INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

THE UNIVERSITY OF CALGARY

Cluster Analysis of Symptoms Reported by Silicone Breast Implant Recipients and Cosmetic Surgery Controls

by

Susan G. Barr

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMMUNITY HEALTH SCIENCES

CALGARY, ALBERTA

JULY, 1998

ABSTRACT

Controversy exists as to whether adverse health effects result from exposure to silicone breast implants (SBI). **Objectives:** 1) To compare incident symptoms and the health status of SBI-exposed women with cosmetic surgery controls. 2) To explore whether SBI recipients develop a unique cluster of multiple symptoms. **Design:** Secondary analysis of a population-based retrospective cohort study. **Patients:** Included 609 controls, 1016 silicone gel and 309 saline breast implant recipients without defined systemic diseases. **Findings:** Symptom frequency was greater in the silicone gel group than controls, and intermediate in the saline group. Cluster analysis identified a natural symptom grouping in the silicone gel group (poor memory, insomnia, numbness, myalgia, arthralgia, joint swelling, headache and dyspepsia). The same cluster was observed in the saline and control groups. The number of cluster symptoms per subject was moderately correlated with health status and the presence of fibromyalgia in all 3 exposure groups.

ACKNOWLEDGEMENTS

Special thanks to my supervisor, Dr. Steve Edworthy, for inviting me to participate in the original historical cohort study, and for his guidance with this thesis as well as during the development of my career goals. I would also like to thank the other members of my supervisory committee, Drs. Rollin Brant, Liam Martin and Marja Verhoef, for their insightful suggestions and for helping me to develop a feasible protocol for this work. Thanks also to Dr. Margaret Russell, who took time out of her busy schedule on short notice in order to act as an examiner during my oral defense.

I would also like to thank Dr. Rosario Talavera and Rafael Talavera for their assistance, particularly with the Medlog database. Many thanks to Elisia Teixeira for her patience and excellent secretarial support during the distribution of this thesis.

Finally, thanks to Dr. Derek Exner for his unfailing support, encouragement and editorial contributions.

To
Derek

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BMI	body mass index
CI	confidence interval
CTD	connective tissue disease
ICD	International Classification of Diseases
kg	kilograms
LQ	lower quartile
m^2	meters-squared
SBI	silicone breast implants
UQ	upper quartile

1.0 INTRODUCTION

1.1 Statement Of The Problem

The putative association between silicone breast implants (SBI) and connective tissue diseases (CTDs) is controversial. Examples of CTDs reported in SBI recipients include scleroderma, rheumatoid arthritis, systemic lupus erythematosus and Sjögren's syndrome. However, the majority of symptomatic women exposed to SBI develop nonspecific multisystem complaints that do *not* meet accepted case definitions for CTD (1). The most commonly reported symptoms in case series have included joint pain, muscle pain, fatigue, memory problems, dry eyes and mouth, hair loss, Raynaud's phenomenon and rash. It is not clear whether SBI-exposed women are at increased risk of developing a new constellation of symptoms, referred to by some as "atypical CTD" or "systemic silicone-related disease" (2). Alternatively, these individuals may fulfill existing case definitions of other conditions, such as chronic fatigue syndrome (3) or fibromyalgia (4). Adequate controlled studies of symptoms in women exposed to breast implants are lacking.

1.2 Significance Of The Problem

In 1992, silicone gel-filled implants were banned in North America pending further research because their manufacturers had failed to provide adequate post-marketing surveillance for adverse effects during 30 years of use (5). Saline-filled breast implants, which consist of a silicone elastomer shell filled with saline instead of silicone gel, continue to be widely used despite incomplete information on their safety. Thus, the

potential harm arising from SBI remains an important public health issue since approximately 1% of North American women currently have breast implants, and because saline implants continue to be used (6,7). Women with implants require valid safety information in order to put the potential risks into perspective and make intelligent decisions related to accessing specialist care and possible implant removal. Such information is also required by those considering augmentation with saline implants or replacement of their silicone gel prostheses with saline-filled devices. Many women with SBI perceive that their symptoms are not being taken seriously by health care professionals, but rather are being dismissed as trivial, imaginary or attributable to secondary gain (8).

A resounding criticism of the existing literature is that researchers have failed to evaluate whether a subset of SBI recipients may have an atypical CTD or other syndrome (9-14). This has been hampered, in part, by the lack of a clear case definition of such a disorder (15-17). Furthermore, there are no diagnostic laboratory tests or physical findings that are consistently abnormal in symptomatic women with breast implants. Nevertheless, there are examples in the literature of consensus case definitions for syndromes that consist of a minimum set of symptoms, including chronic fatigue syndrome (3). Although the study of symptom complexes can be problematic, this should be considered a credible area of research which may advance our understanding of common subjective complaints (18).

1.3 Objectives

A. *i)* To describe symptoms reported by women exposed to silicone gel breast implants, saline breast implants and other forms of non-silicone cosmetic surgery.

ii) To evaluate the agreement between symptoms self-reported on a mailed questionnaire and those recorded by a rheumatologist after clarification of symptom interpretation during a blinded face-to-face interview.

B. *i)* To explore the symptoms reported by women with silicone gel breast implants using cluster analysis in order to discover whether natural groupings of symptoms are present.

ii) To explore the symptoms reported by women with saline breast implants and cosmetic surgery controls using cluster analysis in order to allow qualitative comparison of symptom clusters with those of the silicone gel implant group.

iii) To externally validate the results of the cluster analyses utilizing variables that were not used to generate the cluster solution.

C. *i)* To describe the health status of women exposed to silicone gel breast implants, saline breast implants and other forms of non-silicone cosmetic surgery.

ii) To evaluate whether there is an association between the number of self-reported symptoms per subject and health status among women with silicone gel breast implants, saline breast implants and non-silicone cosmetic surgery controls.

1.4 Literature Review

1.4.1 Do breast implants cause connective tissue diseases? The question of whether silicone breast implants may play a causal role in the development of systemic rheumatic diseases and symptoms has been addressed by several epidemiologic studies. Although it may be theoretically impossible to "prove" the causal nature of an association, public health problems often demand action, and judgments regarding causal inference must be made despite imperfect knowledge (19). Hill's viewpoints on distinguishing causal from non-causal associations provide a useful framework for discussion (20). In particular, those aspects pertaining to analogy, plausibility, dose-response relationships, experimental evidence, temporality, consistency, and strength of association are considered below.

Early case reports and case series of patients with SBI and CTD played an important role in alerting the medical community to this potential association (21). A disproportionate number of scleroderma cases were reported in SBI-exposed women (22) which is interesting given that scleroderma is one of the rarest CTDs with an estimated prevalence of 1 in 10,000 women (23). A number of environmental exposures have been associated with disorders resembling scleroderma, including inhalation of silica in mining (24), vinyl chloride exposure (25) and ingestion of aniline-contaminated rapeseed oil (26). Therefore, by *analogy* it was reasonable to suspect that breast implants might be associated with an atypical CTD similar to scleroderma. Case series represent the weakest form of evidence, however these reports served to stimulate research into

potential biological mechanisms as well as the conduct of controlled epidemiologic studies.

The literature suggests that it is biologically *plausible* for SBI to cause systemic symptoms. A capsule of inflammatory cells and fibrosis forms around the silicone elastomer shell common to all types of breast implants and probably represents a foreign body reaction. Contractures and hardening of the capsule, which cause breast deformity, pain and tenderness, may result from various factors including bacterial colonization (27-29). Elevated levels of inflammatory mediators have been detected in the capsule (30) and could produce nonspecific flu-like symptoms if released systemically. However, there is no objective evidence of a systemic inflammatory response since studies to date have failed to demonstrate abnormal blood levels of inflammatory mediators in symptomatic SBI recipients (30,31). Alternatively, silicone might act as a stimulant of the immune system which could lead to autoimmune disease in genetically susceptible individuals (32,33). Such a response could be evoked by the silicone elastomer shell or silicone that has leaked from gel-filled implants. While silicone causes nonspecific immune stimulation in rats (34,35), a similar adjuvant effect has not been confirmed in humans. In addition, specific antigenic stimulation of the immune system by silicone resulting in arthritis has also been observed in rats (35). Whether silicone can induce a specific immune response in humans remains an area of controversy (36-46). Blinded (47) and unblinded (48) testing for autoantibodies that are markers of classical autoimmune diseases has failed to find a difference between SBI-exposed women and cosmetic surgery controls.

In general, a *dose-response relationship* between levels of exposure and the risk of disease supports the presence of a causal association. However, the association between dose and response does not prove causation since similar findings may result from confounding (19). Furthermore, the absence of a dose-response relationship does not disprove causation. For example, the health-related outcome of interest may occur only in genetically susceptible individuals (32,33). Alternatively, a threshold effect may occur in which the outcome is observed only at a particular level or duration of exposure.

The potential for a biological gradient in response to silicone has not been adequately examined to date. Moreover, it is not clear whether the risk of adverse effects varies with implant type, or the presence of local breast complications including rupture. Indeed, the majority of studies have failed to document implant type. While all SBI have a silicone elastomer shell, a higher level of exposure occurs with gel-filled implants than saline since silicone gel "bleeds" through the elastomer shell. Furthermore, local release of silicone gel can result from shell rupture which appears to be associated with implant aging. Rupture of saline-filled implants results in deflation of the implant and resorption of the salt and water solution by the body. In contrast, rupture of silicone gel-filled devices is diagnosed reliably only by direct visualization during surgery. Therefore, the rate of rupture is known only for selected women who elect to have their gel-filled implants removed due to health concerns. Nevertheless, case series of women undergoing explantation report rupture rates of approximately 50% at 8 years, and up to 95% at 12 years after insertion (27,49-51). This could be one explanation for the estimated latency period of 9 years before systemic symptoms arise (1).

While *experiments* provide the strongest form of evidence for causal inference, results from animal models may not reflect the human experience. Furthermore, it is not ethical to randomly assign women to receive either SBI or non-silicone breast augmentation. Research is therefore limited to observational study designs. Case-control studies and historical cohort studies represent efficient methods of collecting data in a relatively short period of time. An important weakness of these retrospective designs is the potential difficulty in establishing a temporal relationship between exposure and disease, particularly when this information is based on recall by study subjects. Women with implants may be more likely to recall adverse health effects and to report their onset as occurring after cosmetic surgery compared with controls (i.e. recall bias). *Temporality* requires that the cause precede the effect in time, and is perhaps the most important criterion for causal inference (19). Therefore, retrospective studies should validate exposure and disease status, including dates, using objective sources such as medical records whenever possible. Alternatively, prospective cohort studies can establish the temporal relationship between exposure and disease, however, they are not efficient for rare diseases or if a long latency period exists between exposure and disease onset.

The existing literature of controlled analytic studies of the association between SBI and CTDs includes 8 case-control (52-56)(57-59) and 10 retrospective cohort studies (60-69). These studies have ruled out the presence of a moderate to large increase in risk of the combined end point of any CTD. While each of these studies has its own strengths and weaknesses, the *consistency* of these results argues against the possibility of a causal association. Nevertheless, only three studies were able to exclude a doubling of risk of

CTD (55,65,66), and none had sufficient power to rule out a two-fold increase in the risk of scleroderma. The association between SBI and scleroderma has been extremely difficult to study because both the exposure and the outcome are uncommon. Meta-analysis of the three case-control studies of scleroderma (53,56,57) suggests that the risk is not increased in women exposed to SBI [odds ratio (OR) 1.02; 95% confidence interval (CI): 0.56, 1.84] (70).

In general, the *strength of an association* argues in favor of that association being causal since large risks are unlikely to be due to confounding alone. However, weak associations can be important and smaller risk estimates between 1.0 and 2.0 do not disprove an association. This is particularly true if the association occurs only in a fraction of the population, for example those who might be genetically predisposed to adverse effects from SBI (32,33).

Only one study by Hennekens and colleagues has documented a small increase in the risk of CTD (RR 1.24, 95% CI 1.08, 1.41). However, this study has important flaws which might have resulted in an overestimate of risk (65). In this cross-sectional survey of the Women's Health Study cohort, exposure and disease status and associated dates were ascertained by self-report without validation by chart review or physical examination. The authors argue that the diagnoses reported are likely to be accurate because the study population consisted entirely of registered nurses. However, rheumatic diseases are diagnostically complex and are therefore less likely to be reported accurately by patients (71-73), including those who are nurses (74). Others have found that 28% of women are able to identify the year of SBI surgery correctly, and only 9% report the

correct month and year of surgery (75). Thus, recall bias regarding the temporal association of disease onset with respect to exposure may also have affected the risk estimates. This study has been described as representing "a worst case scenario" which may give some indication of the upper limit of risk of CTD (76).

1.4.2 Do breast implants cause systemic symptoms? Hill's viewpoints on causality discussed above also apply when considering symptoms associated with SBI. However, the issue of temporality deserves additional comment in this setting. Because symptoms are subjective in nature, we must rely on self-report in order to document their presence. Symptoms may be particularly vulnerable to recall bias, as it is difficult to verify their occurrence and date of onset. Whereas the diagnosis of a disease may occur at a well-defined point in time, the onset date of a symptom is often more difficult to pinpoint, particularly if it was gradual in onset. Furthermore, the presence or absence of all possible symptoms is usually not documented in medical records. Thus, historical records generally cannot be relied upon due to missing information.

There are numerous case reports and cases series which describe the complaints of hundreds of symptomatic women with breast implants, many of whom were referred to the authors for medical assessments regarding breast implant litigation. The most common symptoms, as reviewed by Borenstein (77), are: fatigue (83% of cases); arthralgias (62%); myalgias (49%); cognitive dysfunction (46%); dry mouth or sore throat (44%); hair loss (38%); skin rash (36%); and lymphadenopathy (35%). In the United States, concerns have been expressed that women opting for settlement from large

compensation funds were sent lists of criteria for compensation to be awarded in various categories (78). An independent medical examination was necessary in order to obtain a monetary settlement, however, many of the criteria require only the demonstration of subjective symptoms without any objective evidence of disease. It is not clear how these circumstances may have influenced the reporting of various symptoms. Furthermore, there is little data on the actual prevalence of symptoms in an unselected population of symptomatic and asymptomatic SBI-exposed subjects and an appropriately selected control group. The inclusion of a control group is important because a substantial proportion of the general population may report symptoms in the absence of any clear diagnosis (79-81). To date, four controlled historical cohort studies have reported data on symptoms and are summarized below (7,62,63,69).

Gabriel and coworkers (7) performed a retrospective chart review of patients enrolled in the Rochester Epidemiology Project (Mayo Clinic) in order to assess the risk of CTD among women with SBI. Data was collected on a large number of variables, including six symptoms (oral ulcers, photosensitivity, morning stiffness, serositis, arthritis, dry eyes and dry mouth). Because medical records generally do not contain complete information on the presence or absence of all relevant symptoms (82), the incidence rates for these symptoms are likely underestimated. The only significant difference between the SBI group and controls was a greater incidence of morning stiffness and serositis, however, stratified analysis revealed that the risk was increased only in women with breast cancer. This serves to emphasize the importance of breast cancer as a potential confounder.

Giltay and colleagues (62) assembled a retrospective cohort of all women with silicone gel-filled devices (N=374) implanted at the Free University Hospital, Amsterdam between 1978 and 1990. A female control, matched on age and year of non-silicone cosmetic surgery, was selected for each case from the same department of plastic surgery. In June 1992, a questionnaire was mailed requesting information on 7 symptoms and their year of onset. Among those contacted, 235 SBI-exposed (63%) and 210 controls (56%) were eligible and returned the questionnaire. A greater proportion of subjects with SBI reported at least one symptom arising after surgery (37.5%) compared to controls (21%). Individual complaints reported more frequently by SBI-exposed subjects included arthralgia (20% vs 9%), burning eyes (16% vs 7%) and photosensitive rash (9% vs 2%). For the remaining 4 symptoms, the study lacked sufficient power to exclude a moderate to large increase in risk. Furthermore, the relatively short mean duration of 6.5 years since surgery may have resulted in an underestimate of risk. In addition, these results may be confounded since 59 (25%) of the SBI respondents had a history of breast cancer while none of the controls did. This difference between groups could explain the increased risk of symptoms observed in the SBI group.

In a similarly designed study, Wells and coworkers (63) sent questionnaires to 826 women aged 20 to 60 years from the practice of one Florida-based plastic surgeon. The questionnaire asked whether 18 symptoms had been experienced and their time of onset relative to cosmetic surgery. Of 826 subjects, 222 SBI-exposed (43%) and 80 controls (26%) completed the questionnaire. None of the subjects had a history of breast cancer. Although surgical records were used to ascertain exposure status, implant type

was not detailed. The SBI group was 9.5 years younger than the control group and had a longer median duration between the time of cosmetic surgery and the survey (5 years compared to 3 years for controls). The short duration of exposure relative to the possible latency period of 9 years may have lead to an underestimate of the risk of symptoms. This study lacked sufficient power to exclude a three-fold increase in risk of most of the symptoms. Only "swollen glands under the arms" and "tender glands under the arms" were significantly more frequent in the SBI group. The authors suspect that this adenopathy occurred in the immediate postoperative period as a normal response to surgical trauma. Conversely, change in skin color, breathing difficulty and skin thickening were more common in the control group. The authors speculate that the breathing symptoms were related to the fact that many of the controls had undergone rhinoplasty, and that skin color changes were related to postoperative bruising. Therefore, self-reported symptoms in this study may not have measured the intended constructs. This misinterpretation of symptoms may have been avoided by including a blinded interview and physical examination.

Recently, Park et al (69) published a historical cohort study in which all women in southeast Scotland who had silicone gel breast implants between 1982 and 1991 (N=492) were invited to participate. All procedures were performed at a single hospital, and operating room records were used to obtain patient names and surgical history. Consenting participants attended the Plastic Surgery clinic for a history and physical assessment. This SBI cohort was divided two groups according to reason for mammoplasty, and were analyzed separately using different control groups. The first

group consisted of 203 women that had SBI for cosmetic reasons; 110 (54%) agreed to participate. Female plastic surgery outpatients (N=126) of similar age formed the comparison group. The second group consisted of 289 patients that had reconstructive surgery post-mastectomy for breast cancer; 207 (72%) consented to the clinical assessment. These women were compared to 88 consenting breast cancer patients who underwent surgery without reconstruction (88 refused to participate). No differences between the two groups of study patients and their respective controls were found among the 16 symptoms studied. However, this study lacked sufficient power to exclude a three-fold increase in risk of any symptom among the augmentation patients. A strength of this study was the inclusion of a face-to-face interview and physical examination in the study design, however, the surgeons were not blinded to exposure status such that the results may have been biased. Furthermore, no attempt was made to ascertain whether the symptoms started before or after the date of cosmetic surgery.

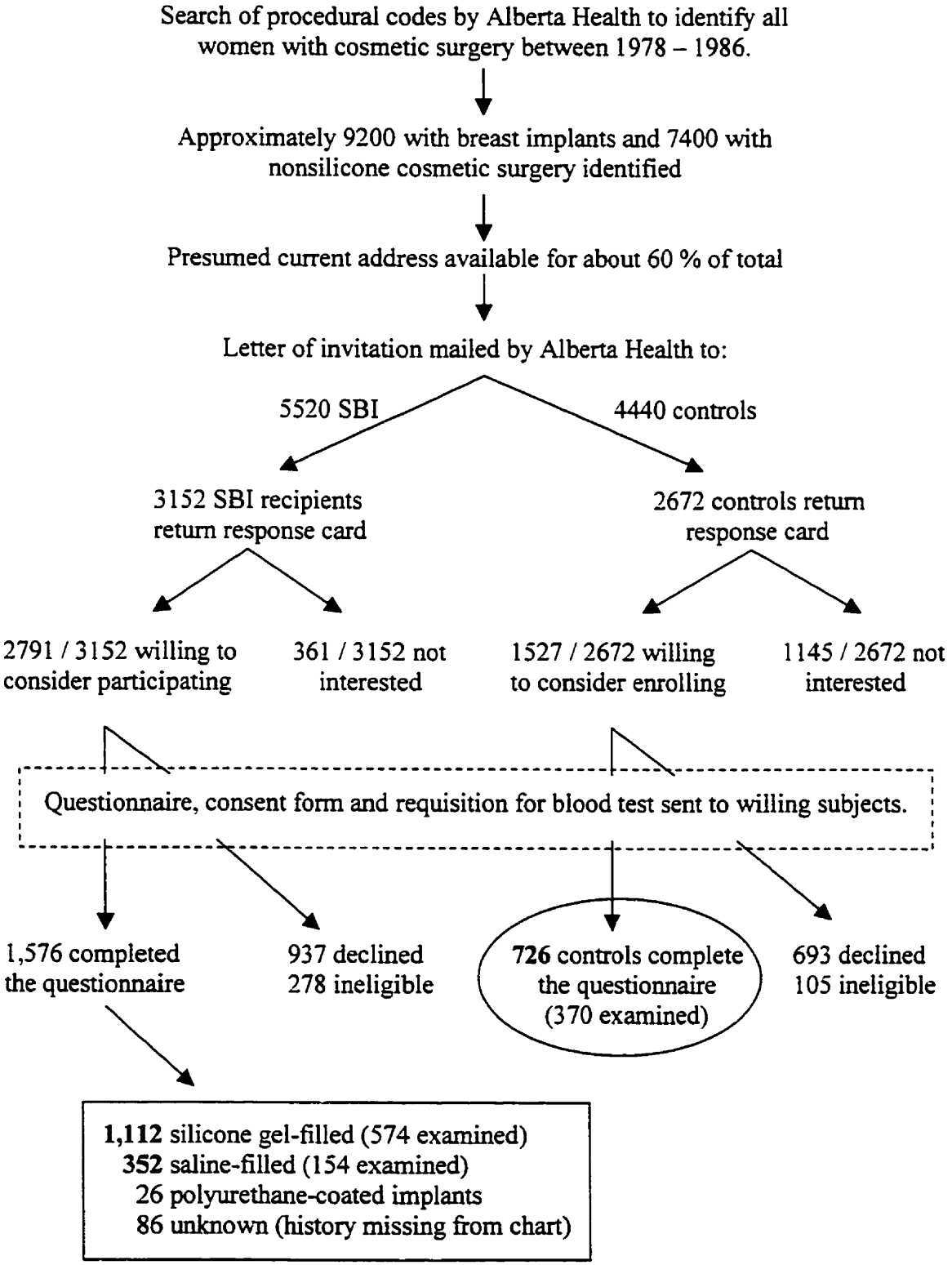
In summary, none of the above investigations included a complete list of the symptoms commonly linked with breast implants. Furthermore, these studies had insufficient power to exclude a moderate increase in the risk of most symptoms. While subjects were of comparable age to those reported in other epidemiological studies of SBI, additional potential confounders were generally not considered. In particular, the presence of comorbid conditions that might cause symptoms was not documented. Finally, these research efforts did not evaluate whether a subset of women with SBI develop a unique constellation of multiple symptoms, nor did they attempt to measure the possible impact of symptoms on health status or quality of life.

