibrin glue) versus drainage, use of fibrin glue versus without fibrin glue in no drainage
Lumachi, 2004 (51)	2	RCT	92	MRM, BCS	Use of ultrasound scissors versus Scissors and ligation in axillary dissection 20 versus 40 (NS)†	↑ →	Tumor size (larger), total drainage volume (greater), no. of involved nodes (greater), type of mastectomy (MRM > BCS) Age, side, pT, grade, ER, nodal status, number of LNs, use of ultrasonic scissors
Mustonen, 2004 (52)	2	RCT	40	MRM	Use of fibrin glue and fibrinolysis inhibitor (aprotinin) versus None (NS)	→	Use of fibrin glue and fibrinolysis inhibitor versus without fibrin glue
Chintamani, 2005 (53)	2	RCT	85	MRM	Half (350 mg/m^2) versus full (700 mg/m^2) vacuum suction drainage 3 versus 4 (NS)	→	Extent of negative suction pressure
Johnson, 2005 (54)	2	RCT	82	MRM, BCS, Simple, ALND, SLNB	Use of fibrin glue without drainage versus Drainage 36.8 versus 45.5 (NS)	→	Use of fibrin glue without drainage or drainage
Purshotham, 2005 (55)	1	RCT	298	MRM, BCS	SLNB12288;versus ALND‡ 11 versus 24 ($P < 0.01$)	↓	–
Soon, 2005 (56)	2	RCT	87	ALND	Drain versus no drain 94 versus 96 (NS)	→	With or without drainage
Shamley, 2005 (5)	2	Meta-analysis of 5 RCTs (11,32,33,71,72)	444	MRM, BCS	Delay versus early shoulder exercise 27 versus 46 Odds ratio: 0.41 (95%CI 0.27–0.61)	↑	Early shoulder exercise

Abbreviations: ALND, axillary lymph node dissection; BCS, breast-conserving surgery; BMI, body mass index; ER, estrogen receptor; Ext, extended mastectomy; IR, immediate reconstruction; LN, lymph node; MRM, modified radical mastectomy; NOS, not otherwise specified; NS, not significant; PgR, Progesterone receptor; POD, postoperative day; Rad, radical mastectomy; RCT, randomized controlled trial; SLNB, sentinel lymph node biopsy.

*Direction of association: ↑ = increase, current evidence demonstrates an association with significant increase of seroma formation; → = no association, current evidence demonstrates no association with seroma formation; ↓ = decrease, current evidence demonstrates an association with significant decrease of seroma formation.

†According to author's reply, difference between MRM and BCS was significant ($P < 0.01$), type of mastectomy was strictly related to the tumor size, and also use of ultrasonic scissors was associated with seroma formation in logistic regression analysis.

‡ALND was performed in node-positive patients as a second procedure if the SLN was positive for metastasis.

RCT of Purushotham et al. (55) has demonstrated that sentinel LN biopsy (SLNB) is associated with significantly less seroma formation than conventional axillary dissection.

WOUND DRAINAGE

In this category, intensity of negative suction pressure, no drainage, number of drains, type of drainage (closed suction versus passive drainage), type of drainage unit (evacuated bottle type versus bellow type) and type of drainage tube (multiple hole type versus multiple channel type) were assessed.

As for no drainage, 5 out of 8 studies had reported that this policy increases seroma formation (10,17,37,50,54,56,61,62). However, seroma formation was not influenced by the intensity of negative suction pressure (8,25,30,53,57), by the number of drains (16,18,47), or by the choice of closed suction drainage or passive drainage (6,7,23). These findings were consistent among studies. Similarly, in an RCT by Britton et al. (8), choice of evacuated bottle type or bellow type did not affect the number of aspirations required. In contrast, in a study of Porter et al. (36), a flat-type drain with multiple channels running the length of the drain reduced seroma formation as compared with a flat-type drain with multiple holes. It was speculated that the holes might clog more easily than the channels, which could lead to premature removal of drains. However, this study was not primarily planned to assess drain type, and the drain was selected according to the attending surgeon's preference.

SURGICAL DEVICES

Several surgical devices including electrocautery, laser scalpel, argon diathermy, ultrasonic scalpel and ultrasonic scissors have been investigated in an effort to achieve better hemostasis and to reduce seroma formation. However, no individual study has shown a significant effect on seroma formation with or without use of argon diathermy (28), a laser scalpel (21) or an ultrasonic scalpel (63). In contrast, according to the author's reply, the use of ultrasonic scissors has reduced seroma formation in an RCT by Lumachi et al. (51). In this study, level I and II axillary dissection was performed using either ultrasound scissors, or scissors and ligation, and electrocautery was used only for skin flap dissection. On the other hand, the use of electrocautery was significantly associated with increased seroma formation in an RCT by Porter et al. (36). In this study, the flap and fascia were dissected either by electrocautery or by scalpel, while control of small bleeding vessels with electrocautery and sharp dissection of the axillary nodes was performed in both groups.

SUTURE FLAP FIXATION

Suture flap fixation is a surgical technique for securing flaps to underlying tissues to close the dead space with sutures. Although this technique is not commonly performed, it is interesting to note that an RCT by Coveney et al. (19) has demonstrated that this technique reduces seroma formation in patients undergoing mastectomy. In association with this, an RCT by Purushotham et al. (45) has demonstrated that mastectomy without drainage does not increase seroma formation when this technique is applied. Also, a prospective study by Schuijtvlot et al. (61) has demonstrated that seroma formation is reduced by the use of this technique in patients undergoing BCT without axillary drainage.

MISCELLANEOUS

Moreover, several factors such as previous biopsy, blood loss, blood transfusion, operation time, skin incision, skin graft, surgeon and type of anesthesia have been assessed, and individual study has demonstrated that a longer operation time and diagonal skin incision as compared to vertical skin incision increase seroma formation (32,64). On the other hand, no association was found for previous biopsy (16,57), type of anesthesia (local or general) (17) or blood transfusion (58). Available evidence was inconclusive for whether or not skill or experience of the surgeon influences seroma formation (20,32,57), for quantity of blood loss (17,36), and for use or non-use of a skin graft (64,68).

POSTOPERATIVE MANAGEMENT

This category was subdivided into drainage volume, timing of drain removal, timing of shoulder movement, immobilization of the shoulder and use of an external pressure garment or compression dressing.

DRAINAGE VOLUME

Drainage volume during the initial 3 or 5 postoperative days (POD), total drainage volume and drainage volume in the 24 h before drain removal were assessed in terms of seroma formation. Among them, a positive association between drainage volume during the initial 3 POD and seroma formation was consistent between two RCTs (31,39). In contrast, evidence was inconclusive for drainage volume in the 24 h before drain removal (29,57,60), total drainage volume during the initial 5 POD (29,30) or total drainage volume (14,29,51,57,58). In this respect, the effect of total drainage volume might be confounded by duration of drainage and vice versa. However, interaction between total drainage volume and duration of drainage was not always documented in detail.