Our research group has recently completed a historical cohort study (68) in order to determine whether women with SBI have a higher incidence of CTD than controls. This study was designed to overcome some of the weaknesses present in previously reported investigations. In particular, it included a blinded face-to-face interview and physical examination of subjects by one of three rheumatologists in order to verify the presence or absence of disease, including symptoms. This thesis is based on a secondary analysis of the symptom data collected during that study. The primary cohort study will be summarized in the next section, and the variables and procedures used in this analysis will be presented in the Methods section.

1.5 Description Of Primary Cohort Study

1.5.1 Sampling procedure. The sampling frame for the original historical cohort study included all women in the province of Alberta who received breast implants for cosmetic reasons or who had other non-silicone cosmetic surgery between 1978 and 1986 (68). Patients with reconstructive surgery for breast cancer were excluded. Assembly of this population-based cohort was made possible through cooperation with the Alberta Department of Health who searched procedural codes of the Alberta Health Registry and mailed invitations to participate in the study. A detailed health questionnaire, consent form, and requisition for a blood test were mailed to respondents who expressed interest in participating. The recruitment process is summarized in flow chart format (Figure 1). Of the 16,600 patients that underwent cosmetic procedures in this time frame, 17% percent of the SBI (N=1,576) and 10% (N=726) of the control population participated.

Figure 1. Flow chart of enrollment of subjects into the silicone breast implant study.



The remaining subjects could not be contacted (40% of SBI, 40% of controls), did not respond to the invitation (26% of SBI, 24% of controls), refused to participate (14% of SBI, 25% of controls), or were ineligible (3% of SBI, 1% of controls). The mean duration since cosmetic surgery was 12 years. Cosmetic procedures among controls included liposuction (5%), lipectomy (15%), breast reduction (32%), rhinoplasty (30%), and facial surgery (18%).

1.5.2 Data collection. The self-administered questionnaire contained questions pertaining to demographics, symptoms, medical and surgical history and use of anti-rheumatic drugs. Subjects were asked to indicate whether each symptom had been present "before" the cosmetic surgery, "after" the cosmetic surgery or both "before and after" (Appendix A). An experienced medical records analyst traveled throughout Alberta to review surgical records in order to validate exposure status and the type of implant received. Blood samples were provided by 90% of the subjects for blinded testing of rheumatoid factor and antinuclear antibodies. Blood results were used to assist in the diagnosis of CTDs. Subjects were asked to attend an interview and physical examination if they met any of the following criteria: 1) history of rheumatic, neurologic, renal or thyroid disease at any time 2) two or more symptoms suggestive of rheumatic disease with onset "after" cosmetic surgery; 3) use of antirheumatic drugs; 4) a mean disability score $\geq$ 1.5 on the modified Stanford Health Assessment Questionnaire (83); or 5) a positive blood test result.

In total, 1,345 subjects were examined 6 to 12 months after the mailed questionnaire in order to validate diagnoses, thereby minimizing the potential for misclassification of disease status. In order to maintain physician blinding, all participants were instructed to wear their brassiere, a hospital gown and a canvas bib, and were also reminded not to divulge their exposure status. The examining rheumatologist was provided with a summary of positive symptoms, diagnoses and family history from the self-administered questionnaire to allow clarification of this information during the interview. The blinded rheumatologists then verified diagnoses during a standardized rheumatic disease history and physical examination. In order to minimize the chance of becoming unblinded, the rheumatologist recorded symptoms as "present" or "absent" but did not record whether they had onset before or after cosmetic surgery (Appendix B). A symptom was recorded as "absent" if, in the physician's judgment, the complaint represented a normal bodily sensation or if the subject had misinterpreted the intended meaning of the symptom. For example, "dry eyes" was recorded as absent if the patient indicated a history of eye strain secondary to overuse rather than a typical description of sicca symptoms.

1.5.3 Data analysis. All data collected in the historical cohort study were abstracted into Medlog (Medlog Clinical Data Management System, Information Analysis Corporation, Incline Village, NV). The database was examined for unusual values due to coding errors, and a sample of the data was reviewed along with the raw data in order to detect data entry problems. The primary outcome of the original cohort

study was a diagnosis of any of the following connective tissue diseases, using accepted criteria: rheumatoid arthritis (84), systemic lupus erythematosus (85), scleroderma (86,86) or Sjögren's syndrome. The diagnosis was made by the rheumatologist immediately after examining the subject based on clinical assessment and blood test results. Only outcomes in which the disease onset date occurred *after* the date of index surgery were considered as incident cases.

1.5.4 Findings. Based on aggregate data provided by Alberta Health, health care utilization was greater among study participants than nonparticipants irrespective of exposure group. This difference may have occurred because people who are familiar with the health care system may be more likely to volunteer for research studies. Furthermore, age-adjusted utilization rates suggested a possible bias in favor of detecting higher rates of rheumatic disease in the SBI group, which would not be expected to alter the main conclusions of the study. The age-adjusted relative risk of developing any one of the four CTD of interest was 1.00 (95% CI 0.45, 2.22). Although this result suggests that the risk of CTD is not increased in women with silicone gel breast implants, this study could not exclude a two-fold increase in risk.

2.0 METHODS

A secondary analysis of the historical cohort study conducted by our research group (Section 1.4) was undertaken in order to evaluate the symptoms and health status of women with silicone gel breast implants, saline breast implants and non-silicone cosmetic surgery controls.

2.1 Ethics

This secondary analysis did not involve contact with the study subjects, such that no further consent was required. While ethical approval was obtained for the original study, the question posed by this secondary analysis was not stated in the original protocol. Therefore, approval of this project was obtained from the Conjoint Medical Ethics Committee of the University of Calgary.

Nonmaleficence and beneficence refer to the ethical obligations of the investigator to minimize harm, and maximize the potential benefits to persons involved in the research project (87). In order for the present study to provide valid information that may benefit the subjects, as well as society, it must be conducted in a scientifically rigorous manner. Potential harm could arise if invalid results were published leading to inappropriate advice and management directed toward women with SBI. Respect for persons implies that the researcher will protect the subject's autonomy and maintain confidentiality (87). The author of this thesis is one of the rheumatologists that conducted interviews and physical examinations for the primary cohort study, such that access to the data did not present problems with confidentiality. Furthermore, hard

copies of the data are stored in a locked facility, and the computerized database is protected by a password.

2.2 Database and Statistical Software

The data from the primary cohort study was entered into Medlog (Medlog Clinical Data Management System, Version 95.6b, Information Analysis Corporation, Incline Village, NV). While this is a powerful data management program, the menu of statistical analyses available in Medlog is limited. The variables of interest for this study were therefore exported as a comma delimited file using the Communicate program in Medlog, and imported into Excel (Microsoft ® Excel 97, Microsoft Corporation, Redmond, WA). The Excel file was then imported into SYSTAT (SYSTAT ® 7.0 for Windows ®, SPSS Inc., Chicago, IL), which was used for all subsequent data manipulation and statistical analyses.

Quality control and data cleaning was performed during the original cohort study. This was accomplished by comparing a sample of the entered data with the raw data in order to find coding errors or data entry problems. The variables used in the present analysis were examined further for unusual values, and discrepancies corrected after reviewing the paper charts. The variables of interest included sociodemographic information, twenty-five self-reported symptoms, eighteen symptoms recorded by a physician during the face-to-face interview, and several scales measuring disability, quality of life, global health, and the intrusiveness of cosmetic surgery complications. These variables are described further in Section 2.4.

2.3 Subjects

This secondary analysis used data collected from three distinct exposure groups, including the 1,112 women with silicone gel-filled implants, the 352 women with saline implants, and the 726 controls who completed the study. In order to study unexplained symptoms in otherwise healthy women who have undergone cosmetic surgery, it was necessary to exclude subjects diagnosed with diseases that are known to cause systemic symptoms. All International Classification of Diseases (ICD-9) codes entered into the database were tabulated with the corresponding disease name. A selection of conditions that can cause multiple systemic symptoms was generated from this list, including malignancies, endocrinopathies, neuropsychiatric disorders, plus systemic rheumatic, cardiopulmonary, renal and gastrointestinal diseases. Some ICD-9 codes are relatively nonspecific and may not reflect the nature or severity of the disorder. Therefore, the paper chart of each subject with one or more of the selected conditions was reviewed in order to confirm that the correct ICD-9 code had been assigned and to determine whether the exclusion criteria described below were present.

Patients with the following conditions were excluded: 1) any cancer except non-melanoma skin cancer or cervical dysplasia; 2) well documented systemic rheumatic diseases, including rheumatoid arthritis, systemic or discoid lupus erythematosus, Sjogren's syndrome, scleroderma spectrum disorders, dermatomyositis, vasculitis, polymyalgia rheumatica, sarcoidosis, and seronegative spondyloarthropathies; 3) idiopathic thrombocytopenic purpura requiring prednisone or splenectomy; 4)

gastrointestinal disease, including Crohn's disease, ulcerative colitis, Celiac disease, chronic autoimmune hepatitis, past hepatitis B infection, and chronic pancreatitis; 5) endocrinopathies, including diabetes mellitus requiring oral hypoglycemic agents or insulin, hyperparathyroidism, hyperprolactinemia, and hypothyroidism with arthropathy (subjects on thyroid hormone replacement therapy were not excluded); 6) severe obesity, based on the exclusion criterion used for chronic fatigue syndrome (3), was defined as a body mass index $\geq$ 45 kg/m^2; 7) neuropsychiatric disorders, including multiple sclerosis, multifocal peripheral neuropathy, stroke, narcolepsy, reflex sympathetic dystrophy, encephalitis with permanent deficits, post-polio syndrome, traumatic or alcohol-induced neuropathies, schizophrenia; 8) significant cardiac disease, including previous myocardial infarction, congestive heart failure, and angina requiring anti-anginal medications or revascularization (angioplasty or coronary artery bypass grafting); 9) obstructive pulmonary disease requiring prednisone therapy or admission to hospital; 10) severe trauma resulting in chronic pain syndrome or major disability.

2.4 Description of Variables

2.4.1 Sociodemographic variables. A number of continuous and categorical baseline characteristics were analyzed in order to ensure that the silicone gel, saline and control groups were comparable to one another and to subjects described in the literature. Age in years was calculated based on the date that the self-reported questionnaire was completed. The time in years since surgery was calculated by subtracting the date of the cosmetic procedure from the date of the questionnaire. Self-reported height and weight

were used to calculate the body mass index (BMI), which is equal to the weight in kilograms divided by the height in meters squared. Use of the BMI allowed estimation of the proportion of women that met sex-specific definitions for being underweight (BMI < 19.0), overweight (BMI $\geq$ 27.3) and severely overweight (BMI $\geq$ 32.3)(88). The total number of years of education completed was also summarized.

Ethnic origin was recorded for each subject according to the following categories: white, black, Asian or Pacific Islander, First Nations, Hispanic, East Indian, or other. Other categorical variables included current marital status (single, married, separated, divorced, widowed), occupation and health behaviors. Occupations were coded according to the Canadian National Occupational Classification (89) and collapsed into broad categories for current employment only. Subjects with "no occupation" were not further categorized in the database as being homemakers, unemployed, retired, or disabled. As part of the health behavior evaluation, subjects were asked "How often do you exercise?". The 4 possible responses (never, 1-2 times per month, 1-3 times per week, > 3 times per week) were collapsed in order to categorize subjects according to whether they exercised regularly ($\geq$ 1-3 times per week) or not. No additional information regarding the type of exercise was available. Based on questions regarding current and past smoking history, a dichotomous variable was created in order to classify subjects as having "ever" smoked or "never" smoked. A similar dichotomous variable was created for alcohol use "ever". Alcohol consumption was further classified according to whether the subject had ever consumed 7 or more drinks per week. This cutoff was chosen to allow comparison with published data (90).

2.4.2 Self-reported symptoms. The self-administered questionnaire asked subjects to review a list of 30 symptoms and, if they had experienced any of the symptoms, to place a check mark indicating whether each symptom had occurred "before" cosmetic surgery, "after" cosmetic surgery or both "before and after" cosmetic surgery (Appendix A). If a symptom was not experienced, respondents were asked to leave the question blank. While this format does not allow the detection of missing data, the same design has been used in other self-administered questionnaires, including the Sickness Impact Profile (91).

Eleven of the symptoms were considered to be typical complaints associated with CTD, although most can occur alone or with other conditions. The symptoms "rash over the cheeks (butterfly rash)" and "sun sensitivity" refer to inflammatory rashes that are part of the diagnostic criteria for systemic lupus erythematosus (i.e. malar rash and photosensitivity, respectively). The symptom "skin tightening" was included as a screen for the highly characteristic skin changes of scleroderma, in which the skin becomes thick, tight and bound down to underlying tissues. "Skin ulcers" can occur with any of the CTD, but are particularly common in patients with scleroderma. "Red, white and blue skin color change in the fingers on exposure to cold or with emotional upset" is known as Raynaud's phenomenon. This spasm of blood vessels can occur alone or in association with any of the CTD, particularly scleroderma. "Difficulty swallowing or the feeling of food getting stuck", referred to as dysphagia, is also very common in

scleroderma. "Seizures" and "miscarriages" are features of systemic lupus erythematosus and lupus-like disorders.

Additional symptoms of CTD relate to joint involvement. "Joint swelling" may result from inflammatory arthritis due to soft tissue swelling and joint fluid accumulation. Bony enlargement due to noninflammatory osteoarthritis can also produce joint swelling. Arthritis may damage joints and surrounding tissues leading to "hand deformities" which tend to be most severe in association with rheumatoid arthritis. Prolonged and severe "morning stiffness" that improves with activity is a classical feature of inflammatory arthritis. For this study, morning stiffness was coded as present if it lasted at least 45 minutes.

The remaining 19 symptoms may also occur with increased frequency in women with CTD, but are much less specific. These nonspecific symptoms are self-explanatory and include: enlargement of lymph nodes; dry eyes; dry mouth; mouth sores (ulcers); nose sores (ulcers); hair loss; body rash; facial rash; heartburn, indigestion or belching (i.e. dyspepsia); chest pain on taking a deep breath; muscle pain; hand joint pain; headaches; numbness; waking at night (insomnia); trouble thinking and remembering; abnormal vaginal bleeding; menstrual problems; and cessation of menstrual periods.

Self-reported symptoms were among the criteria used in the primary cohort study to determine which patients should undergo a clinical evaluation for classical or atypical CTD. Women with two or more of the eleven CTD-associated symptoms were called for an assessment. Subjects that had 4 or more of 15 selected nonspecific symptoms were also evaluated. Heartburn, insomnia, and menstrual symptoms were excluded from the

latter criterion as they are very common in the general population and have not been associated with SBI in the past. Individuals with one or more CTD-associated symptoms plus 2 or more nonspecific symptoms were also invited for an interview and examination.

2.4.3 Physician-recorded symptoms. The blinded rheumatologist reviewed 35 symptoms with each subject that attended the clinical examination, 18 of which had also been included in the self-administered questionnaire (Appendix B). These 18 symptoms were used to assess agreement between self-report and interview data (see Section 2.5.2). The self-reported symptoms "rash on face" and "rash over cheeks" were combined for comparison with the symptom "malar rash" recorded by the physician. In addition, "mouth ulcers" and "nose ulcers" were combined for comparison with the symptom "mouth/nose ulcers" recorded during the clinical examination. As well, responses to the self-reported symptom "hand joint pain" were compared with those recorded by the physician as joint pain at any location (i.e. "arthralgia"). The four additional symptom variables appearing in Appendix B (fatigue, diarrhea, constipation, and abdominal pain) were used in the cluster analysis (Section 2.5.3).

2.4.4 Health Status Measures. Health status was evaluated using five validated instruments. Self-reported quality of life was measured using a 10 cm visual analogue scale (VAS) with the terminal descriptors "lowest quality" (0 cm) and "highest quality" (10 cm). Definitions of the terminal descriptors are also included as a guide for individuals completing the VAS (Appendix A). This scale was originally developed by

Spitzer and colleagues for use by health care professionals in the evaluation of cancer patients (92). While this scale can discriminate between healthy subjects and diseased patients, it was not designed to measure differences in quality of life among healthy individuals (92).

Subjects were asked to consider all aspects of their health and to rate their health on a 10 cm VAS with five verbal descriptors (Appendix A). A score of 10 cm corresponds to very poor health. Global health status was also rated by the examining physician using a double-anchored VAS scored from 0 cm ("very poor") to 10 cm ("best"). The Stanford Health Assessment Questionnaire (HAQ) was originally used for evaluating patients with rheumatoid arthritis (93), but has since been applied to a number of other rheumatic diseases (94). The disability index of the HAQ has been modified to include one item, instead of three, from each of eight domains pertaining to the level of difficulty performing activities of daily living (83). Responses are graded as "0" (without any difficulty), "1" (with some difficulty), "2" (with much difficulty) and "3" (unable to do) such that total scores range between 0 and 24 (Appendix A). This instrument is scored by taking the average of the eight responses to obtain a mean HAQ score.

The above generic scales measure complex constructs related to perceived "health" and "quality of life", and are therefore likely to incorporate a wide variety of factors into the final score. Since we are specifically interested in possible adverse effects of cosmetic surgery, the Illness Intrusiveness Rating Scale was also included (95). Subjects were asked whether they had experienced complications from their cosmetic surgery, and those with complications rated the degree to which the adverse effects had

impacted on thirteen different aspects of life (Appendix A). Seven-point Likert scales with two terminal descriptors, "not very much" ("1") and "very much" ("7") were scored for each item. The value of "1" indicated that aspect of life was not affected by the complications. A total score was obtained by summing the scores of the thirteen items, giving a range of possible scores of 13 (no complications or no effect) to 91 (severe adverse effects).

2.4.5 Diagnostic Certainty of Autoimmune Disease and Fibromyalgia. The blinded rheumatologists used clinical judgment in order to determine whether the subjects selected for a clinical examination had a number of different disorders, including fibromyalgia, current autoimmune disease or a past autoimmune disease. After reviewing the history, physical examination and laboratory tests, we recorded our impressions as a percent certainty for each diagnosis using 10 cm VAS. Fibromyalgia is a syndrome that is diagnosed based on the presence of widespread musculoskeletal pain and the presence of pain on digital palpation of at least 11 out of 18 tender points at defined anatomical locations (4). A specific case definition for current or past autoimmune disease was not detailed. The latter two variables were designed to capture clinical suspicions that an autoimmune condition may have been present but could not be classified according to diagnostic criteria for known disorders. In rheumatological practice, it is well recognized that certain rheumatic diseases may present with undifferentiated symptoms that may take years to evolve to the point where a firm

diagnosis can be established. Therefore, the intent was to screen for possible early or atypical autoimmune disease.

2.5 STATISTICAL ANALYSES

2.5.1 Descriptive Analysis of Symptom Data. A descriptive analysis of symptoms was undertaken in order to quantitate the frequency of each symptom in our sample of cosmetic surgery patients, which included women who were exposed to SBI for a minimum of 8 years. Bar charts were used to summarize the proportion of subjects in each group with each symptom, and the time of onset of each symptom in relation to cosmetic surgery.

Symptoms reported to have onset "after" plastic surgery were considered to be incident symptoms relative to the time of the procedure. This assumes that patient recall of the temporal relationship of symptom onset with respect to surgery is accurate. Because the precise date of onset of each symptom is unknown, it is not possible to express the result as an incidence density (i.e. number of events per person-years of follow-up). Instead, the number of subjects who developed a given symptom after surgery was divided by the number of subjects who were free of that symptom at the start of the exposure period. In other words, the denominator excluded individuals who reported that they had the symptom either "before" or "before and after" the cosmetic procedure. These proportions are referred to as the frequency of incident symptoms. The calculation is analogous to a cumulative incidence rate, which is defined as the proportion of subjects who experience the onset of a given event during a specified time interval that

is generally the same for all subjects (96). However, the proportions calculated do not adhere strictly to the latter definition since the length of follow-up varied between 8 to 16 years. The longer the duration of follow-up, the greater the opportunity for individuals to develop new symptoms. Due to the varying lengths of follow-up, and the inability to precisely define the presence and date of onset of each symptom, reporting of incidence rates and relative risks would be misleading. Furthermore, the main focus of this work is to explore the presence of multiple symptoms rather than to estimate the relative risk of individual symptoms.

The proportion of subjects with each of the self-reported symptoms "before" or "before and after" cosmetic surgery was also estimated using the total number of participants in each group as the denominator. Each proportion was considered to be an estimate of the prevalence of that symptom at the time of cosmetic surgery. This information was of interest mainly to assess whether the exposure groups had similar perceptions regarding their symptomatic state at the time of surgery. If women with SBI are more likely to attribute their current symptoms to their breast implants, then they may report a lower prevalence of symptoms with onset before surgery and a higher frequency of symptoms with onset after surgery compared with controls.

Symptoms reported as occurring "after" or "before and after" cosmetic surgery were assumed to represent prevalent symptoms at the time the current study was conducted. The frequency of incident symptoms, the prevalence of symptoms at the time of surgery and the prevalence of current symptoms were expressed as percents for the three exposure groups and tabulated separately. While statistical testing to compare the

frequency of each symptom between groups was not a planned objective of the analysis, chi-squared statistics comparing the three groups were performed to facilitate appreciation of the main differences between groups.

The number of self-reported symptoms with onset after surgery was summed for each individual. A frequency histogram of symptom number per subject was plotted for each exposure group in order to assess the shape of the distribution and the proportion of subjects with multiple symptoms. Because the distribution of symptom number per subject was highly skewed, the median and interquartile range are presented and nonparametric tests were used to assess differences between groups. The three groups were compared using the Kruskal-Wallis test, and pairwise differences were assessed using the Mann-Whitney U test (97).

2.5.2 Descriptive Analysis of Health Status. The distribution of scores for the quality of life, global health, and Illness Intrusiveness Rating Scale were presented as box-and-whisker plots for each group. The median and interquartile range were presented as the measure of central tendency since the distributions were skewed. The mean and 95% confidence intervals were also presented to allow comparison between this study and values reported in the literature.

2.5.3 Agreement Between Patient- and Physician-Recorded Symptoms. Paired responses for the 18 symptoms recorded on the self-administered questionnaire and the physician interview were cross-tabulated. The resulting two-by-two tables were

pooled to allow an overall assessment of the agreement between the self-report and interview data. Because the blinded physician recorded current symptoms but not the date of onset, the physician record was compared to self-reported symptoms present "after" or "before and after" surgery (i.e. *prevalent* symptoms). A kappa statistic with 95% confidence intervals was calculated as a measure of agreement that corrects for the occurrence of the same response due to chance alone. Interpretation of kappa (κ) followed general guidelines (98) indicating that agreement was very good ($\kappa = 0.81 - 1.00$), good ($\kappa = 0.61 - 0.80$), moderate ($\kappa = 0.41 - 0.60$), fair ($\kappa = 0.21 - 0.40$) or poor ($\kappa < 0.20$).

2.5.4 Cluster Analysis of Symptoms. Cluster analysis is an exploratory technique used to search for structure or natural groupings in data based on measures of similarity or association between variables (99). In the present study, a cluster is considered to be a grouping of symptoms that are associated with one another. Cluster analysis is usually applied in order to detect previously unnoticed, potentially useful groupings that may lead to the generation of hypotheses that can be studied in the future.

The first step in a cluster analysis is to choose the variables to be analyzed since this becomes the frame of reference within which the clusters are established. This choice should reflect the researcher's judgment of relevance based on existing knowledge (100). It is not appropriate to input huge numbers of variables with the hope that cluster analysis will find some useful structure. In the present study, dichotomous variables representing the presence or absence of 23 self-reported incident symptoms among

women with silicone gel-filled breast implants were analyzed. These symptoms were originally recorded because they include typical features of CTD, as well as the most frequent complaints associated with breast implants. Seven of the 30 self-reported symptoms (Appendix A) were excluded in order to obtain a more manageable number of variables. The following symptoms were judged to have less relevance because they are not typical concerns associated with breast implants: abnormal vaginal bleeding, menstrual problems, cessation of menstrual periods, miscarriages, and seizures. The symptom "skin ulcers" was omitted because it was so rare. Finally, the symptom "hand deformities" was excluded because it has not typically been associated with SBI but was originally included to screen for severe arthritis associated with the CTDs described in the primary cohort study. The 23 symptoms analyzed represent a wide range of specific and nonspecific symptoms. It was felt that further exclusion of symptoms might be too restrictive, thereby imposing structure on the data based on preconceptions of what the symptom groupings should be.

The next step in cluster analysis is to choose a measure of "similarity" between variables, referred to as a similarity or association coefficient. Symptoms represent binary data, such that pairs of symptoms can be arranged in the form of 2 x 2 contingency tables and a similarity coefficient calculated as the measure of association. There are a number of similarity coefficients available (99), however, Jaccard's coefficient shown below was selected because it gives a 1-1 match greater weight (i.e. both symptoms present) than a 0-0 match (i.e. both symptoms absent). This coefficient makes sense since it implies that the presence of both symptoms is stronger evidence of similarity than

Figure 2. Formula for calculation of Jaccard's similarity coefficient.

		Symptom 1	
		Yes	No
Symptom 2	Yes	a	b
	No	c	d

Similarity Coefficient (S) = $\frac{a}{a+b+c}$

the absence of both symptoms (99). A matrix of similarity coefficients between all pairs of symptoms was constructed for use in the cluster analysis.

A clustering algorithm known as hierarchical agglomeration was used because it is a data-dependent technique that does not require the investigator to make arbitrary *a priori* assumptions regarding the number or nature of the clusters (99). This algorithm starts with each symptom variable as an individual cluster. Symptom pairs with the highest similarity coefficients are the most closely associated and are merged first to form new groups. These initial clusters continue to be joined with other clusters and eventually, as the similarity decreases, all subgroups are fused into a single group. The mergers made at successive levels during the cluster analysis are displayed in a two-dimensional tree diagram or dendrogram. The dendrogram includes a "distance axis" in which the degree of dissimilarity at each merger is illustrated graphically by the length of the branch. The longer the branches before two symptoms join, the larger the distance and the less similar they are to one another. An example of a dendrogram will be presented in the Results (Section 3.5).

It was also necessary to select linkage criteria that the clustering algorithm uses to decide whether one variable or cluster is similar enough to another to be joined with it. It is recommended that the cluster analysis be repeated using several different linkage methods as a means of validating the cluster solution. Therefore, the current analysis used the four most common techniques, including single linkage, complete linkage, average linkage and the minimum variance method (99,100). These methods are also referred to in the literature by a number of other names (101). Each method is described briefly below, and a detailed example of the average linkage method is also presented in the Results.

The single linkage method defines the distance between two groups as the distance between their closest members (Figure 3a). Since single linkage joins clusters by the shortest link between them, it has difficulty discriminating between poorly separated clusters. The complete linkage method is exactly the opposite of single linkage. The distance between groups is defined as the distance between their most remote pair of members (Figure 3b). Thus complete linkage ensures that all members of a cluster are within some maximum distance of each other. Average linkage is intermediate between these two extremes and has been used extensively in biological sciences (102). It defines the distance between groups as the average of the distances between all pairs of members in the two groups (Figure 3c). Finally, the minimum variance method considers all possible pairs of clusters and, at each step, joins the pair whose fusion results in the smallest increase of the within-group variance.

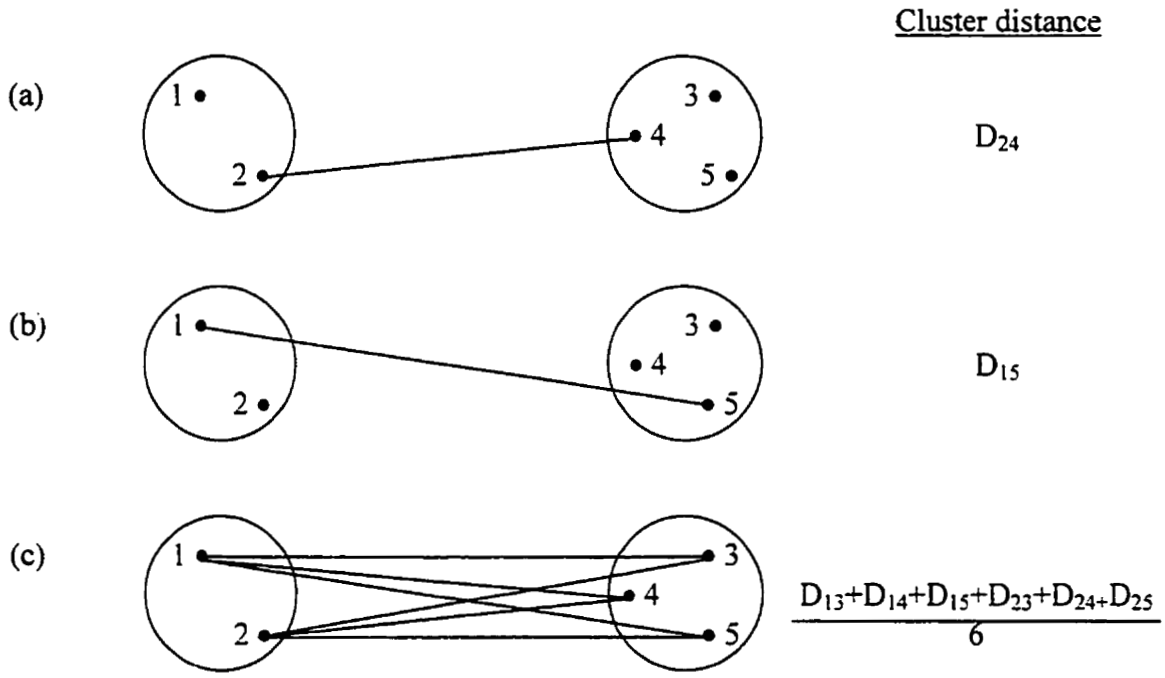


Figure 3: Schematic diagram illustrating inter-cluster distance (dissimilarity) for (a) single linkage, (b) complete linkage, and (c) average linkage [from Johnson and Wichern (99)].

As mentioned previously, each cluster analysis began with 23 clusters consisting of one symptom, and ended with a single large cluster containing all 23 symptoms. In most applications, the intermediate results consisting of a small number of clusters are of interest. There are no reliable quantitative methods for deciding how many clusters may be present in the data (100). The most useful method was applied in the present study, which is to simply examine the dendrogram for clusters that are visually apparent. A large increase in the distance at which the next symptom joins a cluster is often a clue that the symptom does not belong to that grouping.

Rigorous validation of a cluster solution involves replicating the solution across different linkage methods, across parallel data sets, and across a different collection of

variables or after introducing minor changes to the data (100). External validation of the cluster analysis is discussed in the next section (Section 2.5.5). Each analysis was performed using the four linkage methods discussed above. Parallel data sets were created by randomly dividing the silicone gel group into two halves, and performing independent analyses on each half. The random number generator available in SYSTAT was used for this purpose. Although the analysis of incident symptoms was of primary interest in the present study, minor changes to the data were introduced by using prevalent symptoms in the analysis. Finally, the analysis was replicated using a different set of variables based on physician-recorded symptoms.

Clusters that consistently emerged through the above validation procedures were considered to be useful solutions. For a given cluster of symptoms, a "cluster score" was calculated for each subject by summing the number of symptoms they had experienced from that grouping. Cluster scores were compared between the three exposure groups using the same nonparametric statistics described above (Section 2.5.1) because the distribution of scores was highly skewed.

In order to qualitatively compare symptom clusters from women with silicone gel implants with the other two groups, the above cluster analyses and validation procedures were also used to evaluate symptoms reported by the saline and control groups. The only exception was that these exploratory analyses were run only once using all the data from each group rather than dividing the data into two parallel data sets, as described for the silicone gel group.

2.5.5 Correlation Between Cluster Score and Health Status. Correlations between cluster scores and health status were analyzed for two main reasons. First, we were interested in assessing whether cluster symptoms might adversely affect health status. Second, one of the best ways to validate a clustering solution is to perform statistical analyses with relevant variables *not* used to generate the cluster solution (102). This procedure is referred to as "external validation" since it assesses the generality of the solution using variables that are external to the cluster analysis. If the members of a cluster differ from non-members for a number of relevant characteristics that were not used to determine membership, for example demographics or specific health outcomes, then this provides additional evidence that a useful result has been obtained. The value of a cluster solution that has been externally validated is said to be greater than a solution that has not (102). Nevertheless, even cluster results that have been externally validated should be considered exploratory and require testing in other samples of subjects.

The choice of relevant external criteria may be difficult in exploratory analyses when there is no content knowledge to help the investigator judge what criteria are appropriate. In the current study, we were interested in determining whether women with breast implants developed an identifiable cluster of symptoms and whether these symptoms were associated with reduced health status, specific demographic characteristics or the presence of an autoimmune syndrome. These associations were explored by assessing whether cluster scores were correlated with age, body mass index, quality of life, global health, illness intrusiveness, or the diagnostic certainty that the patient had an autoimmune disorder or fibromyalgia. Spearman's rank correlation

coefficients were used instead of Pearson's correlation coefficients since the former method has the advantage of assessing general associations while the latter method assesses linear associations (103). This choice was appropriate since scatter plots showed no evidence of a linear relationship between the variables of interest. Furthermore, calculation of Pearson coefficients requires at least one of each pair of variables to be Normally distributed (103). With the exception of age and body mass index, the variables included in this analysis were not Normally distributed. Tests to determine whether the correlation coefficients differed significantly from zero were performed using SYSTAT based on a *t* distribution with (n – 2) degrees of freedom.

3.0 RESULTS

3.1 Excluded Subjects

A search of the SBI cohort database identified 369 participants with silicone gel-filled breast implants, 120 subjects with saline implants and 271 controls with possible diseases associated with multiple symptoms that would require exclusion from the symptom analysis. The paper charts of these 760 participants were reviewed in order to verify the presence of disease and therefore determine whether exclusion from the current study was necessary.

The chart review identified 96 silicone (8.6% of total), 43 saline (12.2%) and 117 (16.1%) control subjects who met the exclusion criteria of this secondary analysis (Table 1). Of the eight categories of disorders listed in Table 1, the most common reasons for exclusion in all three exposure groups were cancer and rheumatic diseases. Gastrointestinal and neurological disorders occurred with similar frequency in the three groups. The category "other neurological" disorders included permanent deficits resulting from: stroke (1 control); encephalitis (1 control); polio (1 saline); peripheral multifocal neuropathy (1 silicone); and spinal stenosis (1 silicone). Diabetes mellitus and extreme obesity (BMI $\geq$ 45) were more frequent among controls than women with breast implants. Cardiac disorders, including ischemic heart disease (10 controls; 3 silicone) and congestive heart failure (2 controls; 2 silicone), were also more frequent in control subjects. Less frequent disorders, included in Table 1 under the category "Other", were: schizophrenia (1 control); narcolepsy (2 controls); reflex sympathetic dystrophy (1 control, 1 silicone); severe trauma (2 controls); diabetes insipidus (1 control);

Table 1. Proportion of subjects (%) in the silicone implant, saline implant and control groups excluded from the secondary analysis due to the presence of selected diseases.

Disorder	Silicone gel (N=1112) †	Saline (N=352)	Controls (N=726)
Cancer	**3.60**	**6.25**	**6.06**
Rheumatic Disease:	**2.52**	**3.11**	**4.01**
Rheumatoid arthritis	0.90	1.14	1.38
Systemic lupus	0.45	0.28	0.83
Vasculitides	0.27	0.28	0.14
Sjogren's syndrome	0.18	0	0.41
Juvenile rheumatoid	0.18	0.28	0.14
Psoriatic arthritis	0.18	0	0.14
Spondyloarthropathies	0.09	0.57	0.41
Sarcoidosis	0.09	0.28	0
Polymyalgia rheumatica	0.09	0	0.14
Discoid lupus	0.09	0	0.14
Dermatomyositis	0	0.28	0.14
Morphea	0	0	0.14
Gastrointestinal:	**0.90**	**1.42**	**1.79**
Crohn's disease	0.27	0.28	0.41
Ulcerative colitis	0.27	0.28	0.41
Celiac disease	0.27	0.57	0.14
Previous hepatitis B	0.09	0.28	0.69
Autoimmune hepatitis	0	0	0.14
Neurological:	**0.54**	**1.42**	**0.69**
Multiple sclerosis	0.36	1.14	0.41
Other neurological	0.18	0.28	0.28
Metabolic:	**0.27**	**0.28**	**3.17**
Type II diabetes mellitus	0	0	1.93
Severe obesity (BMI≥45)	0.27	0.28	1.24
Cardiac disease	**0.45**	**0**	**1.65**
Obstructive lung disease	**0.45**	**0.57**	**1.10**
Other disorders	**0.72**	**0.28**	**1.24**
Proportion Excluded ‡	**8.63 %**	**12.22 %**	**16.12 %**

† Values represent the per cent in each group with the diagnosis, and includes disorders that onset "before" or "after" cosmetic surgery.

‡ The sum of the column values is not equal to the proportion excluded because some subjects had more than one diagnosis.

hyperparathyroidism (1 silicone); hyperprolactinemia (1 silicone); untreated hypothyroidism with arthropathy (1 control); idiopathic thrombocytopenic purpura requiring splenectomy (2 silicone); alcohol abuse with subdural hematoma (1 silicone); alcoholic neuropathy due to vitamin B deficiency (1 silicone); and chronic pancreatitis (1 silicone, 1 saline, 1 control). Overall, a greater proportion of control subjects were excluded compared to the silicone group, primarily due to a higher prevalence of cancer, severe obesity, Type II diabetes mellitus and cardiac disease.

3.2 Analysis of Sociodemographic Data

Control participants were, on average, older than women with SBI by more than four years (Table 2a). The average time interval since cosmetic surgery was 12 years in the three exposure groups. Height was similar between groups, however, the controls were about 5 kg heavier compared to subjects with breast implants. This difference is also reflected in the body mass index values, and the proportion of subjects fulfilling accepted definitions for being underweight (BMI < 19), overweight (BMI > 27.3) and severely overweight (BMI >32.3) (88). Women with breast implants were twice as likely to be underweight compared to controls, while controls were about 5 times more likely to be severely overweight. Level of education was similar in the three groups.

A similar proportion of women in the three groups were employed in professional, technical or managerial positions, and in farming or other labor (Table 2b). Controls were less likely to work in clerical, sales or service industries and a greater proportion of controls were currently not employed outside the home (i.e. listed under

Table 2a. Continuous demographic characteristics of subjects in the silicone breast implant, saline breast implant and control groups.

Characteristic	**Silicone**, N=1016 Mean (95% CI)	**Saline**, N=309 Mean (95% CI)	**Control**, N=609 Mean (95% CI)
Age (years)	43.0 (42.5, 43.5)	43.4 (42.5, 44.2)	47.6 (46.6, 48.6)
Duration† (years)	11.8 (11.6, 12.0)	12.4 (12.0, 12.7)	11.7 (11.5, 11.9)
Height (cm)	164.2 (163.8, 164.6)	164.2 (163.8, 164.6)	163.7 (163.2, 164.3)
Missing (n)	5	0	3
Weight (kg)	61.3 (60.7, 61.9)	62.4 (61.2, 63.5)	67.8 (66.7, 68.9)
Missing (n)	9	3	10
Body mass index	22.7 (22.5, 22.9)	23.1 (22.7, 23.5)	25.3 (24.9, 25.7)
Missing (n)	11	3	11
< 19.0, % (95 CI)	7.6 (5.8, 9.6)	6.2 (3.4, 10.0)	3.3 (1.9, 5.4)
≥ 27.3, % (95 CI)	9.9 (7.8, 12.1)	12.4 (8.4, 17.2)	30.6 (26.4, 35.0)
≥ 32.3, % (95 CI)	1.6 (0.8, 2.7)	1.6 (0.4, 4.1)	7.9 (5.5, 10.6)
Education (years):	12.9 (12.8, 13.0)	13.1 (12.8, 13.3)	13.3 (13.1, 13.5)
Missing (n)	3	0	3
≥ 16 yr, %, 95 CI	15.9 (13.4, 18.6)	19.1 (14.2, 24.6)	22.1 (18.3, 26.0)

† Duration refers to the number of years since the date of cosmetic surgery.

occupation as "none") relative to the breast implant groups. Over 95% of subjects were Caucasian in each of the three groups (Table 2b). Women with silicone gel implants were slightly more likely to be married than controls, and less likely to be single or widowed. The proportion of subjects that were separated or divorced was similar between groups. In general, controls were less likely to have ever used tobacco relative to breast implant recipients. Alcohol use was similar in the three groups of women. About two thirds of subjects reported that they exercised regularly (≥ 1 to 3 times per week), irrespective of exposure status.

Table 2b. Categorical demographic characteristics of subjects in the silicone breast implant, saline breast implant and control groups.

Feature	**Silicone**, N=1016 % (95% CI)	**Saline**, N=309 % (95% CI)	**Control**, N=609 % (95% CI)
Occupation:			
Professional/Technical/ Managerial	27.4 (24.2, 30.6)	34.6 (28.5, 41.0)	27.8 (23.7, 32.1)
Clerical/Sales/Service	47.9 (44.3, 51.4)	44.8 (38.2, 51.2)	40.4 (35.9, 45.0)
Farming/Other Labor	3.9 (2.6, 5.4)	4.2 (2.0, 7.5)	4.5 (2.8, 6.7)
None	20.9 (18.1, 23.9)	16.3 (11.8, 21.6)	27.3 (23.3, 31.6)
Missing (n)	7	3	5
Ethnic Origin:			
Caucasian	96.5 (94.9, 97.6)	97.4 (94.4, 99.0)	99.0 (97.6, 99.7)
Marital Status:			
Married	81.0 (78.0, 83.7)	78.0 (72.1, 83.0)	74.9 (70.6, 78.7)
Separated/divorced	12.0 (9.8, 14.5)	13.9 (9.7, 18.9)	10.5 (7.9, 13.6)
Widowed	1.5 (0.7, 2.6)	0.6 (0.02, 2.5)	4.6 (2.9, 6.8)
Single	5.5 (4.0, 7.3)	7.4 (4.4, 11.4)	10.0 (7.4, 13.0)
Health Behaviours:			
Regular exercise	67.8 (64.4, 71.1)	66.6 (60.1, 72.4)	71.0 (66.6, 75.0)
Missing (n)	3	1	3
Tobacco use ever	61.4 (57.8, 64.8)	65.4 (58.9, 71.3)	55.8 (51.0, 60.2)
Missing (n)	1	0	0
Alcohol use ever	50.5 (46.9, 54.0)	60.8 (54.2, 67.0)	50.9 (46.2, 55.5)
$\geq$ 7 drinks/week ever	14.9 (12.4, 17.5)	15.9 (11.4, 21.0)	14.5 (11.4, 17.9)
Missing (n)	0	0	1

3.3 Descriptive Analysis of Self-Reported Symptoms

The purpose of this analysis was to report the frequency of symptoms and their temporal relationship to cosmetic surgery. Figures 2a, 2b and 2c are bar charts that summarize responses to the symptom questionnaire (Appendix A) for the silicone gel, saline and control groups, respectively. The symptoms are listed in ascending order of frequency for those with onset "after" surgery among the silicone gel breast implant

Figure 4a. Proportion of silicone gel implant recipients with each symptom and the time of occurrence relative to cosmetic surgery.