TIMING OF DRAIN REMOVAL

Several RCTs comparing timing of drain removal have provided complicated results. For example, in an RCT comparing removal of the drain on the fifth POD with removal on the eighth POD, the incidence of seroma formation was significantly high in the former (43). However, in that study, the drain was also removed when drainage volume fell to 30 ml or less per day for 2 consecutive days, and the actual day of drain removal between two groups was not provided. In contrast, in two RCTs comparing drain removal on the fifth

POD and removal when daily drainage volume became minimal (14,29), the timing of drain removal did not affect seroma formation. In the study by Inwang et al. (14) drains were usually removed at the 10th to 14th POD in the latter group, while the actual day of drain removal was not reported in the study by Ackroyd et al. (29).

On the other hand, evidence was inconclusive when seroma formation was compared between drain removal on the first or third POD, and when drainage volume fell to a minimal level (39,49,60,62). In two RCTs (39,49), early removal of drains increased seroma formation, whereas two other prospective studies did not find a significant association (60,62). In addition, in a study by Parikh et al. (15), there was no significant difference in the incidence of seroma formation between removal on the third POD and on the sixth POD.

Timing of shoulder movement, immobilization, and use of an external pressure garment or compression dressing

With respect to the timing of shoulder movement, five RCTs have found no significant influence when compared between shoulder movement from the first or third POD and that from the 5th to 10th POD (9,12,13,27,38), whereas a meta-analysis of five RCTs (11,32,33,71,72) found that early shoulder movement increased seroma formation (5). In contrast, three RCTs investigating the effect of shoulder immobilization using a bandage, collar and cuff or a sling, on seroma formation (9,11,27), consistently found that temporary shoulder immobilization did not have a significant effect on seroma formation. Evidence was inconclusive for the use of an external pressure garment or compression dressing (34,40), although the concept is to obliterate dead space by applying external pressure to the flaps and to encourage adhesion of the flaps to the underlying muscles.

Nonsurgical Modalities

This category includes radiation, neoadjuvant chemotherapy, use of adhesive glue and antifibrinolytic agents. With respect to radiation, a retrospective study of Say et al. (64) has demonstrated that pre- or postoperative radiation therapy does not affect seroma formation in patients who have undergone radical mastectomy. Similarly, neoadjuvant chemotherapy did not influence seroma formation in an RCT comparing neoadjuvant chemotherapy with immediate surgery (24).

With regard to the use of adhesive glue such as fibrin glue or bovine thrombin, nine RCTs found no significant effect on seroma formation (20,31,35,42,44,46,48,52,59), and an RCT by Vaxman et al. (26) even revealed that the use of fibrin glue increased seroma formation. Similarly, in an RCT by Jain et al. (50), patients were randomized to receive suction drainage or to receive no drain, and those allocated to no drainage were further randomized for application of fibrin sealant to the dissected area or to no intervention. Overall, this RCT failed to show any significant effect of the use of fibrin sealant on

Table 2. Direction and strength of each risk factor for seroma formation

Grade	Direction of association		
	Increase	No association	Decrease
Grade A	None	None	None
Grade B	Body weight (heavier), Extended radical mastectomy (versus simple mastectomy), Total drainage volume during the initial 3 days (greater)	Duration of drainage, Hormone receptor status, Immobilization of shoulder, Intensity of negative suction pressure, LN status or positivity of LNs, Number of drains, Number of removed LNs, Previous biopsy, Removal of drain on the fifth POD versus when daily drainage volume fell to a minimal, Stage, Type of drainage (closed suction drain versus static drainage), Use of fibrinolysis inhibitor	Sentinel LN biopsy (versus Axillary LN dissection)
Grade C	Diagonal skin incision (versus vertical skin incision), Hypertension, Multiple holes type drains (versus multiple channel type drain), No drainage (versus drainage), Obesity, Operation time (longer), Removal of drain on the 5 POD (versus on the 8 POD), Use of electrocautery in flap and fascia dissection (versus cold scalpel)	Anemia, Blood transfusion, Breast size, Diabetes mellitus, Extent of LN dissection, Grade, Histological type, Neoadjuvant therapy, No drainage with suture flap fixation versus drainage without suture flap fixation, Pathological tumor size, Radiation, Removal of drain on the 3 POD versus on the 6 POD, Removal or preservation of pectoral fascia, Smoking, Side, Specimen size, Specimen weight, Timing of shoulder movement, Type of anesthesia, Type of drainage unit (evacuated versus bellow typr), Tumor location, Use of laser scalpel, argon diathermy, and ultrasound scalpel, Use of adhesive glue	MRM + immediate reconstruction (versus MRM), Suture flap fixation, Use of ultrasonic scissors
Grade D	Inconclusive	Age, Blood loss, Body mass index/obesity index, Drainage volume in the 24 h before removal, Number of positive LNs, Removal of drain within the third POD versus removal of drain when daily drainage volume fell to a minimal amount, Skin graft, Surgeon, Total drainage volume, Total drainage volume during the initial 5days, Tumor size, Type of mastectomy (Radical mastectomy versus MRM), Use of pressure garment or external compression dressing	
	Troublingly inconsistent	Type of mastectomy (MRM versus BCS)	

Abbreviations: BCS, breast-conserving surgery; LN, lymph node; MRM, modified radical mastectomy; POD, postoperative day.

seroma formation. In addition, an RCT by Johnson et al. (54) failed to show any advantage to using fibrin glue without drainage over a drain. Similarly, the use of fibrin glue and fibrinolysis inhibitor or perioperative and postoperative administration of fibrinolysis inhibitor did not reduce seroma formation (22,52). The concept of the use of fibrinolysis inhibitor was based on the hypothesis that fibrinolytic activity of the plasmin system in serum and lymph might contribute to fluid accumulation.

SUMMARY OF EVIDENCE ON RISK FACTORS FOR SEROMA FORMATION

Each factor for seroma formation is categorized as shown in Table 2 according to the direction of the association and strength of evidence. Although there was no risk factor supported by strong evidence, there was moderate evidence to support the risk of seroma formation in individuals with heavier body weight, extended radical mastectomy as compared with simple mastectomy, and a greater initial three-day drainage volume. On the other hand, the following factors did not have a significant influence on seroma formation: the duration of drainage; hormone receptor status; immobilization of the shoulder; intensity of negative suction pressure; LN status or positivity of LNs; number of drains; number of removed LNs; previous biopsy; removal of drains on the fifth POD versus when the daily drainage volume fell to a minimal; stage; type of drainage (closed suction versus static drainage); and use of fibrinolysis inhibitor. In contrast, as might have been expected, SLNB reduced seroma formation. For the other factors that were commonly sited in the literature, evidence was weak or unproven.

Thus, although a number of factors have been correlated with seroma formation, strong data on factors associated with seroma formation are still rare, and it seems to be difficult to identify patients who will ultimately suffer from seroma. However, this study has provided findings that are useful for identifying commonly cited risk factors that have no evidence to support them.