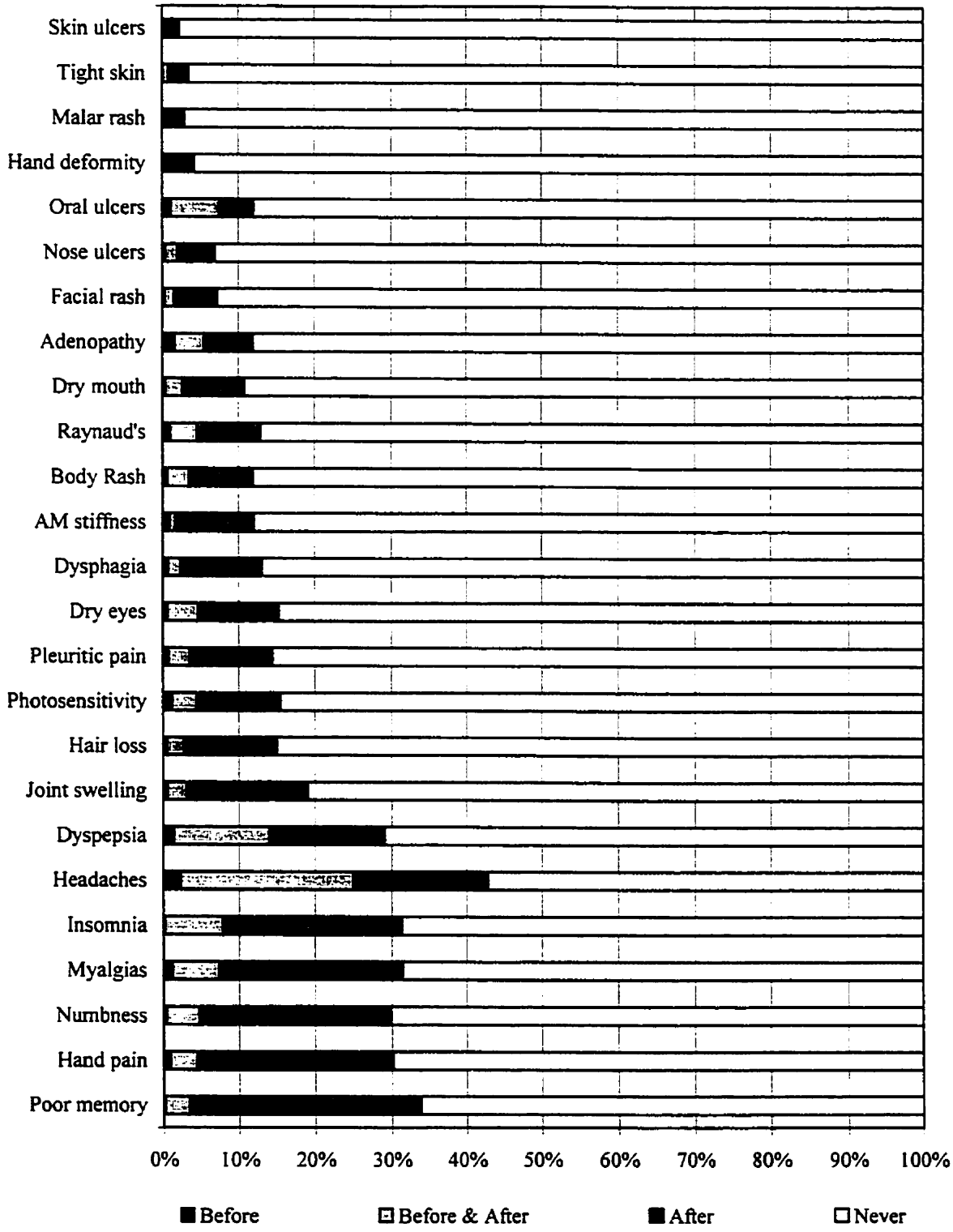


Figure 4b. Proportion of saline implant recipients with each symptom and the time of occurence relative to cosmetic surgery.

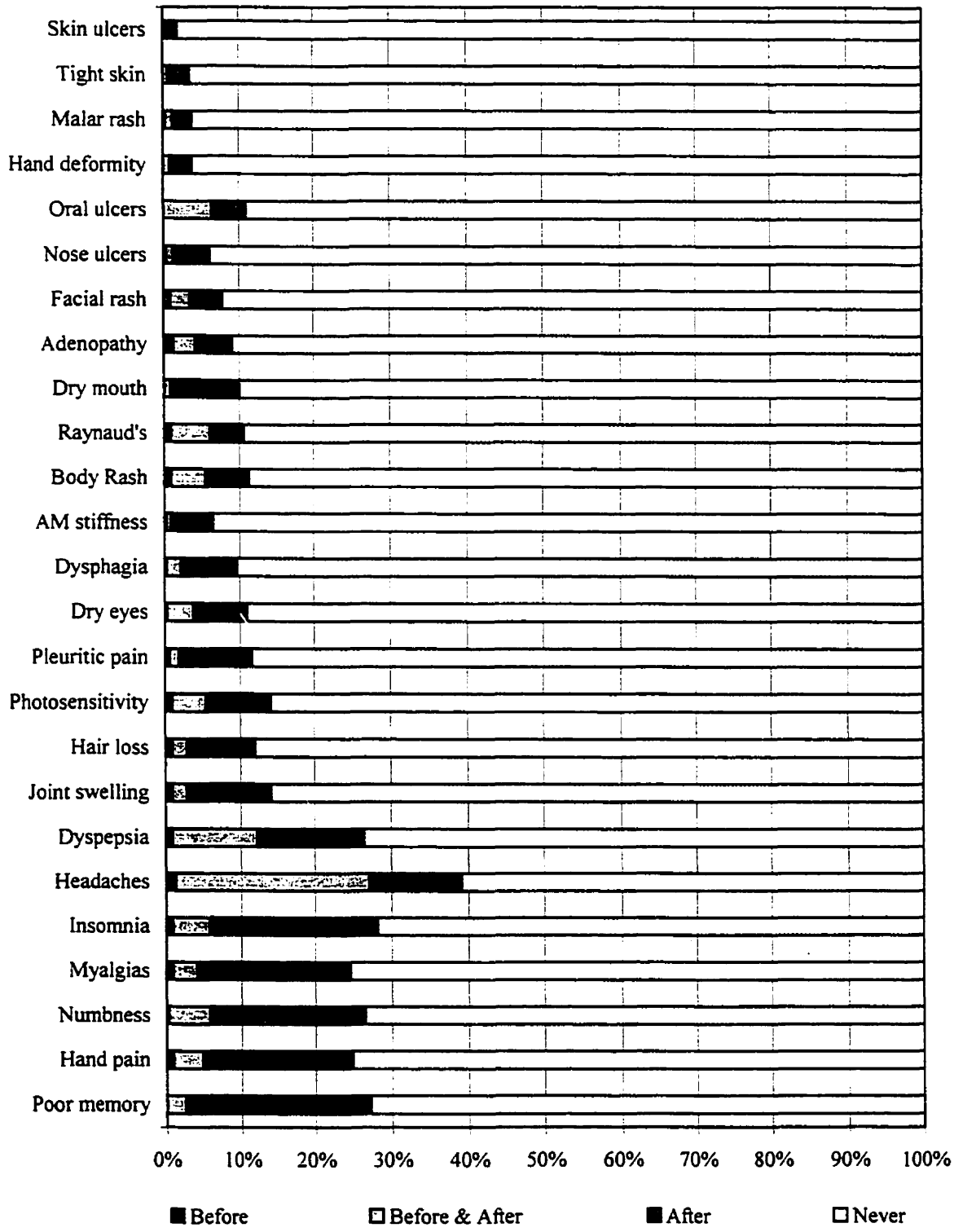
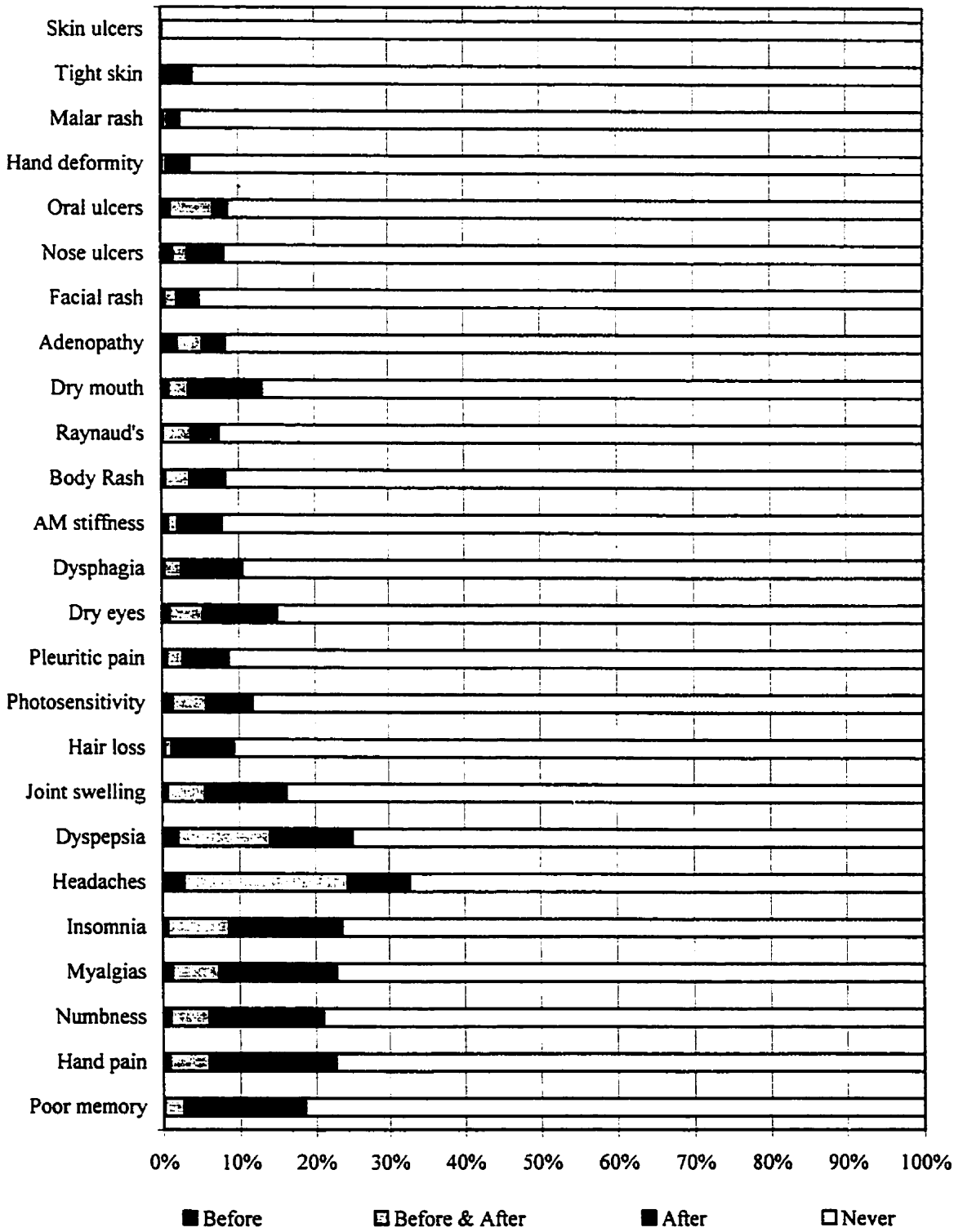


Figure 4c. Proportion of control subjects with each symptom and the time of symptom occurrence relative to cosmetic surgery.



group; this order is maintained in all three figures for simplicity. The majority of subjects in all three groups indicated that they had not experienced each symptom (labelled "never" in the figure legends). It was unusual for individuals to report any symptom as occurring only "before" their plastic surgery. Among silicone gel breast implant recipients who had experienced each symptom at some time after their cosmetic surgery ("after" or "before and after"), most recorded the onset of the symptom as starting "after" the surgery (Figure 4a). This was true for all symptoms except for oral ulcers and headaches. This trend was also observed for the saline (Figure 4b) and control groups (Figure 4c), but was less prominent. The data from these bar charts have been reorganized according to time of symptom onset and presented in reverse order in Tables 3a, 3b and 3c.

In general, the frequency of incident symptoms tended to be higher in the silicone gel group than the other two groups (Table 3a). Based on exploratory Pearson chi-squared analyses, the largest differences between groups were observed for poor memory, hand arthralgias, numbness, myalgia, insomnia, headaches, morning stiffness, and Raynaud's phenomenon. The chi-squared statistic tells us that a significant difference exists between at least two of the three groups for each of these symptoms, but does not tell us which of the three pairs of groups differ from one another. If the primary objective of the study was to compare the frequency of individual symptoms, then 24 pairwise comparisons for the above 8 symptoms could be performed using Fisher's exact tests. Simple inspection of the data reveals that the greatest difference was between the

Table 3a. Frequency (%) of self-reported incident symptoms in the silicone gel breast implant, saline breast implant and non-silicone cosmetic surgery control groups.

Symptom	**Silicone** (%)	**Saline** (%)	**Control** (%)	χ^2	**p value** †
Poor memory	31.5	25.3	16.2	45.23	<0.0005
Hand arthralgia	26.9	21.1	17.7	18.13	<0.0005
Numbness	26.4	22.0	15.9	22.66	<0.0005
Myalgia	26.0	21.3	16.7	17.78	<0.0005
Insomnia	25.5	23.7	16.4	17.05	<0.0005
Headaches	23.8	16.4	11.1	31.22	<0.0005
Dyspepsia	17.6	16.2	12.8	5.69	0.058
Joint swelling	16.5	11.7	11.3	3.58	0.167
Hair loss (scalp)	12.5	9.3	8.2	7.92	0.019
Photosensitivity	11.4	9.3	6.5	10.48	0.005
Pleuritic pain	11.2	9.9	6.1	11.59	0.003
Dry eyes	11.1	7.4	10.3	3.30	0.192
Dysphagia	10.9	7.6	8.1	4.87	0.088
AM stiffness	10.5	5.6	5.9	14.21	0.001
Body rash	8.7	6.2	4.8	8.91	0.012
Raynaud's	8.7	4.8	3.8	16.04	<0.0005
Dry mouth	8.4	9.2	9.9	1.01	0.603
Adenopathy	6.7	5.1	3.1	9.07	0.011
Facial rash	5.9	4.4	2.9	7.84	0.020
Nose ulcers	5.0	4.9	4.9	0.01	0.995
Oral ulcers	4.7	4.8	1.9	8.16	0.017
Hand deformity	3.5	2.9	3.0	0.39	0.822
Malar rash	2.9	2.6	1.5	3.10	0.212
Tight skin	2.6	2.9	3.6	1.44	0.487
Skin ulcers	1.6	1.6	0.2	7.55	0.023

† Incident symptoms are those that onset "after" cosmetic surgery.

Table 3b. Prevalence (%) of self-reported symptoms at the time of cosmetic surgery† in the silicone gel breast implant, saline breast implant and cosmetic surgery control groups.

Symptom	Silicone (N=1016)	Saline (N=309)	Control (N=609)	χ^2	p value‡
Poor memory	3.4	2.6	2.8	0.86	0.649
Hand arthralgia	4.5	4.9	6.1	1.93	0.381
Numbness	4.8	5.8	6.1	1.32	0.517
Myalgia	7.5	4.2	7.4	4.23	0.121
Insomnia	8.0	5.8	8.7	2.38	0.304
Headaches	25.0	27.2	24.5	0.84	0.657
Dyspepsia	14.0	12.3	14.1	0.66	0.717
Joint swelling	3.2	2.9	5.6	6.97	0.031
Hair loss (scalp)	2.9	2.9	1.3	4.35	0.114
Photosensitivity	4.5	5.5	5.8	1.32	0.516
Pleuritic pain	3.5	1.9	2.8	2.26	0.324
Dry eyes	4.7	3.9	5.4	1.09	0.580
Dysphagia	2.4	2.3	2.6	0.16	0.925
AM stiffness	1.7	1.0	2.1	1.67	0.434
Body rash	3.5	5.5	3.8	2.46	0.292
Raynaud's	4.6	6.2	3.9	2.26	0.324
Dry mouth	2.7	1.0	3.6	5.48	0.065
Adenopathy	5.6	4.2	5.4	0.94	0.625
Facial rash	1.6	3.6	2.1	4.63	0.099
Nose ulcers	2.1	1.3	3.5	4.99	0.082
Oral ulcers	7.7	6.5	6.9	0.67	0.716
Hand deformity	0.8	1.0	0.8	0.10	0.952
Malar rash	0.1	1.3	1.0	8.72	0.013
Tight skin	0.9	0.7	0.5	0.85	0.654
Skin ulcers	0.7	0.3	0	4.46	0.108

† Includes symptoms occurring "before" or "before and after" cosmetic surgery.

Table 3c. Prevalence (%) of current self-reported symptoms† in the silicone gel breast implant, saline breast implant and non-silicone cosmetic surgery control groups.

Symptom	Silicone (N=1016)	Saline (N=309)	Control (N=609)	χ^2	p value‡
Poor memory	33.7	27.2	18.4	44.29	<0.0005
Hand arthralgia	29.3	24.0	21.8	11.97	0.003
Numbness	29.5	26.2	20.0	17.85	<0.0005
Myalgia	30.3	23.6	21.7	16.17	<0.0005
Insomnia	31.2	27.2	23.0	12.86	0.002
Headaches	40.6	37.9	30.1	18.23	<0.0005
Dyspepsia	27.8	25.6	23.2	4.23	0.121
Joint swelling	18.5	13.3	15.6	5.51	0.064
Hair loss (scalp)	14.5	11.0	9.0	11.00	0.004
Photosensitivity	14.3	13.3	10.5	4.83	0.089
Pleuritic pain	13.7	11.0	8.1	12.72	0.002
Dry eyes	14.7	10.7	14.1	3.21	0.201
Dysphagia	12.3	9.4	10.2	2.91	0.234
AM stiffness	11.1	6.2	7.1	11.39	0.003
Body rash	11.3	10.4	7.9	4.99	0.082
Raynaud's	11.9	9.7	7.4	8.62	0.013
Dry mouth	10.4	10.0	12.2	1.45	0.485
Adenopathy	10.3	7.8	6.4	7.81	0.020
Facial rash	7.1	6.8	4.4	4.83	0.090
Nose ulcers	6.6	5.8	6.7	0.30	0.861
Oral ulcers	10.9	11.0	7.6	5.39	0.068
Hand deformity	3.8	3.9	3.6	0.07	0.968
Malar rash	3.0	3.6	2.1	1.74	0.419
Tight skin	3.4	3.6	4.1	0.48	0.786
Skin ulcers	1.8	1.6	0.2	8.46	0.015

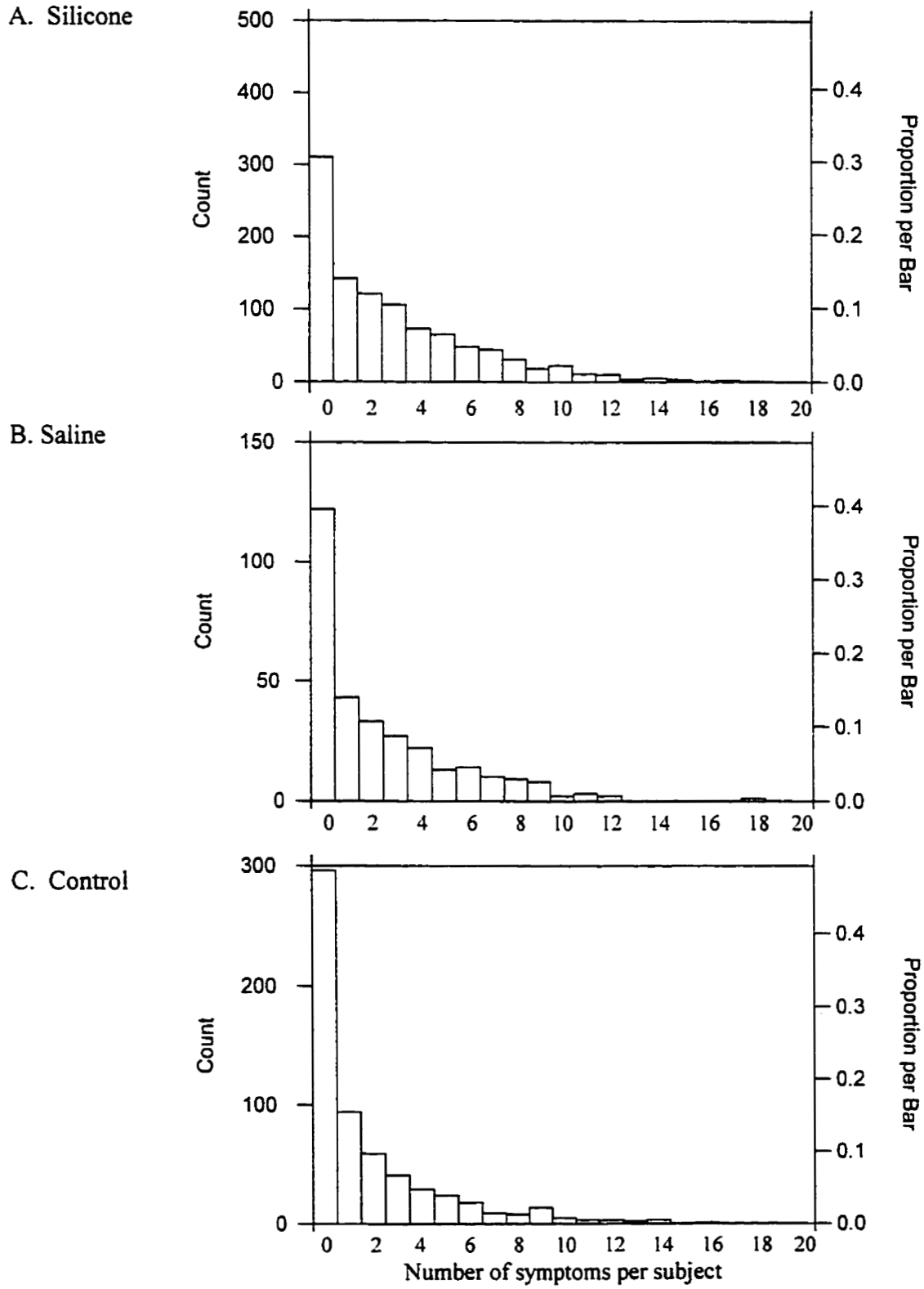
† Includes symptoms occurring "after" or "before and after" cosmetic surgery.

silicone gel and control groups, while the frequency among saline breast implant recipients was intermediate between these two groups.

The prevalence of all symptoms occurring "before" or "before and after" cosmetic surgery was remarkably similar between groups (Table 3b). This suggests that subjects in the three exposure groups had similar perceptions regarding their symptomatic state at the time of cosmetic surgery. The prevalence of current symptoms occurring "after" or "before and after" surgery (Table 3c) show the same trends as the incidence data (Table 3a). This is not surprising given that the prevalence of symptoms before surgery was similar between the groups, while the frequency of incident symptoms was higher in the silicone group.

In order to assess whether a subgroup of patients complain of multiple symptoms, the number of incident symptoms reported by each individual was summed and frequency histograms were plotted for each group (Figure 5). A greater proportion of women with silicone gel implants reported one or more incident symptoms (69.5%, 95% CI: 66.1, 72.7) compared to women with saline implants (60.5%, 95% CI: 53.9, 66.7) or controls (51.4%, 95% CI: 46.7, 55.9). Similar trends were apparent for the proportion of subjects with 4 or more incident symptoms among the silicone (33.2%, 95% CI: 29.8, 36.6), saline (27.2%, 95% CI: 21.6, 33.2) and control (19.5%, 95% CI: 16.0, 23.4) groups. The distribution of symptom number per subject was skewed such that the median and upper (UQ) and lower quartiles (LQ) were reported as the measure of central tendency. Overall, women in the silicone gel group were more likely to have multiple incident symptoms, with a median of 2 symptoms per subject (LQ, UQ: 0, 5) compared to

Figure 5. Histograms of number of incident symptoms per subject for the three groups.



the control median of 1 symptom per subject (LQ, UQ: 0, 3). The saline recipients had a median of 1 symptom per person (LQ, UQ: 0, 4). The Kruskal-Wallis test found the number of incident symptoms to differ significantly ($p<0.0005$) between the three groups. Mann-Whitney U tests determined that all three groups differed significantly from one another with respect to the number of incident symptoms per subject ($p \leq 0.002$).

3.4. Agreement Between Physician- and Self-Reported Symptoms

Of the 1,934 subjects included in this analysis, 932 met criteria for, and consented to, a clinical examination by a rheumatologist. Because symptomatic subjects were invited for a clinical assessment, the frequency of symptoms is inflated in this selected group such that a descriptive analysis of physician-reported symptoms has been omitted. Nevertheless, it was possible to use this data to assess agreement between physician-recorded and self-reported symptoms using the kappa statistic in order to see whether responses had changed during the face-to-face interview. In order to maintain blinding, the physician was unable to document the time of symptom onset relative to cosmetic surgery. Thus, the physician was able to record only symptoms that were prevalent at the time of the study such that agreement with self-reported prevalent (rather than incident) symptoms was analyzed. Figure 6 presents the pooled results for paired responses to the 18 symptom questions appearing on both the self- and physician-administered questionnaires. The numbers in parentheses represent the percent of responses falling into each cell of the 2x2 table.

Figure 6. Agreement between physician-recorded and self-reported prevalent symptoms.

A. SILICONE: Kappa (95 % CI) = 0.43 (0.40, 0.45)

	MD Yes	MD No	
Patient Yes	1098 (0.12)	1078 (0.12)	2176 (0.24)
Patient No	708 (0.08)	6125 (0.68)	6833 (0.76)
	1806 (0.20)	7203 (0.80)	9009 (1.0)

B. SALINE: Kappa (95 % CI) = 0.46 (0.42, 0.51)

	MD Yes	MD No	
Patient Yes	280 (0.13)	227 (0.10)	507 (0.23)
Patient No	181 (0.08)	1527 (0.69)	1708 (0.77)
	461 (0.21)	1754 (0.79)	2215 (1.0)

C. CONTROL: Kappa (95 % CI) = 0.45 (0.42, 0.48)

	MD Yes	MD No	
Patient Yes	578 (0.11)	450 (0.09)	1028 (0.20)
Patient No	458 (0.09)	3694 (0.71)	4152 (0.80)
	1036 (0.20)	4144 (0.80)	5180 (1.0)

Overall, 20% of responses were discordant and 80% were in agreement for all three exposure groups. The chance-corrected proportional agreement, or kappa (κ), was similar for the silicone (κ=0.43), saline (κ=0.46) and control (κ=0.45) groups. Review of the discordant pairs shows that the most frequent discrepancy among women with breast implants was the report of a symptom as being present when the physician recorded it as absent. In contrast, the two types of discrepancy occurred with equal frequency in the control group. In general, there was fair to moderate agreement (κ = 0.21 to 0.61) for the majority of individual symptoms (data not shown). The poorest agreement was observed for tight skin, with a kappa consistently $\leq$ 0.11. For the more common symptoms (myalgia, arthralgia, insomnia, headaches, joint swelling, numbness), the kappa values ranged between 0.33 and 0.61. In summary, there were no major differences in agreement between physician- and self-reported symptoms among the three groups. However, because overall agreement was only fair to moderate, exploratory cluster analysis of both self-reported and physician-recorded symptoms was performed.

3.5 Cluster Analysis of Self-Reported Incident Symptoms

Although it is difficult to assess the degree of recall bias in reporting of the temporal relationship of symptoms with respect to surgery, a decision was made to simplify the remaining analyses by focussing on *incident* self-reported symptoms. This choice is appropriate because we are most interested in exploring symptoms that might be attributable to breast implants based on the data available from this study. Furthermore,

exploratory analyses of prevalent symptoms found similar results such that this choice did not materially alter the findings of this study.

In order facilitate understanding of cluster analysis, a step-by-step description based on the first analysis will be given prior to presenting the remainder of the results. The first step of this procedure is to select the variables to be clustered and use them to generate a matrix of similarity coefficients. Twenty-three symptoms self-reported by half of the silicone gel group selected at random (Group 1) were included in the first analysis. Jaccard's similarity coefficient (S) was computed for each pair of symptoms, as shown in Figure 7a. The particular coefficient shown can be interpreted as the proportion of subjects with either insomnia or poor memory who have both of these symptoms (i.e. 38.2%). In the final results, cluster analysis traditionally reports the "distance" between two variables that are being joined to form a cluster, which is simply 1.0 minus the similarity coefficient (e.g. 1.0 – 0.382 = 0.618).

Figure 7a. Sample calculation of Jaccard's similarity coefficient.

		Insomnia Yes	No	
Poor Memory	Yes	81	87	168
	No	44	296	340
		125	383	508

Similarity Coefficient (S) $= \frac{a}{a+b+c} = \frac{81}{81+87+44} =$ **0.382**

The similarity coefficients are entered into a matrix for use by the cluster analysis algorithm. An abbreviated sample matrix is shown in Figure 7b (Step 1). At the beginning of the analysis, each individual symptom is considered to be a separate cluster. The matrix is reviewed by the computer, and the two symptoms with the largest similarity coefficient (i.e. smallest distance) are merged into a cluster. In this example, poor memory and insomnia join to form a new cluster "AB" which has two members. This requires the shaded columns to be collapsed into one column using one of the available linkage methods. The "average linkage" method shown in this example calculates the average similarity coefficient for all pairs of symptoms in the two clusters being joined. This value is entered into a new matrix (Step 2). For example, the new value in row C, column AB is equal to (0.307 + 0.262)/2 = 0.285. The process is then repeated using the new matrix (Step 2), such that hand pain and joint swelling have the highest similarity coefficient and are joined next. This results in two clusters with two members (AB and CD) and two clusters with one member (E and F) as shown in Step 3. The next step joins muscle pain (E) with cluster CD. The final step shown joins cluster AB with cluster CDE, again linking the shaded cells using average linkage as demonstrated in the sample calculation for Step 4. On completion of the analysis, symptoms (or clusters) with progressively smaller similarity coefficients are eventually joined to form one large cluster.

The sequence of steps resulting from the analysis of all incident symptoms reported by Group 1 is summarized in Table 4a, however, the results are presented most efficiently in the form of a dendrogram or tree diagram (Figure 8a). The lower the association

Figure 7b. Example of a similarity matrix and the average linkage method based on cluster analysis of incident symptoms for Group 1 of the silicone gel implant group.

Step 1:

	A. Poor Memory	B. Insomnia	C. Hand Pain	D. Joints Swell	E. Muscle Pain	F. Numb-ness
A. Poor memory	1.000	-	-	-	-	-
B. Insomnia	**0.382**	1.000	-	-	-	-
C. Hand Pain	0.307	0.262	1.000	-	-	-
D. Joints Swell	0.228	0.244	0.351	1.000	-	-
E. Muscle Pain	0.339	0.304	0.344	0.338	1.000	-
F. Numbness	0.319	0.335	0.340	0.300	0.337	1.000

Step 2:

	AB	C. Hand Pain	D. Joints Swell	E. Muscle Pain	F. Numbness
AB	1.000	-	-	-	-
C. Hand Pain	0.285	1.000	-	-	-
D. Joints Swell	0.236	**0.351**	1.000	-	-
E. Muscle Pain	0.322	0.344	0.338	1.000	-
F. Numbness	0.327	0.340	0.300	0.337	1.000

Step 3:

	AB	CD	E. Muscle Pain	F. Numbness
AB	1.000	-	-	-
CD	0.260	1.000	-	-
E. Muscle Pain	0.322	**0.341**	1.000	-
F. Numbness	0.327	0.320	0.337	1.000

Step 4:

	AB	CDE	F. Numbness
AB	1.000	-	-
CDE	0.281†	1.000	-
F. Numbness	**0.327**	0.326	1.000

† Sample Calculation for Similarity coefficient (AB)(CDE) using values from Step1

$= (S_{AC} + S_{AD} + S_{AE} + S_{BC} + S_{BD} + S_{BE}) / 6$

$= (0.307 + 0.228 + 0.339 + 0.262 + 0.244 + 0.304) / 6$

$=$ 0.281

between symptoms, the longer the "branches" of the dendrogram and the greater the distances (D) at which the symptoms join. The subgroup of symptoms consisting of poor memory, insomnia and numbness joined with a second subgroup consisting of hand pain, joint swelling and muscle pain early in the analysis (D=0.704). Headaches (D=0.751) and heartburn (D=0.763) were added last to form a natural grouping of 8 member symptoms. These 8 symptoms are left unshaded in the dendrogram (Figure 8a). Six other symptoms were then joined to form 3 two-member clusters which were unrelated to the initial 8 member cluster, including: malar rash and facial rash (D=0.765); dry eyes and dry mouth (D=0.789); photosensitivity and body rash (D=0.791). Thereafter, the remaining symptoms were joined at greater distances to form one large cluster of 23 symptoms.

As a means of validation, the above analysis was repeated for the second half of subjects (Group 2) with silicone gel breast implants. The same six symptom cluster formed early in the analysis, with headaches and heartburn again added last to form the same 8 member cluster (Table 4b, Figure 8b). Four additional small clusters were formed before all the symptoms joined together, including: malar rash and facial rash; dry mouth and dry eyes; body rash and hair loss; Raynaud's and photosensitivity. Only the first two are considered "stable" in that they joined together in the analyses for both Group 1 and Group 2.

Cluster analysis consistently identified the natural grouping of 8 symptoms using four different linkage methods (average, single, complete, minimum variance) applied to the two randomly selected halves of subjects with silicone gel implants. These 8 self-

Table 4a. Results of cluster analysis† of incident symptoms for Group 1 of the silicone gel breast implant group.

Cluster Containing AND Symptom	Cluster Containing Symptom	Were Joined at Distance	# of Members in New Cluster
Poor Memory	Insomnia	0.618	2
Hand Pain	Joint Swelling	0.649	2
Muscle Pain	Hand Arthralgia	0.659	3
Poor Memory	Numbness	0.673	3
Poor Memory	Muscle Pain	0.704	6
Poor Memory	Headaches	0.751	7
Poor Memory	Heartburn	0.763	8
Malar Rash	Facial Rash	0.765	2
Dry Mouth	Dry Eyes	0.789	2
Photosensitivity	Body Rash	0.791	2
AM Stiffness	Poor Memory	0.795	9
Dysphagia	Raynaud's	0.802	2
Dysphagia	AM Stiffness	0.819	11
Chest Pain	Adenopathy	0.841	2
Dry Mouth	Chest Pain	0.855	4
Dysphagia	Dry Mouth	0.860	15
Dysphagia	Hair Loss	0.867	16
Photosensitivity	Dysphagia	0.876	18
Malar Rash	Mouth Ulcers	0.876	3
Malar Rash	Nose Ulcers	0.911	4
Photosensitivity	Malar Rash	0.923	22
Photosensitivity	Tight Skin	0.966	23

† Cluster analysis was performed using the average linkage method.

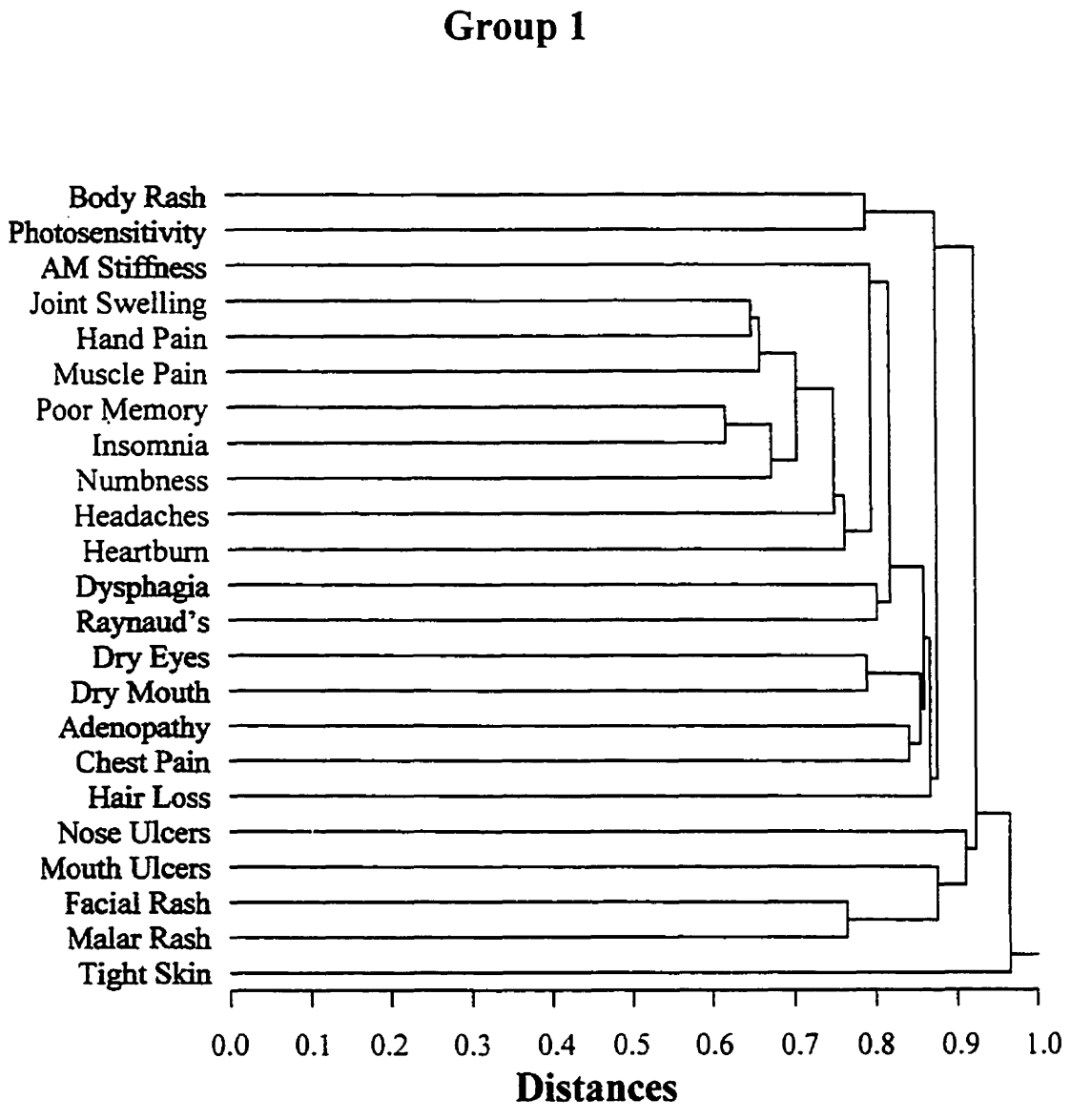


Figure 8a. Average linkage dendrogram for incident symptoms reported by Group 1 of the silicone gel breast implant group.

Table 4b. Results of cluster analysis† of incident symptoms for Group 2 of the silicone gel breast implant group.

Cluster Containing Symptom	AND	Cluster Containing Symptom	Were Joined at Distance	# of Members in New Cluster
Hand Pain		Joint Swelling	0.596	2
Poor Memory		Insomnia	0.597	2
Numbness		Muscle Pain	0.654	2
Numbness		Hand Pain	0.698	4
Numbness		Poor Memory	0.733	6
Malar Rash		Facial Rash	0.757	2
Headaches		Numbness	0.769	7
Dry Mouth		Dry Eyes	0.775	2
Heartburn		Headaches	0.785	8
Body Rash		Hair Loss	0.808	2
Raynaud's		Photosensitivity	0.811	2
Dysphagia		Dry Mouth	0.824	3
Heartburn		Chest Pain	0.826	9
Dysphagia		Heartburn	0.842	12
Raynaud's		Body Rash	0.855	4
Nose Ulcers		Malar Rash	0.865	3
Dysphagia		Raynaud's	0.876	16
AM Stiffness		Dysphagia	0.892	17
Nose Ulcers		AM Stiffness	0.904	20
Nose Ulcers		Tight Skin	0.911	21
Nose Ulcers		Adenopathy	0.919	22
Nose Ulcers		Mouth Ulcers	0.924	23

† Cluster analysis was performed using the average linkage method.

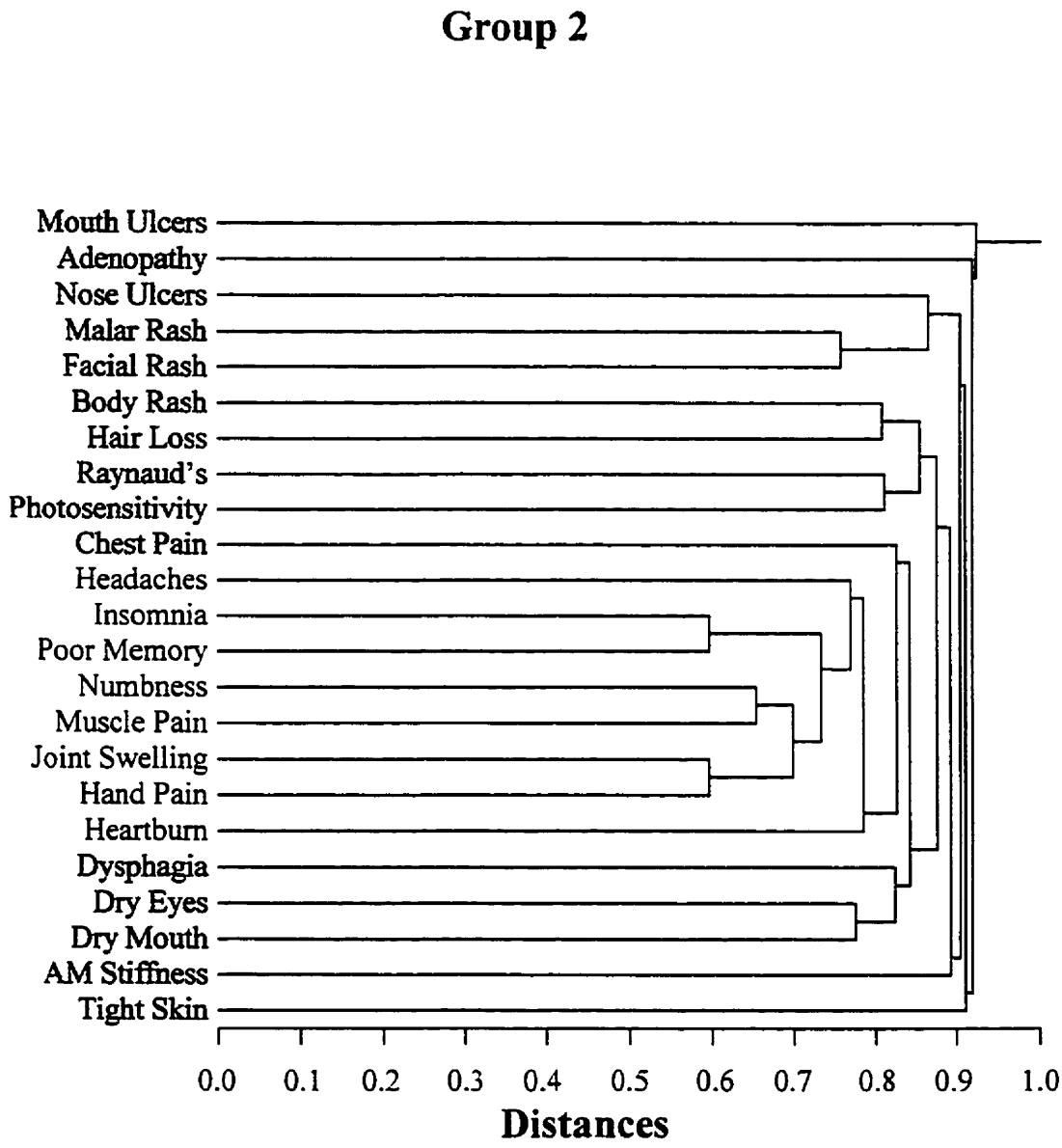


Figure 8b. Average linkage dendrogram for incident symptoms reported by Group 2 of the silicone gel breast implant group.

reported symptoms will be referred to as "cluster symptoms" throughout the remainder of this thesis. The same results were obtained irrespective of whether incident or prevalent symptoms (data not shown) were analyzed. Interestingly, similar findings were obtained on exploratory analysis of symptoms self-reported by the saline and control groups, irrespective of linkage method or whether incident or prevalent symptoms were used. The dendrograms for the above analyses have been excluded for the sake of brevity.

3.6 Cluster Analysis of Physician-Recorded Symptoms

Among the subjects included in this secondary analysis, 513 silicone gel, 127 saline and 292 control subjects underwent a clinical assessment by a blinded rheumatologist because of abnormal serology, or to confirm self-reported diagnoses and symptomatology. The two random halves of silicone gel recipients were again analyzed separately, including the 217 subjects in Group 1 and the 221 subjects in Group 2 with complete data for all 22 symptoms. The 75 subjects with incomplete data were excluded. Analysis of Group 1 symptoms joined fatigue, insomnia, muscle pain, joint pain and headaches first to form a 5 member cluster (Table 5a, Figure 9a). Numbness (D=0.718) and joint swelling (D=0.747) were then added to this initial cluster. Two additional small groupings were also formed at similar distances, including: dry eyes and dry mouth (D=0.699); and abdominal pain, constipation, and diarrhea (D=0.736). The remaining symptoms were not as strongly associated with one another and were ultimately joined to form the final 22 member cluster.

Table 5a. Results of cluster analysis† of physician-recorded symptoms for Group 1 of the silicone gel breast implant group.

Cluster Containing Symptom	AND	Cluster Containing Symptom	Were Joined at Distance	# of Members in New Cluster
Fatigue		Insomnia	0.532	2
Muscle Pain		Fatigue	0.551	3
Muscle Pain		Joint Pain	0.578	4
Muscle Pain		Headaches	0.649	5
Dry Mouth		Dry Eyes	0.699	2
Abdominal Pain		Constipation	0.702	2
Numbness		Muscle Pain	0.718	6
Diarrhea		Abdominal Pain	0.736	3
Joint Swelling		Numbness	0.747	7
Joint Swelling		Dry Mouth	0.765	9
Body Rash		Joint Swelling	0.805	10
Body Rash		Diarrhea	0.816	13
AM Stiffness		Pleuritic Pain	0.857	2
Dysphagia		Raynaud's	0.868	2
Photosensitivity		Body Rash	0.875	14
Dysphagia		Photosensitivity	0.878	16
Oral/Nasal Ulcers		AM Stiffness	0.894	3
Dysphagia		Oral/Nasal Ulcers	0.908	19
Tight Skin		Hair Loss	0.917	2
Tight Skin		Dysphagia	0.959	21
Tight Skin		Malar Rash	0.976	22

† Cluster analysis was performed using the average linkage method.