References

1. Aitken DR, Minton JP. Complications associated with mastectomy. *Surg Clin North Am* 1983;63:1331–52.
2. Pogson CJ, Adwani A, Ebbs SR. Seroma following breast cancer surgery. *Eur J Surg Oncol* 2003;29:711–17.
3. Tadych K, Donegan WL. Postmastectomy seromas and wound drainage. *Surg Gynecol Obstet* 1987;165:483–7.
4. http://www.cebm.net/.
5. Shamley DR, Barker K, Simonite V, Beardshaw A. Delayed versus immediate exercises following surgery for breast cancer: a systematic review. *Breast Cancer Res Treat* 2005;90:263–71.
6. Morris AM. A controlled trial of closed wound suction. *Br J Surg* 1973;60:357–9.
7. Bourke JB, Balfour TW, Hardcastle JD, Wilkins JL. A comparison between suction and corrugated drainage after simple mastectomy: a report of a controlled trial. *Br J Surg* 1976;63:67–9.
8. Britton BJ, Gilmore OJ, Lumley JS, Castleden WM. A comparison between disposable and non-disposable suction drainage units: a report of a controlled trial. *Br J Surg* 1979;66:279–80.
9. Flew TJ. Wound drainage following radical mastectomy: the effect of restriction of shoulder movement. *Br J Surg* 1979;66:302–5.
10. Cameron AE, Ebbs SR, Wylie F, Baum M. Suction drainage of the axilla: a prospective randomized trial. *Br J Surg* 1988;75:1211.
11. Dawson I, Stam L, Heslinga JM, Kalsbeek HL. Effect of shoulder immobilization on wound seroma and shoulder dysfunction following modified radical mastectomy: a randomized prospective clinical trial. *Br J Surg* 1989;76:311–12.
12. Jansen RF, van Geel AN, de Groot HG, Rottier AB, Olthuis GA, van Putten WL. Immediate versus delayed shoulder exercises after axillary lymph node dissection. *Am J Surg* 1990;160:481–4.
13. Petrek JA, Peters MM, Nori S, Knauer C, Kinne DW, Rogatko A. Axillary lymphadenectomy. A prospective, randomized trial of 13 factors influencing drainage, including early or delayed arm mobilization. *Arch Surg* 1990;125:378–82.
14. Inwang R, Hamed H, Chaudary MA, Fentiman IS. A controlled trial of short-term versus standard axillary drainage after axillary clearance and iridium implant treatment of early breast cancer. *Ann R Coll Surg Engl* 1991;73:326–8.
15. Parikh HK, Badwe RA, Ash CM, Hamed H, Freitas R Jr, Chaudary MA, et al. Early drain removal following modified radical mastectomy: a randomized trial. *J Surg Oncol* 1992;51:266–9.
16. Petrek JA, Peters MM, Cirrincione C, Thaler HT. A prospective randomized trial of single versus multiple drains in the axilla after lymphadenectomy. *Surg Gynecol Obstet* 1992;175:405–9.
17. Somers RG, Jablon LK, Kaplan MJ, Sandler GL, Rosenblatt NK. The use of closed suction drainage after lumpectomy and axillary node dissection for breast cancer. A prospective randomized trial. *Ann Surg* 1992;215: 146–9.
18. Terrell GS, Singer JA. Axillary versus combined axillary and pectoral drainage after modified radical mastectomy. *Surg Gynecol Obstet* 1992;175:437–40.
19. Coveney EC, O'Dwyer PJ, Geraghty JG, O'Higgins NJ. Effect of closing dead space on seroma formation after mastectomy–a prospective randomized clinical trial. *Eur J Surg Oncol* 1993;19:143–6.
20. Uden P, Aspegren K, Balldin G, Garne JP, Larsson SA. Fibrin adhesive in radical mastectomy. *Eur J Surg* 1993;159:263–5.
21. Wyman A, Rogers K. Randomized trial of laser scalpel for modified radical mastectomy. *Br J Surg* 1993;80:871–3.
22. Oertli D, Laffer U, Haberthuer F, Kreuter U, Harder F. Perioperative and postoperative tranexamic acid reduces the local wound complication rate after surgery for breast cancer. *Br J Surg* 1994;81:856–9.
23. Whitfield PC, Rainsbury RM. Suction versus siphon drainage after axillary surgery for breast cancer: a prospective randomized trial. *Br J Surg* 1994;81:547.
24. Forouhi P, Dixon JM, Leonard RC, Chetty U. Prospective randomized study of surgical morbidity following primary systemic therapy for breast cancer. *Br J Surg* 1995;82:79–82.
25. van Heurn LW, Brink PR. Prospective randomized trial of high versus low vacuum drainage after axillary lymphadenectomy. *Br J Surg* 1995;82:931–2.
26. Vaxman F, Kolbe A, Stricher F, Zund D, Volkmar P, Gros D, et al. Does fibrin glue improve drainage after axillary lymph node dissection? Prospective and randomized study in humans. *Eur Surg Res* 1995;27:346–52.
27. Browse DJ, Goble D, Jones PA. Axillary node clearance: who wants to immobilize the shoulder? *Eur J Surg Oncol* 1996;22:569–70.
28. Kerin MJ, O'Hanlon DM, Kenny P, Kent PJ, Given HF. Argon-enhanced cutting and coagulation confers advantages over conventional electrocautery for mastectomy. *Eur J Surg Oncol* 1996;22:571–3.
29. Ackroyd R, Reed MWR. A prospective randomized trial of the management of suction drains following breast cancer surgery with axillary clearance. *The Breast* 1997;6:271–4.
30. Bonnema J, van Geel AN, Ligtenstein DA, Schmitz PI, Wiggers T. A prospective randomized trial of high versus low vacuum drainage after axillary dissection for breast cancer. *Am J Surg* 1997;173:76–9.
31. Burak WE, Jr., Goodman PS, Young DC, Farrar WB. Seroma formation following axillary dissection for breast cancer: risk factors and lack of influence of bovine thrombin. *J Surg Oncol* 1997;64:27–31.
32. Schultz I, Barholm M, Grondal S. Delayed shoulder exercises in reducing seroma frequency after modified radical mastectomy: a prospective randomized study. *Ann Surg Oncol* 1997;4:293–7.
33. Abe M, Iwase T, Takeuchi T, Murai H, Miura S. A Randomized Controlled Trial on the Prevention of Seroma after Partial or Total Mastectomy and