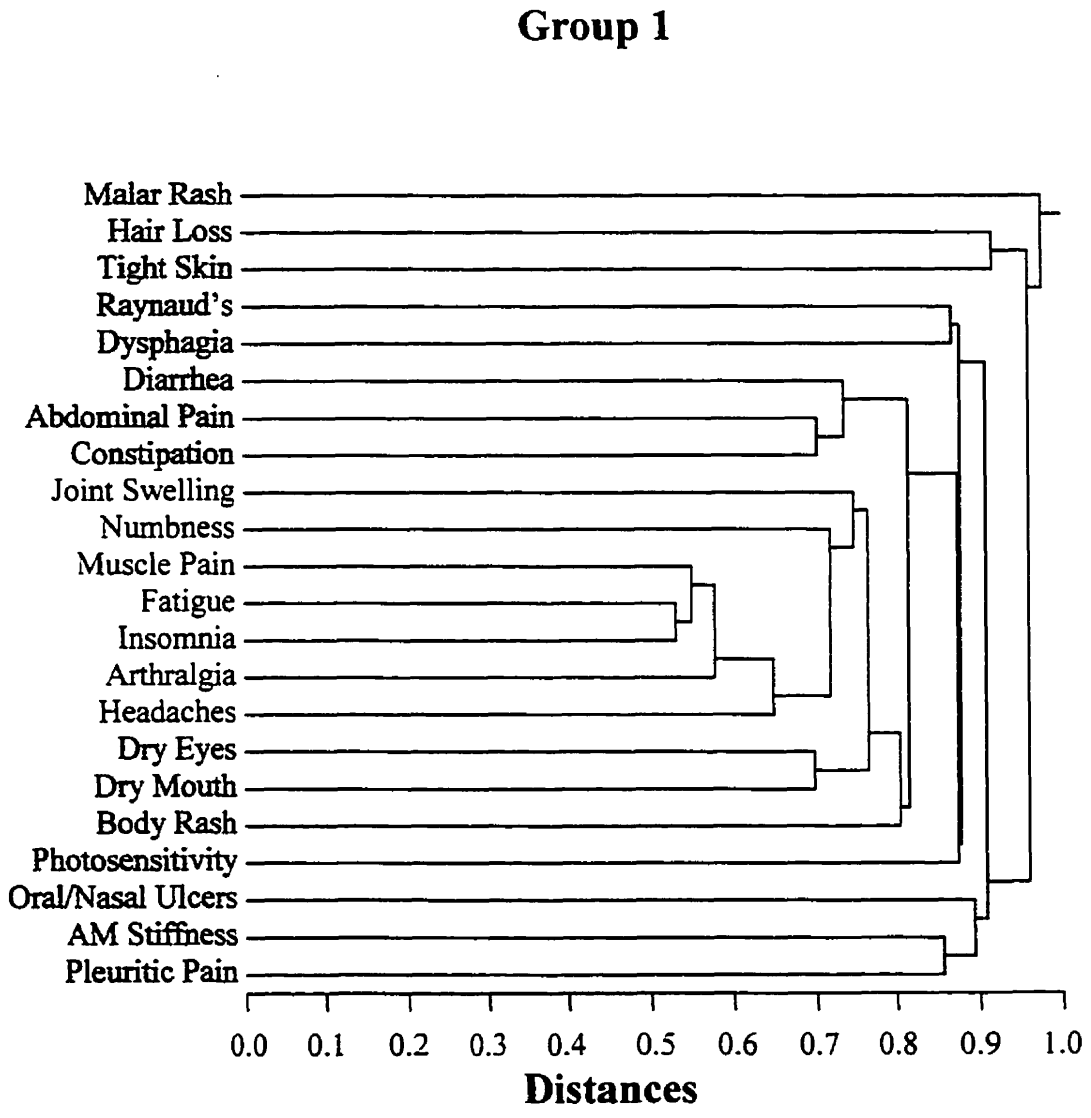


Figure 9a. Average linkage dendrogram of physician-recorded symptoms for Group 1 of the silicone gel breast implant group.

Table 5b. Results of cluster analysis† of physician-recorded symptoms for Group 2 of the silicone gel breast implant group.

Cluster Containing Symptom	AND	Cluster Containing Symptom	Were Joined at Distance	# of Members in New Cluster
Headaches		Joint Pain	0.554	2
Fatigue		Insomnia	0.591	2
Muscle Pain		Headaches	0.644	3
Fatigue		Muscle Pain	0.649	5
Numbness		Fatigue	0.715	6
Dry Mouth		Dry Eyes	0.729	2
Abdominal Pain		Diarrhea	0.756	2
Joint Swelling		Numbness	0.758	7
Constipation		Dry Mouth	0.769	3
Joint Swelling		Constipation	0.778	10
Abdominal Pain		Joint Swelling	0.843	12
Pleuritic Pain		Dysphagia	0.852	2
Abdominal Pain		AM Stiffness	0.861	13
Abdominal Pain		Body Rash	0.870	14
Raynaud's		Oral/Nasal Ulcers	0.889	2
Hair Loss		Chest Pain	0.896	3
Abdominal Pain		Photosensitivity	0.899	15
Hair Loss		Abdominal Pain	0.922	18
Hair Loss		Raynaud's	0.938	20
Hair Loss		Malar Rash	0.969	21
Tight Skin		Hair Loss	0.990	22

† Cluster analysis was performed using the average linkage method.

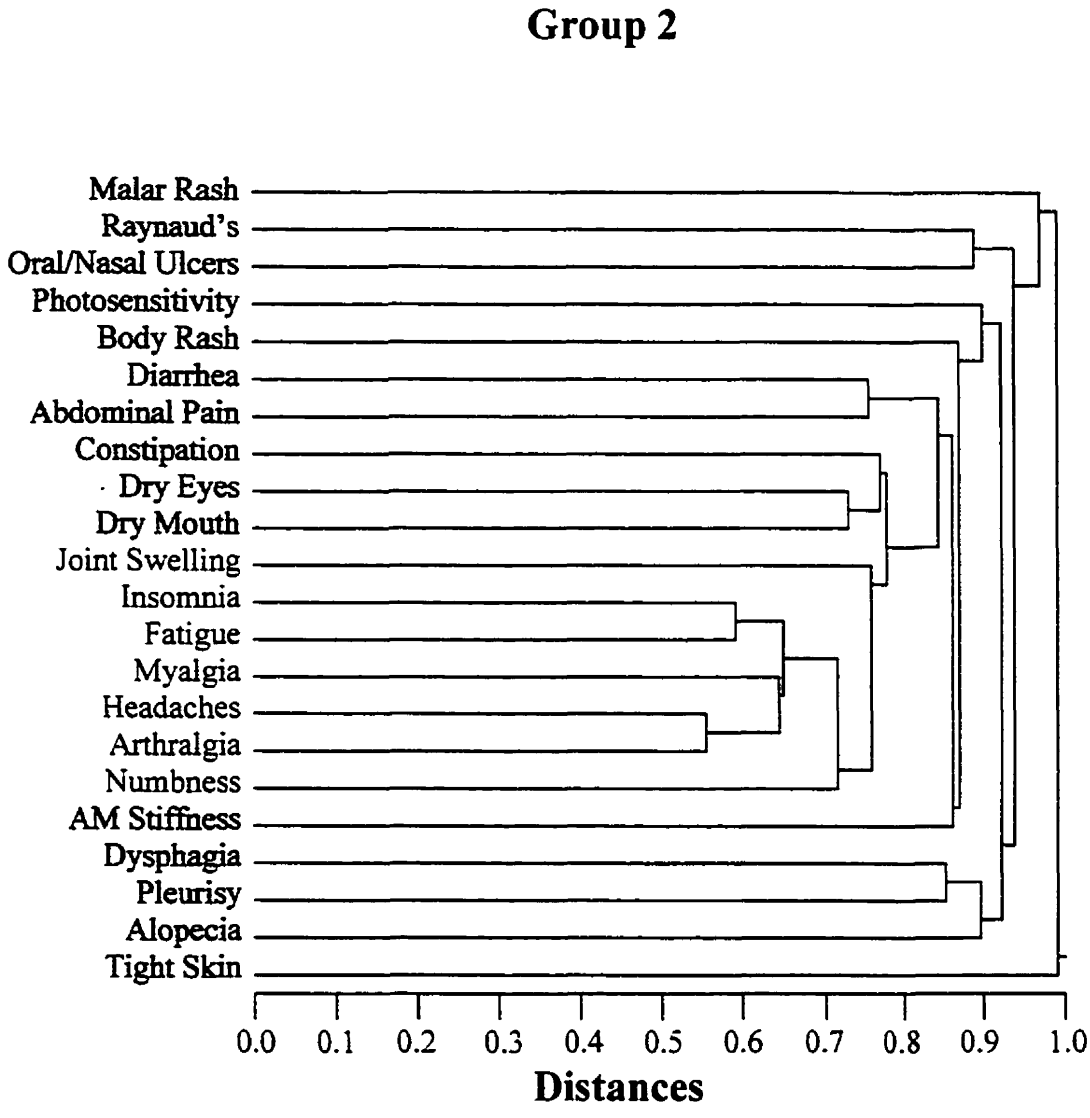


Figure 9b. Average linkage dendrogram of physician-recorded symptoms for Group 2 of the silicone gel breast implant recipients.

Validation using Group 2 data identified the same 7 member cluster consisting of headaches, joint pain, fatigue, insomnia, muscle pain, and numbness, with joint swelling again joining the group last at a distance of 0.758 (Table 5b, Figure 9b). Two small clusters also formed at a distance less than 0.758, including: dry eyes and dry mouth; and abdominal pain and diarrhea. The 7 member cluster was consistently obtained when the analysis was repeated using three alternative linkage methods (single, complete and minimum variance) in the two groups of silicone gel recipients (data not shown). The only exception occurred with complete linkage analysis of Group 1 data, in which a 10 member cluster was formed by adding dry eyes, dry mouth and constipation to the 7 core symptoms.

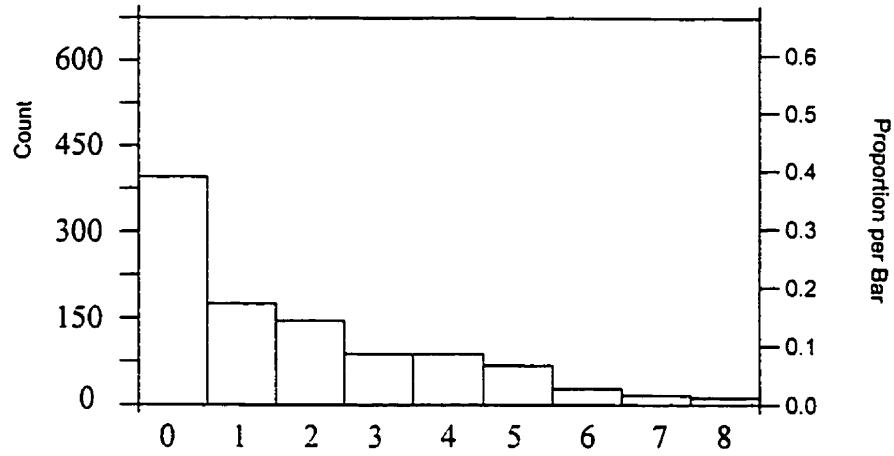
Exploratory analysis of physician-recorded symptoms was also performed for the 108 saline and 255 control subjects that were interviewed and had complete data for all 22 symptoms; 19 saline and 37 control subjects had missing data and were excluded. The main cluster of 7 symptoms described above was also demonstrated in the saline group using average, single, complete and minimum variance linkage methods. Results for the control group were slightly different relative to the two breast implant groups. All four linkage methods identified fatigue, insomnia, and headaches as being closely associated ($D \leq 0.650$). Three of the four linkage techniques also joined numbness, joint pain and swelling to this main cluster of symptoms ($D \leq 0.855$). One of the linkage techniques included muscle pain in this primary cluster for the control group, while the other three techniques joined this symptom with a second cluster containing abdominal pain, diarrhea and constipation.

3.7 Analysis of Cluster Scores

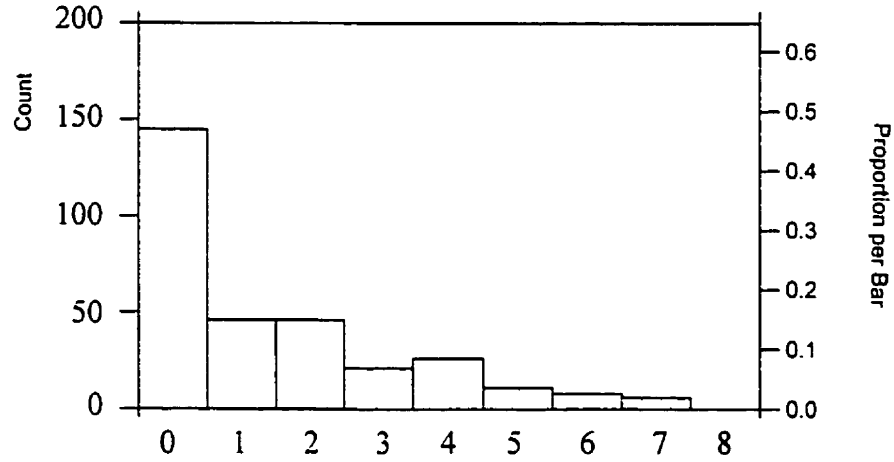
A cluster score was computed by summing the number of cluster symptoms each subject had from the grouping of eight self-reported incident symptoms. The distribution of cluster scores was highly skewed to the right (Figure 10). A median cluster score of 1.0 symptom per subject was observed for the silicone gel (LQ, UQ: 0 to 3) and saline (LQ, UQ: 0 to 2) groups. The median cluster score was 0 for the control group (LQ, UQ: 0 to 2). The three groups differed significantly from one another according to the Kruskal-Wallis test ($p<0.0005$). Mann Whitney U tests found that the silicone and saline groups differed from one another ($p=0.009$), as well as from the control group ($p<0.0005$). Control subjects were less likely to have a cluster score of one or more (40.4%, 95% CI: 35.9, 44.9) compared to the silicone gel (61.1%, 95% CI: 57.6, 64.5) and saline (53.1%, 95% CI: 46.4, 59.4) groups. The proportion of subjects in each group with four or more cluster symptoms was higher in the silicone gel (21.0%, 95% CI: 18.1, 24.0) and saline (16.5%, 95% CI: 12.0, 21.7) groups compared to controls (10.7%, 95% CI: 8.0, 13.8). The correlation between cluster score and health status is presented elsewhere (Section 3.9).

Figure 10: Frequency histograms of cluster scores for the silicone gel breast implant, saline breast implant and non-silicone cosmetic surgery control groups.

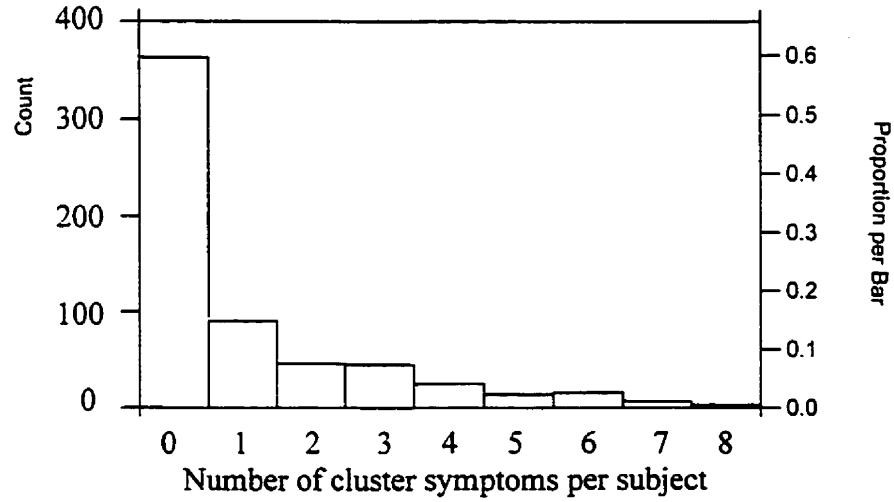
A. Silicone



B. Saline



C. Control



3.8 Descriptive Analysis of Health Status

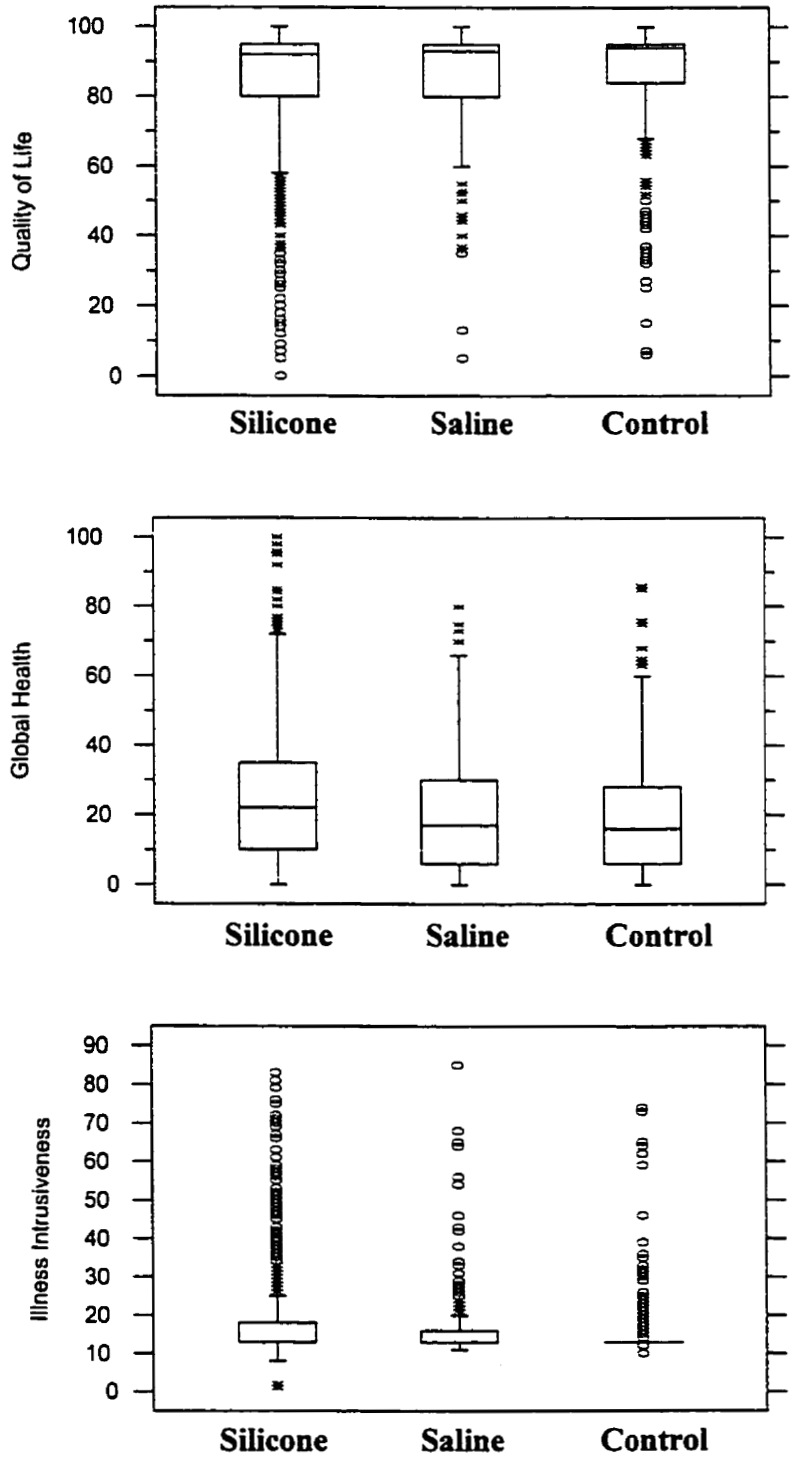
Data from the self-reported health status measures, including quality of life, global health, illness intrusiveness, and disability, were not Normally distributed (Table 6; Figure 11). The majority of subjects reported good health in all three exposure groups. The median quality of life score in millimeters, with 100 mm representing the best quality, was similar in the silicone gel (92 mm), saline (93 mm) and control (94 mm) groups (Table 6). Less than 10% of subjects had a quality of life score of 50 mm or less. Box plots reveal a similar range of values for quality of life in the three groups, with a large number of outliers in all groups (Figure 11a). The boxes indicate the 25^{th} and 75^{th} centiles, and the central line is the 50^{th} centile or median. The difference between the 25^{th} and 75^{th} centiles is the interquartile range. Asterisks appearing on box plots created by SYSTAT represent outliers, which are values that fall a distance of more than 1.5 times the interquartile range from the box. Open circles symbolize extreme outliers which are values that fall more than 3 times the interquartile range from the box.

Median self-reported global health scores were also similar in the silicone (22 mm), saline (17 mm) and control (16 mm) groups (Table 6, Figure 11b). Fair, poor or very poor global health scores ($\geq$ 40 mm on the 100 mm VAS) tended to be more frequent in the silicone group (16.3%) compared to saline (12.9%) and controls (10.5%). The physician-recorded global health scale showed similar findings, however, the spread of the scores was not as wide (Table 6). The median physician score of 90 mm for all groups, which is equivalent to 10 mm on the self-reported scale, suggests that physicians

Table 6. Health status of women with silicone gel breast implants, saline breast implants and non-silicone cosmetic surgery controls.