Axillary Lymph Node Dissection. *Breast Cancer* 1998;5:67–9.
34. Chen CY, Hoe AL, Wong CY. The effect of a pressure garment on post-surgical drainage and seroma formation in breast cancer patients. *Singapore Med J* 1998;39:412–15.
35. Gilly FN, Francois Y, Sayag-Beaujard AC, Glehen O, Brachet A, Vignal J. Prevention of lymphorrhea by means of fibrin glue after axillary lymphadenectomy in breast cancer: prospective randomized trial. *Eur Surg Res* 1998;30:439–43.
36. Porter KA, O'Connor S, Rimm E, Lopez M. Electrocautery as a factor in seroma formation following mastectomy. *Am J Surg* 1998;176:8–11.
37. Zavotsky J, Jones RC, Brennan MB, Giuliano AE. Evaluation of axillary lymphadenectomy without axillary drainage for patients undergoing breast-conserving therapy. *Ann Surg Oncol* 1998;5:227–31.
38. Chen SC, Chen MF. Timing of shoulder exercise after modified radical mastectomy: a prospective study. *Changgeng Yi Xue Za Zhi* 1999;22:37–43.
39. Kopelman D, Klemm O, Bahous H, Klein R, Krausz M, Hashmonai M, et al. Postoperative suction drainage of the axilla: for how long? Prospective randomised trial. *Eur J Surg* 1999;165:117–20; discussion 121–2.
40. O'Hea BJ, Ho MN, Petrek JA. External compression dressing versus standard dressing after axillary lymphadenectomy. *Am J Surg* 1999;177:450–3.
41. Dinsmore RC, Harris JA, Gustafson RJ. Effect of fibrin glue on lymphatic drainage after modified radical mastectomy: a prospective randomized trial. *Am Surg* 2000;66:982–5.
42. Berger A, Tempfer C, Hartmann B, Kornprat P, Rossmann A, Neuwirth G, et al. Sealing of postoperative axillary leakage after axillary lymphadenectomy using a fibrin glue coated collagen patch: a prospective randomised study. *Breast Cancer Res Treat* 2001;67:9–14.
43. Gupta R, Pate K, Varshney S, Goddard J, Royle GT. A comparison of 5-day and 8-day drainage following mastectomy and axillary clearance. *Eur J Surg Oncol* 2001;27:26–30.
44. Moore M, Burak WE, Jr., Nelson E, Kearney T, Simmons R, Mayers L, et al. Fibrin sealant reduces the duration and amount of fluid drainage after axillary dissection: a randomized prospective clinical trial. *J Am Coll Surg* 2001;192:591–9.
45. Purushotham AD, McLatchie E, Young D, George WD, Stallard S, Doughty J, et al. Randomized clinical trial of no wound drains and early discharge in the treatment of women with breast cancer. *Br J Surg* 2002;89:286–92.
46. Langer S, Guenther JM, DiFronzo LA. Does fibrin sealant reduce drain output and allow earlier removal of drainage catheters in women undergoing operation for breast cancer? *Am Surg* 2003;69:77–81.
47. Puttawibul P, Sangthong B, Maipang T, Sampao S, Uttamakul P, Apakupakul N. Mastectomy without drain at pectoral area: a randomized controlled trial. *J Med Assoc Thai* 2003;86:325–31.
48. Ulusoy AN, Polat C, Alvur M, Kandemir B, Bulut F. Effect of fibrin glue on lymphatic drainage and on drain removal time after modified radical mastectomy: a prospective randomized study. *Breast J* 2003;9:393–6.
49. Dalberg K, Johansson H, Signomklao T, Rutqvist LE, Bergkvist L, Frisell J, et al. A randomised study of axillary drainage and pectoral fascia preservation after mastectomy for breast cancer. *Eur J Surg Oncol* 2004;30:602–9.
50. Jain PK, Sowdi R, Anderson AD, MacFie J. Randomized clinical trial investigating the use of drains and fibrin sealant following surgery for breast cancer. *Br J Surg* 2004;91:54–60.
51. Lumachi F, Brandes AA, Burelli P, Basso SM, Iacobone M, Ermani M. Seroma prevention following axillary dissection in patients with breast cancer by using ultrasound scissors: a prospective clinical study. *Eur J Surg Oncol* 2004;30:526–30.
52. Mustonen PK, Harma MA, Eskelinen MJ. The effect of fibrin sealant combined with fibrinolysis inhibitor on reducing the amount of lymphatic leakage after axillary evacuation in breast cancer. A prospective randomized clinical trial. *Scand J Surg* 2004;93:209–12.
53. Chintamani, Singhal V, Singh J, Bansal A, Saxena S. Half versus full vacuum suction drainage after modified radical mastectomy for breast cancer—a prospective randomized clinical trial[ISRCTN24484328]. *BMC Cancer* 2005;5:11.
54. Johnson L, Cusick TE, Helmer SD, Osland JS. Influence of fibrin glue on seroma formation after breast surgery. *Am J Surg* 2005;189:319–23.
55. Purushotham AD, Upponi S, Klevesath MB, Bobrow L, Millar K, Myles JP, et al. Morbidity after sentinel lymph node biopsy in primary breast cancer: results from a randomized controlled trial. *J Clin Oncol* 2005;23: 4312–21.
56. Soon PS, Clark J, Magarey CJ. Seroma formation after axillary lymphadenectomy with and without the use of drains. *Breast* 2005;14:103–7.
57. Tejler G, Aspegren K. Complications and hospital stay after surgery for breast cancer: a prospective study of 385 patients. *Br J Surg* 1985; 72:542–4.
58. Kumar S, Lal B, Misra MC. Post-mastectomy seroma: a new look into the aetiology of an old problem. *J R Coll Surg Edinb* 1995;40:292–4.
59. Medl M, Mayerhofer K, Peters-Engl C, Mahrhofer P, Huber S, Buxbaum P, et al. The application of fibrin glue after axillary lymphadenectomy in the surgical treatment of human breast cancer. *Anticancer Res* 1995; 15:2843–5.
60. Yii M, Murphy C, Orr N. Early removal of drains and discharge of breast cancer surgery patients: a controlled prospective clinical trial. *Ann R Coll Surg Engl* 1995;77:377–9.
61. Schuijtvlot M, Sahu AK, Cawthorn SJ. A prospective audit of the use of a buttress suture to reduce seroma formation following axillary node dissection without drains. *Breast* 2002;11:94–6.
62. Talbot ML, Magarey CJ. Reduced use of drains following axillary lymphadenectomy for breast cancer. *ANZ J Surg* 2002;72:488–90.
63. Galatius H, Okholm M, Hoffmann J. Mastectomy using ultrasonic dissection: effect on seroma formation. *Breast* 2003;12:338–41.
64. Say CC, Donegan W. A biostatistical evaluation of complications from mastectomy. *Surg Gynecol Obstet* 1974;138:370–6.
65. Aitken DR, Hunsaker R, James AG. Prevention of seromas following mastectomy and axillary dissection. *Surg Gynecol Obstet* 1984;158: 327–30.
66. Corneillie P, Gruwez JA, Lerut T, Van Elst F. Early and late postoperative sequelae after surgery for carcinoma of the breast. *Acta Chir Belg* 1984;84:227–31.
67. Vinton AL, Traverso LW, Jolly PC. Wound complications after modified radical mastectomy compared with tylectomy with axillary lymph node dissection. *Am J Surg* 1991;161:584–8.
68. Chilson TR, Chan FD, Lonser RR, Wu TM, Aitken DR. Seroma prevention after modified radical mastectomy. *Am Surg* 1992;58:750–4.
69. Woodworth PA, McBoyle MF, Helmer SD, Beamer RL. Seroma formation after breast cancer surgery: incidence and predicting factors. *Am Surg* 2000;66:444–50; discussion 450–1.
70. Gonzalez EA, Saltzstein EC, Riedner CS, Nelson BK. Seroma formation following breast cancer surgery. *Breast J* 2003;9:385–8.
71. Lotze MT, Duncan MA, Gerber LH, Woltering EA, Rosenberg SA. Early versus delayed shoulder motion following axillary dissection: a randomized prospective study. *Ann Surg* 1981;193:288–95.
72. Knight CD, Jr., Griffen FD, Knight CD, Sr. Prevention of seromas in mastectomy wounds. The effect of shoulder immobilization. *Arch Surg* 1995;130:99–101.

Breast Cancer
Vol. 14 No. 1 January 2007

Review Article

Recent Topics of Health Outcomes Research in Oncology

Kojiro Shimozuma[*1,2], Hirohisa Imai[*3], Katsumasa Kuroi[*4], Shozo Ohsumi[*5], and Michikazu Ono[*6]

[*1]Department of Healthcare and Social Services, University of Marketing and Distribution Sciences, [*2]Institute for Stress Science, Public Health Research Foundation, [*3]Department of Epidemiology, National Institute of Public Health, [*4]Division of Surgery/Breast Oncology, Nyuwakai Oikawa Hospital, [*5]Department of Breast Oncology, National Hospital Organization Shikoku Cancer Center, [*6]Department of Health Science and Social Welfare, School of Human Sciences, Waseda University, Japan.

This article reviews recent topics in health outcomes research. First, we discuss the concept and importance of 'subjective' assessment of quality of life (QOL), and introduce new guidance, by the respective medical product regulatory authorities in Europe and the United States, for labeling claims of medical products that are assessed for outcomes related to QOL. Second, we address the application of item response theory (IRT) in developing and assessing QOL measures to compensate for several drawbacks of the classical psychometric approach, which has been commonly used to verify the reliability and validity of QOL instruments. Third, the relevance and determination of the minimally clinically important difference (MID) of QOL scores is discussed. Finally, we address the so-called 'response shift' which may affect the reliability of analysis results of QOL scores in longitudinal studies such as randomized clinical trials.

Breast Cancer 14:60-65, 2007.

Key words: Health outcome, Quality-of-life, Patient-reported outcomes, Item response theory, Computer adaptive testing

Introduction

Breast cancer is a disease that most commonly affects women in their 40s - 60s, a time when they are at their peak and most active. This is not only a great loss to society, but also can have negative effects on the social life of the patient herself. Fortunately, breast cancers differ from other types of solid cancer in that there are a number of effective treatments. However, even with successful treatment there may be damage that the breast cancer patient must live with for a long time; not only to physical functions but also psychosocial damage from a perceived loss of femininity. In breast cancer treatment, therefore, assessment of health outcomes such as quality of life (QOL) or health-related quality of life (HRQOL; a concept that excludes domains of QOL such as social environment and spirituality that are difficult to change through medical treatment or care interventions) is an area that has come to be emphasized strongly worldwide.

Thus, there are an increasing number of cases in recent years in which HRQOL is included as a secondary endpoints both in phase Ⅲ studies of anti-cancer treatments and in some late phase Ⅱ studies. In clinical trials of supportive or palliative therapy as well, it is not unusual for health outcomes themselves to be the primary endpoints.

As described above, demand for health outcome assessments such as HRQOL is rising considerably in the field of medical technology assessment, but at the same time even the two pillars of the QOL concept on which consensus has been reached among health outcome researchers, namely, multi-dimensionality (multi-domain concepts) and subjectivity (information sources are patients' subjective feelings), are not necessarily well understood by general clinicians or specialists. Moreover, clinicians know very little about the theoretical background for ensuring the reliability and

Reprint requests to Kojiro Shimozuma, Professor, Department of Healthcare and Social Services, University of Marketing and Distribution Sciences, 3-1, Gakuennishi-machi, Nishi-ku, Kobe, Hyogo 651-2188, Japan.
E-mail: Kojiro_Shimozuma@red.umds.ac.jp

Abbreviations:
QOL, Quality of life; HRQOL, Health-related quality of life; PRO, Patient-reported outcomes; IRT, Item response theory; CAT, Computer adaptive testing; MID, Minimally important difference; EBM, Evidence-based medicine; ES, Effect size; SD, Standard deviation; SEM, Standard error of measurement; RS, Response shift

validity of quantitative assessments of health outcomes, or the issues that need to be resolved in order to establish that background more firmly.

This article reviews recent topics in health outcomes research to make them more understandable to clinicians and others.

Organization of Concepts from "Quality of Life" (QOL) to "Patient-Reported Outcomes" (PRO), and Guidance for Using Health Outcomes in Medical Product Labeling Claims

Discrepancies that cannot be Ignored between Clinician-Reported Outcomes and Patient-Reported Outcomes

As mentioned in the introduction, subjectivity is the basis for the concepts of QOL/HRQOL. While important, we shall refrain here from philosophical or phenomenological discussions of subjectivity, such as questions of the range indicated by human subjectivity or whether subjectivity actually exists. What is important in health outcomes, particularly QOL assessments, is not whether or not they are objective but how accurately and precisely patients' subjective feelings (if they exist) are or can be understood.

In the range of a patient subjective experience there are problems that have already risen into consciousness, but there are also those which remain latent in the subconscious. Normally diagnosis proceeds with a search for causes starting with the symptoms the patient complains of, and treatment and care based on those complaints is the basis of medical attention. In assessments of QOL, however, we would like to also understand latent problems in the subconscious. Essentially, one of the aims of psychological measures and other such instruments is to bring to light latent problems that even the patient herself has not noticed.

Most doctors want to believe that they can gain a fairly good understanding of patient complaints from their interviews and observations, but in fact it is known that doctors overestimate or underestimate patient complaints from the nature of their symptoms and problems. For example, in a large-scale study of prostate cancer patients[1] doctors reportedly underestimated bone pain and fatigue/energy, while they overestimated erectile dysfunction. In addition, in studies in which we examined the frequency and level of peripheral neuropathy from various aspects in phase Ⅲ studies of breast cancer using taxane chemotherapy, doctors' assessments clearly underestimated both sensory and motor disturbances[2,3].

As stated above, the aim of "wanting to assess the patient's subjectivity" is a major aim of QOL/HRQOL assessments, but that aim and the range of included concepts is difficult to clearly communicate with those words, so the term "patient-reported outcomes" (PRO) started being used about 10 years ago. However, until recently this term had not come to be used by many researchers and medical professionals.

A Sign that the Concept and Term PRO is Widely Recognized: Introduction of Guidance in Cases when Health Outcomes are Used in Medical Product Labeling Claims

From last year through this year, two events have occurred that triggered the more widespread use of the term (concept) PRO. They were the issuing of new guidance, by the respective medical product regulatory authorities in Europe and the United States, for labeling claims of medical products that are assessed for outcomes related to QOL. We shall introduce them here briefly while comparing the differences between the two.