Health Status Instrument	Silicone gel	Saline	Control
Quality of Life (mm)			
Mean (95% CI)	83.9 (82.7, 85.0)	85.6 (83.9, 87.3)	85.5 (84.2, 86.8)
Median (LQ, UQ)	92 (80, 95)	93 (80, 95)	94 (84, 95)
Range	0 – 100	5 – 100	6 – 100
Score ≤ 50 mm (%)	7.6	4.2	6.1
N	1005	308	604
Global Health (mm)			
Mean (95% CI)	24.5 (23.3, 25.6)	21.6 (19.8, 23.6)	21.2 (20.0, 22.5)
Median (LQ, UQ)	22 (10, 35)	17 (6, 30)	16 (6, 28)
Range	0 – 100	0 – 80	0 – 86
Score ≥ 40 mm (%)	16.3	12.9	10.5
N	1010	307	605
MD Global Health (mm)			
Mean (95% CI)	86.4 (85.1, 87.6)	87.8 (85.9, 89.7)	87.2 (85.8, 88.7)
Median (LQ, UQ)	90 (80, 95)	90 (80, 95)	90 (80, 95)
Range	10 – 100	60 – 100	25 – 100
Score ≤ 40 mm (%)	1.2	0	3.4
N	512	126	292
Illness Intrusiveness			
Mean (95% CI)	17.9 (17.2, 18.5)	16.3 (15.3, 17.3)	14.7 (14.2, 15.2)
Median (LQ, UQ)	13 (13, 18)	13 (13, 18)	13 (13, 13)
Range	1 – 83	11 – 85	10 – 74
Score > 13 (%)	37.8	30.7	15.9
N	1004	305	601
MHAQ Disability Score			
Mean (95% CI)	0.05 (0.04, 0.06)	0.04 (0.03, 0.06)	0.06 (0.05, 0.08)
Median (LQ, UQ)	0 (0, 0)	0 (0, 0)	0 (0, 0)
Range	0 – 1.38	0 – 1.38	0 – 1.38
Score > 0 (%)	13.6	13.6	16.6
N	1005	309	606

Figure 11. Box plots of self-reported health status for women with silicone gel breast implants, saline breast implants and cosmetic surgery controls.



rated overall health as being slightly better than that perceived by the subjects themselves.

The Illness Intrusiveness Rating Scale is a means of assessing whether patients perceive that they have had complications due to cosmetic surgery and, if so, how they impact on various aspects of life. The lowest possible score (13.0) indicates that there have been no complications, or that complications did not intrude on daily life. The median score was 13.0 for all three exposure groups (Table 6). The spread of the data points is quite narrow, although a large number of outliers are visible, particularly for the silicone gel group (Figure 11c). The disability scale from the modified Stanford HAQ found no evidence of significant physical limitation in any of the three groups. Only 14% of breast implant recipients and 17% of controls recorded mean HAQ scores greater than zero, with a maximum recorded score of 1.38.

In summary, overall health status was not different in the three exposure groups, however, a large number of outliers were observed particularly for the silicone gel group. This suggests that a subgroup of subjects from each group perceived their health to be suboptimal in the absence of a defined systemic disease.

3.9 Relationship Between Cluster Scores and Health Status

Further exploration of the data was performed in order to assess whether the cluster score was associated with differences in self-reported health status or baseline characteristics. Tables 7a and 7b present Spearman correlation coefficients between cluster score and selected variables. Previous analyses indicated that age and body

Table 7a. Correlation of cluster score with age, body mass index and self-reported health status for the silicone gel breast implant, saline breast implant and non-silicone cosmetic surgery controls.

Feature	Silicone† (N = 976)	Saline† (N = 300)	Control† (N = 584)
Age	0.091 (0.004)	0.039 (0.505)	<0.0005 (>0.999)
Body Mass Index	0.155 (<0.0005)	0.205 (<0.0005)	0.042 (0.311)
Quality of Life	-0.353 (<0.0005)	-0.245 (<0.0005)	-0.203 (<0.0005)
Global Health	0.505 (<0.0005)	0.406 (<0.0005)	0.364 (<0.0005)
Illness Intrusiveness	0.428 (<0.0005)	0.272 (<0.0005)	0.273 (<0.0005)

† Values are Spearman rank correlation coefficients, with p values shown in parentheses.

weight differed between the three exposure groups (Table 2a). It is therefore possible that differences in cluster scores between groups could be related, at least in part, to the confounding effects of these two variables rather than the effect of the exposure alone. Because this analysis was exploratory in nature, it was not appropriate to stratify the analysis on age and weight. However, in order to explore whether cluster score (Section 3.7) was associated with age or weight, correlation coefficients were calculated (Table 7a). The association between number of cluster symptoms and age was negligible for all groups. There was a weak, positive correlation between body mass index and cluster score in the two breast implant groups, but not in controls. Thus, the number of cluster symptoms increased as weight increased among women with breast implants. Because breast implant recipients in the present study weighed an average of 5 kg less than controls, the impact of this weight difference on the results would tend to minimize

differences in cluster scores between groups and should therefore not influence the general conclusions.

Increasing cluster scores were associated with lower quality of life and global health in the three exposure groups (Table 7a). The correlation coefficients were all significantly different from zero ($p<0.0005$), and the size of the coefficients suggested a moderate degree of association. The strength of the association with quality of life and global health was greater in the silicone gel group (-0.353 and 0.505, respectively) than controls (-0.203 and 0.364), while saline recipients had correlation coefficients intermediate between these two groups. Interestingly, the strength of association between illness intrusiveness and cluster score was weaker in the saline (0.272) and control groups (0.273) but was appreciably higher in the silicone group (0.428). This result implies that symptoms may have had a greater impact on the lives of women with silicone gel breast implants, or that this group may be more likely to attribute their symptoms to breast implants.

3.10 Association Between Cluster Score and Fibromyalgia or Autoimmune Disease

The association between cluster score and the percent certainty of fibromyalgia documented by the examining rheumatologist was explored because the eight cluster symptoms identified are very similar to those reported by patients with this condition (81). In addition, correlation with percent certainty of current or past autoimmune disease was also computed in order to assess whether there was any association between increased number of these symptoms and the physicians' perception that an atypical

Table 7b. Correlation of cluster score with diagnostic certainty of fibromyalgia or autoimmune disease in the silicone gel breast implant, saline breast implant and non-silicone cosmetic surgery control groups.

Diagnoses (% certainty)	**Silicone†** (N = 483)	**Saline†** (N = 124)	**Control†** (N = 282)
Fibromyalgia	0.306 (<0.0005)	0.288 (0.001)	0.259 (<0.0005)
Current autoimmune disease	-0.096 (0.036)	-0.041 (0.654)	-0.004 (0.951)
Past autoimmune disease	-0.067 (0.139)	0.132 (0.144)	0.009 (0.882)

† Values are Spearman rank correlation coefficients, with p values shown in parentheses.

autoimmune disorder might be present (Table 7b). Higher cluster scores were significantly associated with increased certainty of fibromyalgia, with a similar strength of association (0.26 to 0.31) in all three groups. No significant association with current or past autoimmune disease was identified in any of the groups.

Table 8 presents the estimated proportion of examined subjects with fibromyalgia in each group and the associated 95% confidence intervals for these estimates. The proportion of subjects diagnosed with fibromyalgia by the examining rheumatologist was about 21% in all three exposure groups (Table 8). This does not represent the prevalence of fibromyalgia in the cohort, however, since symptomatic patients were selected for examination. Examined subjects in each exposure group were then stratified according to whether they had at least four cluster symptoms or less than four cluster symptoms. Participants with four or more cluster symptoms were two to three times more likely to have fibromyalgia compared to subjects with 3 symptoms or less; this was true for comparisons within each exposure group as well as between the three groups. The

Table 8. Proportion of subjects in the silicone gel breast implant, saline breast implant and cosmetic surgery control groups with high and low cluster scores that had fibromyalgia on clinical assessment.

Exposure group	Fibromyalgia† (%)	95% C.I.
Silicone (N=485)	**21.9**	17.7, 26.3
Cluster score < 4 (N=318)	15.1	10.7, 20.1
Cluster score ≥ 4 (N=167)	34.7	26.4, 43.4
Saline (N=124)	**21.0**	13.2, 30.2
Cluster score < 4 (N=82)	12.2	5.1, 22.4
Cluster score ≥ 4 (N=42)	38.1	21.2, 55.9
Control (N=282)	**20.9**	15.6, 26.8
Cluster score < 4 (N=225)	16.0	10.8, 22.2
Cluster score ≥ 4 (N=57)	40.4	25.5, 55.6

† Subjects were classified as having fibromyalgia if the examining rheumatologist indicated a percent certainty of diagnosis of ≥ 75%.

frequency of fibromyalgia within each of these two strata did not differ across the three exposure groups.

4.0 DISCUSSION

4.1 Study Participants

Subjects with known diseases were excluded from this analysis in order to focus the investigation on "unexplained" symptoms that might be related to breast implants. The frequency of certain diseases in the study cohort is unlikely to be representative of women of similar age in the general population since these subjects chose to undergo cosmetic surgery. Women with significant disease are less likely to undergo elective procedures. For example, women with Type I diabetes mellitus may be less likely to undergo cosmetic surgery due to increased risk of complications, such that this disease was under-represented in the cohort. Furthermore, the accuracy of self-reported diseases likely varied among diagnoses in this study. Rheumatic diseases, which were specifically listed on the questionnaire and which were confirmed by a clinical assessment, were likely recorded most accurately. Conversely, other self-reported diagnoses that were not listed on the questionnaire may have been under-reported if patients failed to specifically record them. Because the average age and weight of the control group was higher than the breast implant groups, it is not surprising that they were more likely to be excluded for diseases associated with aging and, or excessive weight, including cancer, Type II diabetes mellitus, cardiac disease and severe obesity.

Others have noted that women with breast implants have different demographic and lifestyle characteristics, which might confound studies of the health effects of SBI. For example, women with implants are less likely to be overweight, and may be more likely to drink alcohol regularly ($\geq$ 7 drinks per week) (90). We did not observe

differences in alcohol intake between groups, but did find that, on average, controls were 4 years older and 5 kg heavier. These differences between SBI and control groups are not surprising since thin women are more likely to obtain breast implants, and since controls had procedures typically associated with obesity and aging (i.e. liposuction, lipectomy, breast reduction, facial surgery). The proportion of breast implant recipients that were overweight (10%) was considerably less than that observed among controls (31%) and among American women over age 20 (36%) (88). Although women with breast implants may be more conscious of body image than other women, we did not find the proportion that were underweight (7%) to differ from that of American women over age 20 (5.7%)(88). The frequency of certain symptoms, such as muscle pain, joint pain and swelling, dyspepsia, and insomnia, might be expected to increase with age and weight. Thus, one might expect to observe a greater frequency of symptoms among controls. Since SBI recipients tended to have more symptoms, confounding due to age and weight would be expected to minimize, rather than exaggerate, between-group differences and should not alter the general conclusions of this study.

The demographic characteristics of women with breast implants included in this analysis are generally comparable to those reported in other epidemiological studies. The mean age ranged from 34 to 43 years in the three cohort studies that reported symptom frequency (7,62,63,69) compared to 43 years in this study. Breast implant recipients in previous reports from North America were predominantly (>90%) Caucasian (7,63,65,90). Our finding that women with silicone gel breast implants were slightly more likely than controls to be married and less likely to be widowed or single is

consistent with previous publications (7,55,57,90). Education level differed little between groups, and was similar to that of subjects recruited into other studies (57,90). Others also agree with our findings that a history of smoking tended to be more frequent in women with SBI compared with controls (7,90).

4.2 Self-Reported Symptoms

The descriptive analysis of incident symptoms identified several complaints that were more frequent in the silicone gel group compared to controls, including: poor memory, arthralgia, numbness, myalgia, insomnia, headaches, morning stiffness and Raynaud's phenomenon. The frequency of symptoms in the saline group was generally intermediate between the two other groups. Although the precise date of onset of each symptom was unknown, the exposed and unexposed groups had a similar average duration of follow-up, such that this analysis provides a reasonable indication of the relative frequency of complaints among the three groups.

The frequency of incident symptoms among women with silicone gel breast implants was compared to data published in the three historical cohort studies that specifically evaluated symptoms (62,63,69). Among the self-reported symptoms described in this study, 12 were also reported in one or more of the other studies. The frequencies of the 12 symptoms are summarized in Figure 12 for women with silicone gel-filled breast implants (62,69) or breast implants of unspecified type (63). Published data represent self-reported incident symptoms (62,63) or physician-recorded prevalent

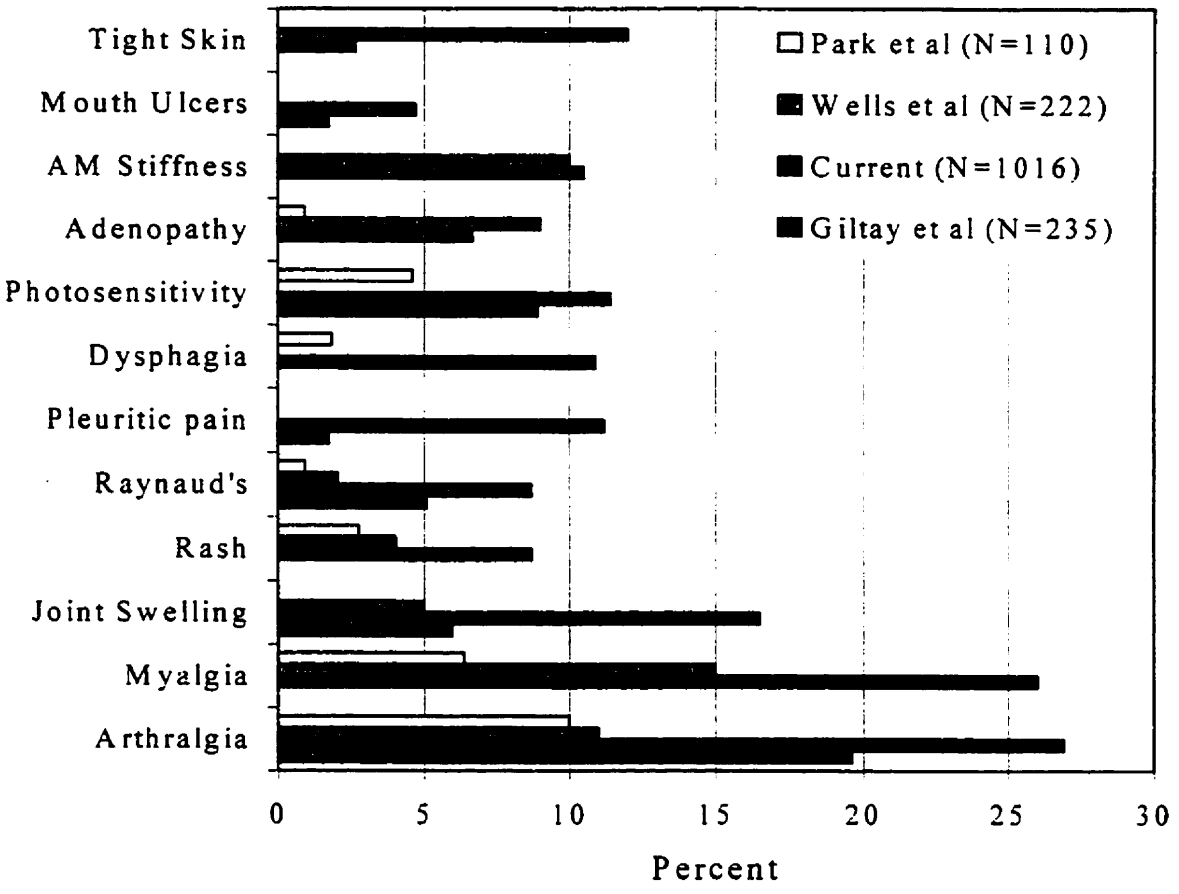


Figure 12: Frequency of symptoms among women with silicone breast implants reported in the literature and in the current study.

symptoms (69). Tight skin is the only symptom reported more commonly in the literature (63) than in this study, however, these publications did not describe how this symptom was defined such that misinterpretation may have occurred. Mouth ulcers (62) and morning stiffness (63) were reported by a similar proportion of cases as in this study. Adenopathy (63) and photosensitivity (62) were also similar in frequency to that reported herein, except in the study of Park et al (69). They found a lower frequency for all symptoms compared to the other studies. Because the latter study documented prevalent

symptoms, one might have expected higher frequencies relative to incident symptoms in the cohorts with a similar duration of follow-up (62,63). It is possible that data collection by an unmasked plastic surgeon may have resulted in a lower estimate of symptom frequency relative to the self-reported data. The frequencies of arthralgia, joint swelling, myalgia, rash, Raynaud's phenomenon, dysphagia and pleuritic pain were 5% to 10% higher in the current study than that reported elsewhere (Figure 12). These differences may be due to the fact that the duration of post-surgical follow-up was at least 6 years longer in the present study. Furthermore, the potential for recall bias may have been greater in 1994 when this study was conducted compared with the others which were completed between 1991 and 1993. Awareness regarding the possible health effects of silicone gel breast implants has continued to grow since widespread media coverage started in 1990, and since the North American moratorium on their use commenced in 1992.

Only one cohort study (62) reported the mean number of incident symptoms per subject for the silicone gel group (0.6 symptoms per subject) and controls (0.3 symptoms per subject) ($p<0.001$). The present study also found a significant difference ($p<0.0005$) between the silicone gel (mean 2.9, median 2.0 per subject) and control subjects (mean 1.8, median 1.0 per subject). In the study of Giltay *et al*, 21% of controls and 37% of the silicone gel group had one or more incident symptoms (62). Analogous results were found in the present study (52% and 70%, respectively), but again subjects in our study tended to have more symptoms. These different results are likely related to the greater

number of symptoms listed on our questionnaire, the greater duration of follow-up in the present study, and possibly recall bias.

Indeed, recall bias represents a potential problem in studying symptoms since the outcome of interest is a subjective complaint that cannot be easily validated using medical records or other sources. A natural reaction of women with SBI to the recent media attention surrounding breast implants would be to pay special attention to any symptoms. It would thus not be surprising for SBI-exposed women to report more symptoms, and also be more likely to recall the onset of the symptoms as occurring "after" the surgery took place instead of "before". In the present analysis, there was no difference in prevalence of symptoms at the time of plastic surgery. This result suggests that major bias in reporting of the temporal relationship of symptoms with respect to surgery was unlikely, assuming that all groups were equally likely to recall and record symptoms in general. However, the latter assumption may not hold true if recall bias was present and subjects with breast implants were more likely to report symptoms overall. In this instance, one cannot conclude that bias in reporting the temporal relationship was absent. For example, controls may have under-reported symptoms occurring either before or after surgery, while the silicone group under-reported those occurring before and over-reported those occurring after. This would exaggerate differences in the frequency of incident symptoms whereas prevalent symptoms at the time of surgery would appear to have the same frequency in the exposed and unexposed groups. It is impossible to assess the degree to which recall bias might have influenced the results since we must rely on patient report as our sole source of information regarding

symptoms and their time of onset. Nevertheless, it is possible to assess whether subjects may have misinterpreted the meaning of the symptom descriptions by comparing the self-reported data with that clarified during the physician interview (Section 4.3).

4.3 Agreement Between Interview and Self-Reported Data

Epidemiological studies make extensive use of questionnaires, however, the validity of self-reported data is often not assessed. Incompleteness and under-reporting tend to be greatest for mailed questionnaires, whereas face-to-face interviews allow for clarification of information and avoidance of unintentional omission of responses (71). While severe diseases and those with clear diagnostic criteria are recorded most accurately, diagnostically complex diseases and ill-defined conditions or symptoms are less likely to be recorded accurately (71,74). For example, in a study of the test-retest reliability of a population survey administered one week apart, 13% fewer symptoms were recorded on the second survey (82).

In the present study, a face-to-face interview was performed within one year of the self-administered questionnaire only for a subgroup of patients who were symptomatic or had possible autoimmune disease based on self-report or blood tests. Eighteen symptoms were recorded by both the patient and the physician and were therefore available for comparison of responses. The chance-corrected proportional agreement (kappa) between the patient and physician was moderate (0.43 to 0.46) for all three exposure groups.

Discordant responses occurred when either the physician or the subject noted that a symptom was present while the other did not. There are a number of possible explanations why the two types of discrepancy may have occurred. For example, the symptom may have started or resolved during the time interval between the questionnaire and the interview. Alternatively, the subject may have misinterpreted the meaning of a symptom descriptor which the physician was able to clarify during the interview. Misinterpretation might be more likely to result in the patient recording a symptom as present when the physician did not. This is because certain nonspecific symptoms, (e.g. hair loss or tight skin) may be extremely common in the general population, while the intended specific medical meaning may be uncommon (e.g. significant alopecia or scleroderma).

Recall bias may also have contributed to discrepancies if women with SBI were more vigilant in reporting minor symptoms which the physician judged to represent normal bodily sensations. Indeed, the current study found that women with silicone gel implants were more likely to report symptoms as being present when the physician did not (12% of paired responses) compared to controls (9% of paired responses). In contrast, control and SBI subjects were equally likely to report a symptom as absent when the physician recorded it as present (9% vs 8% of paired responses, respectively). It is possible that the degree to which controls may have under-reported their symptoms was under-estimated by this analysis, since subjects that did not report any symptoms were less likely to be called in for a physician interview.

In summary, women with breast implants were somewhat more likely than controls to report a symptom as present when the physician did not. Overall, there were no major differences in agreement between physician- and self-reported symptoms among the three groups. However, because agreement was only moderate, exploratory cluster analysis of both self-reported and physician-recorded symptoms was performed.

4.4 Cluster Analysis

Cluster analysis of incident symptoms reported by women with silicone gel breast implants identified eight symptoms that appeared to form a natural grouping. The most closely associated symptoms included poor memory, insomnia, numbness, hand pain, joint swelling, and muscle pain. One may speculate whether or not headaches and heartburn also belong to this cluster since they were more weakly associated and therefore joined the cluster last. The remaining 15 symptoms were not as closely associated with the eight cluster symptoms or with one another, and did not appear to form any additional clusters. Among these 15 symptoms, it was not surprising that "face rash" and "malar rash" were associated with one another since the two items describe variations of the same symptom. Similarly, association of the symptoms "dry eyes" and "dry mouth" makes sense since these two complaints reflect a decrease in mucosal secretions and are commonly observed together in clinical practice.

Having identified this primary grouping of eight symptoms, how do we know whether the cluster is "real"? There are three main ways of validating the results in order to distinguish whether a structure was "forced" on the data or whether a natural cluster

was "discovered" (100). The first is to divide the data randomly in half, and cluster each half independently. If the clusters are stable then similar results should be obtained for the two halves, as was the case in the present study. The second approach to validation involves the use of several different techniques on the same data set; only results produced by all (or the majority) of methods should be accepted. In this study, four different techniques (average, single, complete and Ward's linkage methods) were used on both halves of the data. All methods consistently discovered the 8 member cluster, irrespective of whether incident or prevalent symptom data was used. The third means of validation involves replicating the solution after deleting variables and, or adding new variables of interest. Analysis of the physician-recorded symptoms involved both adding new symptoms not recorded on the self-administered questionnaire, and removing variables that were not recorded at the time of interview. Furthermore, clarification of symptoms during the interview reduced the possibility of misinterpretation. Despite these changes, a similar cluster is evident in both halves of the data and includes six of the cluster symptoms from the self-reported data (joint swelling, numbness, myalgia, fatigue, insomnia, arthralgia and headaches). The two remaining self-reported cluster symptoms, heartburn and poor memory, were not included in the physician interview. From the new symptoms included in this analysis, only fatigue was added to the cluster. Therefore, three different approaches to validation of the cluster solution produced remarkably similar results suggesting that a useful solution was obtained.