1) "Reflection Paper on the Regulatory Guidance for the Use of Health-Related Quality of Life (HRQOL) Measures in the Evaluation of Medicinal Products"[4]

The European Medicines Agency (EMEA) published this guidance in 2005. It is a short document of about 5 pages, which consists of 4 main sections: Ⅰ. Introduction (background), Ⅱ. HRQOL in drug evaluation process, Ⅲ. Study design for HRQOL assessment, and Ⅳ. Statistical analysis. The document defines the concepts of HRQOL and PRO, and describes the points that should be kept in mind when using HRQOL, in particular, in evaluations of medical products.

It should be noted here that in the Introduction section PRO is defined clearly as "Any outcome evaluated directly by the patient himself and based on the patient's perception of a disease and its treatment(s)." Differences with the concept of HRQOL are described as follows: "The term PRO is proposed as an umbrella term to cover both single dimension and multi-dimension measures of symptoms, HRQOL, health status, adherence to

treatment, satisfaction with treatment, etc." Thus, PRO is positioned above HRQOL.

The document also contains a detailed definition of HRQOL, which is that "HRQOL is considered to represent a specific type/subset of PROs, distinguished by its multi-dimensionality. Indeed, HRQOL is a broad concept which can be defined as the patient's subjective perception of the impact of his disease and its treatment(s) on his daily life, physical, psychological and social functioning and well-being. The definition of HRQOL has as a common basis the definition of health given by the WHO in 1948: 'Health is a state of complete physical, mental, and social well-being and not merely the absence of disease' "

A description of Ⅱ-Ⅳ is omitted here.

2) "Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (Draft Guidance)" [5]

This draft guidance was published by the U.S. Department of Health and Human Services in the Food and Drug Administration (FDA) in February 2006. As of August 2006 it remains open to the public for the purpose of gathering public comment, so it may undergo some changes in the final version.

The purpose of this guidance is described as follows: "This guidance describes how the FDA evaluates patient-reported outcome (PRO) instruments used as effectiveness endpoints in clinical trials. It also describes our current thinking on how sponsors can develop and use study results measured by PRO instruments to support claims in approved product labeling. By explicitly addressing the review issues identified in this guidance, sponsors can increase the efficiency of their endpoint discussions with the FDA during the product development process, streamline the FDA's review of PRO endpoint adequacy, and provide optimal information about the patient's perspective of treatment benefit at the time of product approval."

This guidance consists of seven sections: Ⅰ. Introduction, Ⅱ. Background, Ⅲ. PRO - Regulatory perspective, Ⅳ. Evaluating PRO instruments, Ⅴ. Study design, Ⅵ. Data analysis, and Ⅶ. Glossary. It is a detailed work of 36 pages.

The content differs somewhat from the EMEA document in that the focus is on PRO evaluation guidance rather than HRQOL evaluation guidance.

As in the EMEA document, the Introduction gives a clear definition of PRO: "A PRO is a measurement of any aspect of a patient's health status that comes directly from the patient (i.e., without the interpretation of the patient's responses by a physician or anyone else)." Furthermore, "In clinical trials, a PRO instrument can be used to measure the impact of an intervention on one or more aspects of patients' health status, hereafter referred to as PRO concepts, ranging from the purely symptomatic (response of a headache) to more complex concepts (e.g., ability to carry out activities of daily living), to extremely complex concepts such as quality of life, which is widely understood to be a multidomain concept with physical, psychological, and social components. Data generated by a PRO instrument can provide evidence of a treatment benefit from the patient perspective." In this we see that while QOL is a concept included in PRO, a characteristic of this guidance is that it shows clearly that a broad concept such as QOL can also be used as an end point in clinical trials for medical product labeling claims (however, high levels of reliability and validity are required for the evaluation).

Moreover, in section Ⅲ. PRO Regulatory perspective, the document says that there are parts of the treatment effect that only the patient can know, indicating the importance of patients' subjective assessment.

As mentioned above, the importance of patients' subjective assessment is one of the most difficult things for clinicians to understand, and so it is worth devoting an entire section to this point.

After the guidance is amended and completed based on public comment, it will serve to bring more widespread awareness of the concept of PRO.

Application of Item Response Theory (IRT) in Developing and Assessing QOL/PRO Measures

The theoretical basis for the reliability and validity of data when developing measures to quantify QOL or PRO, or conducting health outcome evaluations using these measures, comes from psychometry, which has been developed as a field of psychology.

This theory is also called the classical test model, and has been accepted by many resear-

chers. However, it has several drawbacks.

For example, there are inefficiencies in measures developed using this theory and evaluations using these measures. When using these measures in actual clinical trials, it is rare that the study population has exactly the same attributes as the target group when the measure was developed. Therefore, if the same subjects do not have a previously evaluated history, it is necessary to verify whether or not the same psychometric characteristics as at the time of development were maintained in the study, before the results of the analysis of the obtained data can be interpreted.

This problem also occurs when a translated version of a measure is developed in a language other than the original language. Today, in the development of new medical products that are predicted to have a wide market, it is not unusual for multinational clinical trials to be conducted in Japan, the USA, and Europe. If the psychometric characteristics differ with each translation, the QOL/PRO data obtained cannot be used together in analysis.

Moreover, while it is often the case that several measures exist for the same purpose, it is not rare for there to be slight differences in the scale of the measure or the conceptual structure being measured and evaluated. In these cases, there is a problem in that it is difficult to compare data that have been evaluated with different measures.

One method that has started to be considered in order to overcome problems such as those above is to adopt item response theory (IRT; also called new test theory) to health outcome assessments. IRT has been used mainly in the field of education in preparing test problems for TOEFL or shared tests for entrance to overseas universities.

IRT is fundamentally a multiple logistic model. When only the single parameter of "difficulty" is used it is called the "Rasch model," and special software is commercially available in Japan. Additional methods are a two parameter model made by adding the parameter of "discrimination" to that of "difficulty," and a model with the addition of a third parameter of "guess." The most commonly used is a two parameter logistic model.

Although IRT has been applied in the development of short versions of QOL/PRO measures and in cross-cultural validation, it is still not used by many researchers. Therefore, with the wide application of IRT to item analysis and scale scoring of health outcome assessments, we can look forward to increasing its application in new evaluation methods called computer adaptive testing (CAT), similar to the field of education.

CAT is a method of administering QOL/PRO measures by computer using the psychometric framework of IRT. Items are automatically selected by computer from some item banks on the basis of the patient's responses to previously administered items. This process uses an algorithm to estimate a person's score and the score's reliability and then chooses the best next item, enabling scale administration based on specifications such as content coverage, test length, and standard error. The capacity to rank all patients on the same continuum, even if they have not been given any common items, allows for an assessment that is individually tailored to each person. With item banking, each patient need only answer a subset of items to obtain a measure that accurately estimates what would have been obtained by administering the entire set of items[5].

This promising method has the benefit for researchers and clinicians of reducing the labor and cost to develop several measures for each subject group and goal, while for patients who are evaluation subjects it has the benefit of letting them know where they rank in the entire subject group by answering a minimum number of short questions.

However, several premises and constraints can also be predicted when IRT is applied to CAT. For example, IRT relies on strong assumptions, that is, unidimensionality, local independence, and monotonicity. Things that do not fit these assumptions are also included in the data and concepts of QOL/PRO.

In the USA, a large project called Patient-Reported Outcomes Measurement System (PROMIS) was started in September 2004 with the support of the National Institutes of Health (NIH) and the simultaneous contributions of many researchers[6]. This project was begun to solve many of these problems at once, and there are high expectations for the results.

Minimally Importance Difference (MID)

In clinical studies of health outcome assessments using these measures, the many data that are obtained, whether in a cross-sectional study or a clinical trial, are normally analyzed statistically

and applied to the clinical setting based on the results of statistical significance tests. Of course, this is the fundamental method of evidence-based medicine (EBM), which is normally used in analysis of survival periods and other parameters, and it is scientifically correct. However, in regard to outcomes for qualities such as QOL, many clinicians question why feedback on results cannot be given quickly to the subjects assessed. Moreover, some also say that for QOL, perhaps because individual clinicians and patients can to a certain degree imagine the outcomes, findings of a significant difference do not necessarily match the image of clinicians and patients.

The cause of these doubts is that the actual minimum number of points difference needed to indicate a clinical meaning is unknown, as is whether certain differences can be said to be important. In other words, health outcomes researchers did not seem to think seriously about the minimally (clinically) important difference (MID) in the past.

Meanwhile, it has come to be emphasized that, in demonstrating the effectiveness of new medical products with clinical trials that have health outcome as the primary endpoint in the guidance outlined in section I., data on MID are needed in advance in the sample size estimation (in other words, the use of a measure for which MID is already known is strongly recommended) in the stage of developing the study protocol, and debate is increasing.

Various attempts have been made over the past several years to resolve these issues. The methods may be broadly divided into (1) distribution-based methods and (2) anchor-based methods.

In distribution-based methods, the effect size (ES) has been used often. However, the criticism has been made that although standard deviation (SD) is considered, it is sample size dependent. Therefore, indicators such as standard error of measurement (SEM) have come to be considered[7]. These distribution-based methods are simple, but in the end we cannot expect an answer to the fundamental problem of whether they match human perceptions.

Meanwhile, in anchor-based methods detailed investigations using real people have been accumulated for each measure[8, 9]. Problems have also been indicated with these methods, such as the possibility of differences occurring with each measure, or when measures have been improved or worsened.

Based on these various studies, Dr. David Osoba, a well-known Canadian health outcome researcher, proposed at an educational workshop of the International Society for Quality of Life Research (ISOQOL) Annual Meeting in 2005 several possible definitions of MID, including "7-8 or 10% change of QOL score', '1/3-1/2 of SD', '0.4 × ES,' and wondered if agreement could not be reached. It is desirable that a consensus among researchers be reached at an early date.

Response Shift (Adaptation to Changing Health)

It has long been known that people's value standards change with experience. This is called a response shift (RS). RS is basically divided into 3 categories[10, 11]: "change in internal standards," "change in values," and "reconceptualization."

"Change in internal standards" is the change that occurs in a person's value standards from knowing a higher QOL than before, or conversely a lower QOL, as a result of the occurrence of a major event to that person. "Change in values" is when, for example, a person gives priority to fulfillment in work over family relationships before becoming ill, but after becoming ill comes to place a lower priority on work. "Reconceptualization" is when there is a change in factor structure, such as the inclusion in question items of emotional well-being before an intervention, but inclusion of social well-being after the intervention.

RS is not really a problem in cross-sectional or short-term longitudinal studies, but in randomized clinical trials evaluating the effectiveness of medical products the reliability of statistical significance test results can be affected if a large RS occurs in only one treatment arm of 2 groups.

For all 3 types of RS above, it is important to steadily examine what kinds of problems occur in what situations, and how serious they are. For example, it is necessary to understand the situation from subject attributes, kinds of intervention, study period, baseline QOL, and improving or worsening direction. Clarification of the details of the properties that can occur from these biases may contribute to dramatically raising the reliability of analysis results for QOL/PRO.

Conclusions and Implications

This article outlined the following recent topics and issues in health outcome research, not just in the area of breast cancer, but in medicine in general: (1) introduction of guidance from regulatory authorities related to medical product approval, using the new term of PRO, (2) the possibilities of CAT using IRT, (3) renewed awareness of the importance of MID, and (4) the possibility of improving the reliability of QOL/PRO assessments by promoting RS research.

It is hoped that advances in these research areas will make health outcome research more approachable and beneficial for clinicians and patients alike.

Acknowledgment

This review was partly supported by the grant-in-aid to Kojiro Shimozuma M.D., Ph. D. from the Ministry of Education, Culture, Sports, Science and Technology of Japan (number 08590503).

References

1) Litwin MS, Lubeck DP, Henning JM, Carroll PR: Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. *J Urol* 59:1988-1992, 1998.
2) Shimozuma K, Ohashi Y, Takeuchi A, Morita S, Ohsumi S, Sunada Y, Kuroi K, Makino H, Watanabe T, Hausheer FH: Validation of the Patient Neurotoxicity Questionnaire (PNQ) during taxane chemotherapy in a phase Ⅲ randomized trial in patients with breast cancer: N-SAS BC 02. Proc. of 27th San Antonio Breast Cancer Symposium, 2004.
3) Shimozuma K, Ohashi Y, Takeuchi A, Aranishi T, Morita S, Kuroi K, Ohsumi S, Makino H, Watanabe T, Hausheer FH: Assessment and quantification of taxane-induced neurotoxicity in a phase Ⅲ randomized trial of patients with breast cancer (AC followed by PAC/DOC vs. PAC/DOC alone): N-SAS BC 02. *J Clin Oncol* 24:473s, 2006.
4) 'Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQOL) measures in the evaluation of medicinal products' European Medicines Agency Pre-authorisation Evaluation of Medicines for Human Use, Committee for Medical Products for Human Use (CHMP), London, 27 July 2005, http://www.emea.eu.int/pdfs/human/ewp/13939104en.pdf
5) 'Guidance for industry patient-reported outcome measures: Use in medical product development to support labeling claims (Draft guidance).' U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH) Feb. 2006 http://www.fda.gov/CDER/GUIDANCE/5460dft.pdf
6) Patient-reported Outcomes measurement Information System (PROMIS). http://www.nihpromis.org/, Aug, 2006.
7) Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD: Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 37:469-478, 1999.
8) Jaeschke J, Guyatt GH: Measurement of health status: Ascertaining the minimal clinically important difference. *Control Clin Trials* 10:407-415, 1989.
9) Doyle C, Crump M, Pintilie M, Oza AM: Does palliative chemotherapy palliate? Evaluation of expectations, outcome, and costs in women receiving chemotherapy for advanced ovarian cancer. *J Clin Oncol* 19:1266-1274, 2001.
10) Schwartz CE, Sprangers MAG: Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med* 48:1531-1548, 1999.
11) Sprangers MAG, Schwartz CE: Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med* 48:1507-1515, 1999.

Breast Cancer
Vol. 14 No. 1 January 2007

Review Article

Quality of Life of Breast Cancer Patients and Types of Surgery for Breast Cancer - Current Status and Unresolved Issues -

Shozo Ohsumi[*1], Kojiro Shimozumai[*2], Katsumasa Kuroi[*3], Michikazu Ono[*4], and Hirohisa Imai[*5]

*[1]Department of Breast Oncology, National Hospital Organization Shikoku Cancer Center, *[2]Department of Healthcare and Social Services, University of Marketing and Distribution Sciences, *[3]Division of Surgery and Breast Oncology, Nyuwakai Oikawa Hospital, *[4]Department of Health Science and Social Welfare, School of Human Sciences, Waseda University, *[5]Department of Epidemiology, National Institute of Public Health, Japan

Since standard radical mastectomy was established by Halsted for breast cancer, surgical procedures for breast cancer have been changed according to the results of randomized controlled trials. Breast-conserving treatment is now regarded as a standard local treatment for early breast cancer. More recently, sentinel node biopsy is becoming popular as an alternative procedure to axillary node dissection for nodal staging. These new procedures have been believed to be better in terms of patients' quality of life in comparison with previous surgical procedures without impairing prognosis.

Many studies regarding the quality of life (QOL) of patients after such procedures have been reported. Here we review those data, especially of studies comparing quality of life of patients after mastectomy and breast-conserving treatment, and of those after axillary node dissection and sentinel node biopsy. Viewpoints and issues on surgical treatment-related QOL are discussed.

Breast Cancer 14:66-73, 2007.

Key words: Quality of life (QOL), Breast cancer, Surgery

Introduction

Since standard radical mastectomy was established by Halsted for breast cancer surgery, numerous randomized controlled trials (RCTs) have been carried out. Although at one stage there was a temporary tendency toward extended radical mastectomy with parasternal node dissection in order to improve prognosis, the lack of improvement in outcome discouraged this trend[1,2]. Investigations have sought less extensive surgical approaches without negatively affecting surgical effects. If favorable surgical effects could be established with less extensive surgery (LES), cosmesis would be improved, and furthermore the quality of life (QOL) of patients would be maintained with minimal attenuation of functional aspects of the operated areas. Hitherto, no prognostic differences has been shown between modified radical mastectomy with preserved pectoral muscles and standard radical mastectomy with non-preserved pectoral muscles in RCTs[3,4].

Subsequently, only the area of the breast involved by tumor was resected, and followed by radiation therapy of the preserved breast; i.e. breast-conserving treatment (BCT). RCTs have shown no difference in prognosis between BCT and radical mastectomy[5,6].

With recent extensive applications of sentinel node biopsy (SNB), axillary lymph node dissection (ALND) has been able to be avoided in patients without lymph node metastasis. RCTs comparing the prognoses of operated patients with SNB alone and ALND are ongoing but definitive results have not yet been reported.

Whether LES would contribute to better postoperative QOL in post-surgical patients requires

Reprint requests to Shozo Ohsumi, Department of Breast Oncology, National Hospital Organization Shikoku Cancer Center, 160 Kou Minamiumemoto-machi, Matsuyama 791-0288, Japan
E-mail: sosumi@shikoku-cc.go.jp

Abbreviations:
QOL, quality of life; RCT, randomized controlled trial; LES, less extensive surgery; BCT, breast-conserving treatment; SNB, sentinel node biopsy; ALND, axillary lymph node dissection; ROM, Range of motion; BDI, Beck Depression Inventory; BSI, Breif Symptom Inventory; MAC, Mental Adjustment to Cancer Scale; SF-36, Short Form 36; STAI, State-Trait Anxiety Inventory; FACT, Functional Assessment of Cancer Therapy scale

Table 1. Methodologic framework of studies

study	stage	treatment	sample size	sampling procedure study design	% dropout rate
Fallowfield et al[9]	I+II	BCT + RT	n = 48	consecutive	20
		M	n = 53	retrospective	
De Haes et al[10]	I+II	BCT + RT	n = 17	consecutive	13
		RM	n = 17	longitudinal	
Lasry et al[11]	I+II	BCT	n = 44	consecutive	nr
		BCT + RT	n = 36	retrospective	
		M	n = 43		
Kemeny et al[12]	I+II	BCT	n = 14	consecutive	38
		BCT + RT	n = 11	longitudinal	
		M	n = 27		
Lee et al[13]	I+II	BCT + RT	n = 85	consecutive	12
		MRM	n = 88	longitudinal	
Schain et al[14]	I+II	BCT + RT	n = 76	consecutive	24
		MRM	n = 66	longitudinal	
Poulsen et al[15]		BCT + RT	n = 87	consecutive	13
		MRM ± RT	n = 97	retrospective	
Curran et al[16]	I+II	BCT + RT	n = 151	consecutive	69
		MRM	n = 127	longitudinal	

BCT: breast-conserving treatment, MRM: modified radical mastectomy, RM: radical mastectomy, RT: radiotherapy, M: mastectomy, nr: not reported
(modified from the table in the referred article 7)

an objective evaluation. Therefore, we reviewed studies comparing the post-operative QOL of BCT with that of mastectomy, and studies comparing those of SNB and ALND with reference to the literature.

Post-Operative QOL of BCT vs. Mastectomy in Breast Cancer Patients

Many studies on differences in the outcome between BCT and mastectomy have generated a series of review articles on the post-operative QOL of patients. Most of the studies were not RCTs in which the surgical therapy preferred by the patient had been performed with the findings confined to short-term postoperative outcomes. Results based on RCTs are critical and impartial, as they eliminate the selection bias of patients selecting a certain surgical approach. As such, the present study focused on RCTs comparing BCT and mastectomy first. Although we have reviewed the article by Kiebert *et al.*[7], we chose the RCTs that fulfilled the standard criteria of Irwig and Bennetts[8] from the studies quoted by Kiebert *et al.*, because the former included the results of non-randomized studies and RCTs of low quality as well. In this study, we further complemented the findings of new RCTs that were repoterd after the article of Kiebert *et al.* (Table 1-4)[9-16]. According to Table 1-4, most RCT-based studies involved small sample sizes with relatively low response compliance, yielding findings not of high quality. Based on the results, differences in the respective QOL domains were not established between BCT and mastectomy, except for body image. These findings coincide well with other non-RCT surveys as well[17]. On the other hand, in a prospective non-randomized comparative study by Engel *et al.* using EORTC QLQ-C30, young patients disliked mastectomy, as they were significantly plagued with emotional and social function issues as well as financial and future health concerns after the operation[18].

On the other hand, there are few surveys on long-term post-operative QOL, especially in RCTs. Dorval *et al.*[19] performed a cross-sectional study involving 124 breast cancer survivors with a mean post-operative period of 8.8 (7.8 - 9.3) years, 47 and 77 patients underwent partial and total mastectomies, respectively. The survey concentrated on psychiatric distress using the Psychiatric