Most researchers agree that the validity of clusters should also be judged qualitatively, by subjective evaluation and interpretability. In most instances, the

researcher has enough content knowledge regarding the problem to distinguish "good" groupings that make sense from "bad" groupings that do not (99). In the context of the current study, the main cluster contains the most frequent symptoms reported by the silicone gel group. The grouping makes sense in that it generally excludes symptoms that are associated with classical CTD and inflammatory disorders, with the exception of self-reported joint swelling. When the latter symptom occurs in the setting of an inflammatory arthritis, it is usually associated with morning stiffness which did not join the cluster consistently in the present analyses. This suggests that the joint swelling was probably mild or that patients may have misinterpreted this symptom. The cluster analysis found that the most common symptoms tended to occur together within individuals. Nevertheless, based on the distance values at which the symptoms joined, the cluster symptoms did not appear to be very highly associated. The analysis did not uncover any previously unnoticed and potentially useful groupings, such that one may question whether the cluster analysis provided any useful information beyond the frequency analysis presented earlier. However, the fact that other clusters of closely associated, perhaps less frequent, symptoms were not identified is also an important finding.

In the context of the existing literature, it was not surprising that the main cluster contained some of the most common symptoms reported by women with SBI including fatigue, cognitive dysfunction, insomnia, myalgia, arthralgia and joint swelling (77). In contrast, however, other symptoms that are associated with "siliconosis" in case series were not found to be associated with the cluster, including lymphadenopathy, Raynaud's

phenomenon, alopecia, and skin rashes. Many case series also suggested that dry eyes and dry mouth were over-represented in women with SBI, suggesting the possibility of a Sjögren's-like illness (104). While dry eyes and dry mouth appeared to be associated with one another in the current analysis, these symptoms were not associated with the main grouping. Thus, many of the symptoms previously ascribed to "siliconosis" did not appear in the primary cluster identified in this analysis.

In summary, the primary eight symptom cluster identified was reproducible and appears to make sense, however, the cluster was not unique. The same cluster was consistently observed in the control and saline breast implant groups irrespective of linkage method, or whether incident, prevalent or physician-recorded symptoms were analyzed. As alluded to previously, a cluster analysis solution is the *beginning* of the research process, not the end (105). At this stage, one can generate hypotheses, however, they must be tested on data generated from new subjects and not the data from which the hypotheses were developed (100). In addition, one can describe the cluster further by evaluating variables that were not used to generate the cluster solution as a form of external validation (102) (Section 4.5).

Simple inspection of the cluster symptoms leads readily to the generation of at least one hypothesis. This grouping of symptoms is very reminiscent of those associated with fibromyalgia. According to the American College of Rheumatology criteria (4), a patient may be classified as having fibromyalgia if they have had widespread pain for at least 3 months and have pain in 11 of 18 tender points on digital palpation. Patients with fibromyalgia also frequently complain of subjective joint swelling (70%), numbness

(63%), sleep disturbance (74%), fatigue (66%), cognitive dysfunction and depression (48%) (81). Symptoms of irritable bowel syndrome, including abdominal pain and constipation alternating with diarrhea, are also common in patients with fibromyalgia (48%). Indeed, the above symptoms, along with chronic headaches, used to be considered minor diagnostic criteria for fibromyalgia before the criteria were revised in 1990 (106).

Chronic fatigue syndrome has many overlapping features with fibromyalgia, and it is recognized that patients often fulfill diagnostic criteria for both disorders (107). While any of the above symptoms can be observed in patients with chronic fatigue syndrome, the diagnosis requires the presence of debilitating fatigue for at least six months, plus 4 out of 8 minor criteria (sore throat, lymphadenopathy, arthralgias, myalgias, new headaches, cognitive dysfunction, nonrestorative sleep, and post-exertional malaise) (108). Interestingly, many of the above symptoms are also described in the "Gulf War Syndrome" (109) and "Multiple Chemical Sensitivities" syndrome (110), neither of which have accepted case definitions at this time. Thus, it is possible that these collections of symptoms may represent a non-specific response to a broad range of stressors (111).

Based on the cluster of symptoms described in the present study, one might hypothesize that patients with SBI were more likely to have fibromyalgia or chronic fatigue syndrome. This would also explain why the primary cluster was not unique to the silicone group, but was also identified in the saline implant and control groups. It is conceivable that SBI may be responsible, at least in part, for causing the cluster

symptoms. For example, women with significant breast pain due to local complications might have chronic disruption of their sleep resulting in fibromyalgia. In addition, women with saline breast implants appear to have fewer local breast complications (112), such that one might expect the frequency of symptoms in this group to be lower. Alternatively, it is also conceivable that some additional feature that predisposes women to fibromyalgia also makes them more likely to have breast implants (58,113). As an example, women who seek perfection in all aspects of life, including their physical appearance, may experience enormous stress levels which may in turn increase their risk of developing fibromyalgia. The latter example does not explain why women with saline breast implants had a lower frequency of symptoms relative to the silicone gel group.

Based on data from case series of symptomatic women with breast implants, other researchers have previously suggested that their symptoms resembled fibromyalgia or chronic fatigue syndrome. In one such case series (N=69), 61% of women met criteria for chronic fatigue syndrome, 49% met criteria for fibromyalgia and 43% met criteria for both conditions (114). Solomon and colleagues reported a high frequency of chronic fatigue syndrome (62%) among 639 women referred for assessment of symptoms but only 5% of these patients were diagnosed with fibromyalgia (115). Several smaller series of symptomatic patients reported frequencies of fibromyalgia ranging from 13% to 98% (116-118). Romano's series of 272 symptomatic breast implant recipients demonstrated that a high proportion had fibromyalgia (86%), myofascial pain syndrome (93%) or both conditions (54%)(119).

Few controlled studies have assessed the possible association between fibromyalgia and SBI. In 1995, Wolfe and colleagues reported a case control study in abstract form comparing the odds of having breast implants among 533 women with fibromyalgia, with the odds of breast implants among 479 women with osteoarthritis plus 655 healthy community controls (58). Exposure status was recorded on a mailed questionnaire or telephone interview. Seven fibromyalgia patients reported breast implants (OR 3.86, 95% CI 1.08, 13.75), however, exclusion of the 3 cases diagnosed with fibromyalgia prior to breast augmentation resulted in a lower risk estimate (OR 2.11, 95% CI 0.51, 8.77).

Recently, two population-based historical cohort studies evaluated the risk of CTD in breast implant recipients from Denmark (67) and Sweden (66). These studies used national registries of hospital discharge ICD-8 codes to ascertain exposure and disease status. The Swedish study did not find an increase in the risk of fibromyalgia among 7442 SBI-exposed compared with 3353 breast reduction controls. However, patients are rarely admitted for fibromyalgia and discharge data for non-rheumatology admissions are likely to miss this diagnosis, such that the prevalence of fibromyalgia was severely underestimated by this study (0.1%). Although the same problem was evident in the Danish study, they found a doubling of risk for "muscular rheumatism" in women with cosmetic breast augmentation (N=1,135) as well as in breast reduction controls (N=7,071) compared to the number expected based on national hospital discharge rates. They suggested that breast surgery *per se* may be associated with an apparent increase in muscular rheumatism, and that this relationship requires further study.

4.5 Health Status of Cosmetic Surgery Patients

Descriptive analyses demonstrated that cluster scores were higher in women with silicone gel breast implants compared to cosmetic surgery controls. Despite this difference, scores for quality of life, global health, and the modified HAQ were similar in the three groups. Quality of life and global health scores were in the range expected for healthy individuals, however, a large number of outliers with lower health status were noted. The similarity between groups is not surprising since the quality of life scale used was not designed for use in the general population and does not discriminate adequately among relatively healthy individuals (92). Similarly, the modified HAQ showed no evidence of disability in any group, with scores (<0.1) similar to those reported for healthy individuals (120). This instrument was designed for use in arthritis patients, but has also been applied in studies of fibromyalgia where average disability scores are typically equal to 1.0 (81,121).

Generic quality of life scales measure the overall well-being of individuals, and complex factors which are unrelated to disease may be incorporated into the final score. The Illness Intrusiveness Rating Scale has therefore been developed as a means of directly assessing the psychosocial impact of illness by measuring illness-induced lifestyle disruptions (122). In this study, patients were specifically asked about the effects of complications due to cosmetic surgery on their lifestyle. Not surprisingly, illness intrusiveness scores were greater in the two breast implant groups compared to controls. Although the silicone gel group had the highest mean illness intrusiveness

scores (18, 95% CI 17.2, 18.5), the data suggest that the degree of intrusiveness was relatively mild overall. For example, this score is considerably lower than that of patients with rheumatoid arthritis, whose mean scores may range from 30 to 50 depending on the severity of the arthritis (122). Again, box plots revealed a large number of outliers, particularly in the silicone gel group, indicating that complications from surgery had affected the lives of a subgroup of patients.

It is reassuring that the health status of the cosmetic surgery patients included in this analysis was normal overall, and that illness intrusiveness was relatively low. Nevertheless, some individuals appear to develop multiple unexplained symptoms and a small group of patients have low health status. The data were therefore explored to determine whether there might be any association between the number of cluster symptoms and health status. Cluster score was moderately associated with reduced quality of life and global health in all three exposure groups ($p<0.0005$). Correlation coefficients ranged from 0.20 to 0.51, and were consistently highest for the silicone gel group and lowest for the control group. While global health had the highest correlation ($r=0.51$), this result suggests that cluster score alone accounted for only 25% of the variability in global health of women with silicone gel breast implants. This is not surprising since a large number of factors are likely to be weighed by the patient when completing this scale. Nevertheless, the size of this coefficient is not trivial relative to analogous correlations reported in the literature. For example, a recent study reported global health and pain scores among patients with rheumatoid arthritis to be highly correlated with a coefficient of 0.66 (123).

Illness intrusiveness scores, which relate more specifically to surgical complications, were also moderately associated with cluster score. It is interesting that the strength of association was highest in the silicone gel group (r=0.43) and the same in the saline and control groups (r=0.27). This implies that women with silicone gel breast implants may be more likely to attribute symptoms to their breast implants or that their symptoms had a greater impact on daily life for reasons that cannot be deduced from this analysis. Local complications, including breast pain, hardening and deformity of the implants, may explain some of the remaining variance in illness intrusiveness scores.

As discussed, the cluster symptoms are relatively nonspecific and resemble those experienced by patients with fibromyalgia. While the symptoms might also be related to an atypical autoimmune disorder, this seems less likely since symptoms such as dry eyes, dry mouth, skin rashes and Raynaud's phenomenon did not form part of the cluster. Further exploratory analysis found that the cluster score was moderately associated with the physician's diagnostic certainty that the patient had fibromyalgia for all exposure groups (r=0.26 to 0.31). Moreover, there was no association between symptoms and per cent certainty of any autoimmune disease (r = -0.004 to 0.132) suggesting that the examining physician did not judge the symptoms to be part of an autoimmune syndrome.

Among the subgroup of subjects that underwent a physical examination, the prevalence of fibromyalgia was 21% in all three exposure groups. It is not surprising that this is higher than the estimated prevalence of fibromyalgia of 3.4% among women in the general population (81), since symptomatic patients were selected for the examination. This value is consistent with the estimated prevalence of 10 to 20% among patients

referred to rheumatology clinics (124). Categorization of subjects based on cluster score revealed that subjects with a score of four or more were 2 to 3 times more likely to have fibromyalgia in all three exposure groups.

While 35% to 40% of women with four or more cluster symptoms were found to have fibromyalgia, at least 60% had symptoms that were "unexplained". This was true for women with breast implants and those without, and likely reflects the high prevalence of musculoskeletal symptoms in the general population (80). For example, a recent population survey in Norway found that 17% of participants (N=11,780) had noninflammatory widespread pain during the previous month, with the highest prevalence occurring among low educated, divorced, or widowed middle-aged and non-working women (125). Wolfe and colleagues performed a population survey in the USA (N=3,006), and found that 10.6% of the population had widespread pain that had persisted for at least 3 months (81). However, only 25% of women and 7% of men with chronic widespread pain had fibromyalgia based on a physical examination that was performed on a random sample (N=193) of these subjects. In agreement with the Norwegian study, they also found that divorced females with lower education levels were at higher risk for fibromyalgia. The prevalence of fibromyalgia increased with age, with the highest values attained between 60 and 79 years (81).

Based on demographic characteristics, one might have expected the control group in this study to have a higher frequency of symptoms and fibromyalgia. Further analysis found no correlation between cluster score and age in any group, and only a weak positive correlation between cluster score and body mass index for the two breast implant

groups. It therefore seems unlikely that demographic differences are responsible for the differences in symptom frequency observed between groups.

In summary, the correlation coefficients described can only provide an indication that cluster score was associated with lower health status and fibromyalgia, but a causal relationship cannot be proven. It is also not possible to conclude that breast implants caused the increase in symptoms. Nevertheless, this exploratory analysis indicates that characteristics related to health status and fibromyalgia differed between subjects with multiple cluster symptoms and those without. This form of external validation lends further support that a useful cluster solution was obtained. While the frequency of symptoms was highest in the silicone gel group, the prevalence of fibromyalgia was similar in all three groups. Alternative explanations for symptoms among women with breast implants need to be considered in future studies.

5.0 STRENGTHS AND LIMITATIONS

5.1 Strengths

Selection bias is defined as the error due to systematic differences between those who participate in a study versus those who do not (96). The potential for selection bias was addressed in the original historical cohort study by analyzing aggregate data on health service utilization in participants and nonparticipants. Both controls and SBI recipients who participated used health care resources to a greater extent than did nonparticipants. This difference may have occurred because people who are familiar with the health care system may be more likely to volunteer for research studies. This finding implies that nonparticipants did not refuse because they were too ill to participate.

Misclassification bias is the error that occurs when an individual is assigned to a category other than that to which they belong (96). This form of bias was minimized by validating exposure status, including implant type, using surgical records. Furthermore, subjects had an adequate exposure period of 8 to 16 years in which symptoms could develop, thus minimizing the potential for underestimating the frequency of symptoms. The design included a blinded history and physical examination which allowed an unbiased assessment of fibromyalgia and atypical autoimmune disease, as well as clarification of symptoms.

Confounding is defined as a situation in which a measure of the effect of an exposure on risk is distorted because of the association of exposure with other factor(s) that influence the outcome under study (96). This study attempted to minimize potential confounding by excluding breast cancer patients. A number of other potential

confounders were also assessed to ensure similarity between the comparison groups. Furthermore, subjects with defined systemic diseases were excluded to allow the study of unexplained symptoms. Additional strengths of the present study included its feasibility, non-intrusiveness and low cost since the database was readily available. Although power calculations are not available for cluster analysis, the number of participants was more than adequate for use of this technique.

5.2 Limitations

Assessment of possible selection bias was a strength of the primary cohort study, and it is unlikely that the risk of CTD was underestimated. However, since health care utilization rates may be a surrogate measure of morbidity, this study might have included more women with symptomatic complaints. Since physician service utilization rates were similar in the SBI and control groups, the frequency of symptoms may have been overestimated to the same extent in both groups. Alternatively, women with SBI may have chosen to participate due to anxiety regarding possible implant-related symptoms, while controls participated irrespective of symptoms. This potential form of selection bias may not have been reflected by physician service utilization rates and would tend to exaggerate differences in symptom frequency between groups. Indeed, the higher rate of participation of women with SBI (17%) relative to controls (10%) suggests that selection bias may have occurred.

Given the attention of the media, as well as potential legal and financial ramifications associated with breast implants during our study, recall bias in the implant

group presents the greatest threat to validity of the symptom data. This would tend to overestimate the risk of individual and multiple symptoms among the exposed. Because of the length of time that has elapsed since cosmetic surgery (7 to 16 years), recall as to whether the symptom started before or after surgery may not be accurate. This is particularly true if the symptom started around the time of the operation. However, if the literature is correct in suggesting that symptom onset occurs an average of 8 to 10 years after breast implantation, then it may be easier for patients to accurately recall whether symptoms started after surgery. This would represent a long interval from the time of surgery and a period of only 0 to 6 years before the study started.

Although potential confounders were explored in the present study, participants were not randomized to the two exposure groups such that they may differ with respect to unmeasured characteristics. For example, a characteristic that makes women choose to have breast implants may also increase symptoms. Thus, differences in symptom frequency between SBI-exposed and control subjects could be due to unmeasured confounders rather than to the implants.

Although the results of the cluster analysis were validated and found to be robust, the results should be considered exploratory and are therefore not applicable to all women with breast implants. Additional research with other subjects is required to determine the extent to which these findings might be generalizable.

6.0 CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

A. *i*) In general, individual symptoms with onset after cosmetic surgery were more frequent among women with silicone gel breast implants compared with cosmetic surgery controls. In addition, women in the silicone gel group were more likely to develop multiple incident symptoms compared with the control group. The saline breast implant group had an intermediate number of symptoms relative to the other two groups; this finding has not previously been reported.

ii) Agreement between physician- and self-reported symptoms was moderate in the two breast implant groups and controls. Women with silicone gel breast implants were more likely than controls to report a symptom as present when the physician did not.

B. *i*) Exploratory cluster analysis of 23 self-reported incident symptoms consistently identified a natural grouping of 8 symptoms in women with silicone gel breast implants, including: poor memory, insomnia, numbness, muscle pain, hand joint pain, joint swelling, headaches and heartburn. Validation procedures, including the use of four different linkage methods and analysis of physician-recorded symptoms, obtained analogous results. The strength of association between these cluster symptoms appeared to be moderate.

ii) The symptom cluster was not unique to the silicone gel group since it was also observed on analysis of data from the saline and control groups. Members of the

silicone gel breast implant group had a greater number of cluster symptoms per subject relative to controls. Again, the saline group had an intermediate number of cluster symptoms per subject relative to the other two groups.

C. *i*) Overall health status among study participants was good. No differences between groups were observed for quality of life, global health or disability scores, although a large number of outliers with reduced health status were noted, particularly among the silicone gel group. Complications due to cosmetic surgery leading to lifestyle disruptions were slightly greater in the two breast implant groups relative to controls. While the overall degree of this intrusiveness was relatively mild, a small subgroup of patients experienced complications that had an important impact on lifestyle.

ii) The number of cluster symptoms per subject was moderately associated with a number of outcomes, including reduced health status, greater illness intrusiveness, and fibromyalgia. This was true for all three exposure groups, however, the strongest associations were observed in the silicone gel breast implant group. This form of external validation of the cluster results provides additional support that a useful solution was obtained.

6.2 Recommendations

The results of the present analysis agree with previous literature suggesting that women with silicone breast implants may be more likely to develop multiple symptoms compared with other women. The symptom cluster identified consisted mainly of nonspecific complaints and did not contain many of the symptoms previously attributed to silicone exposure or autoimmune disease. The differences in symptom frequency observed between groups may either reflect the truth or be due to bias. If the differences are real, it is possible that the symptoms are caused by breast implants. Based on available information, it seems unlikely that breast implants cause these symptoms through autoimmune mechanisms. Alternative explanations, including local breast complications, should be considered in future studies. If local complications are related to the nonspecific cluster symptoms, then this might explain why saline breast implant recipients had fewer symptoms relative to the silicone gel group.

Alternatively, real differences in symptom frequencies may be due to confounding factors. Because the women were not randomized to the three exposure groups, they may differ with respect to unmeasured characteristics that might confound the relationship between SBI and symptoms. Personal characteristics that lead women to choose breast augmentation may be associated with an increase in the risk of nonspecific symptoms. Future studies should broaden our understanding of these personality and related factors, which may in turn improve our ability to counsel and care for these women.

As mentioned, between-group differences in symptoms may be due, at least in part, to bias. Selection bias may have occurred if symptomatic women with breast implants

volunteered to participate in this study, while controls participated irrespective of symptoms. More importantly, recall bias alone could explain the differences in symptom frequency. The observation that women with saline implants reported fewer symptoms than those with silicone gel breast implants could be related to the fact that many women perceive saline implants to be completely safe.

Subjective symptoms represent a difficult health outcome to study since we must accept that symptoms are what the patient says they are, with no means of objectively verifying their presence. Nevertheless, the association between breast implants and symptomatic complaints warrants additional investigation. As discussed, further study is required in order to validate the observed cluster solution in other samples of cosmetic surgery patients. If the same cluster is obtained in other cohorts then this would indicate that the results are generalizable beyond the subjects included in the present study. It seems unlikely, however, that this cluster will form the basis of a new case definition for silicone-related disease. This symptom cluster would not likely be capable of differentiating symptomatic women with breast implants from those with chronic fatigue syndrome or fibromyalgia. Thus, future research efforts should focus on existing constructs related to these fatiguing disorders where a significant body of literature already exists.

Future research could be pursued as an extension of our historical cohort study by conducting a prospective follow-up component. This study would be feasible since a recent survey of the cohort found that the majority of women would be willing to participate in future studies. In addition, sociodemographic data has already been

collected, and implant type has been rigorously validated for these women. Exposed subjects should include women with silicone gel breast implants and saline breast implants. Among the unexposed cohort, those undergoing breast reduction should be analyzed separately to allow evaluation of the effects of breast surgery *per se*.

Blinded interviews and examinations should be performed in a manner similar to the primary cohort study, including documentation of local breast complications. Additional information should be collected regarding possible silicone-related symptoms, including symptom severity, time of onset, and patient certainty regarding onset time. In particular, more detailed information should be obtained on symptoms related to fibromyalgia and chronic fatigue syndrome in order to determine whether subjects meet classification criteria for these conditions. Furthermore, validated scales for measuring pain and fatigue may also provide useful information. As alluded to previously, psychological assessments for personality traits and depression should also be considered in future studies. Finally, valid health status instruments appropriate for relatively healthy individuals should be included in future research efforts. For example, the Fibromyalgia Impact Questionnaire is commonly used in the assessment of fibromyalgia patients (126). In addition, the SF36 has been shown to be useful in assessing functional status in patients with fatiguing illnesses (127). The latter instrument has been used extensively for a broad range of conditions such that comparison of results with those published for other patient groups would be possible. Nevertheless, the ability of the SF36 to discriminate amongst subgroups of relatively healthy individuals is uncertain, such that careful consideration of other health status instruments may also be warranted.

7.0 REFERENCES

1. Sánchez-Guerrero J, Schur P, Sergent JS, Liang MH. Silicone breast implants and rheumatic disease: Clinical, immunologic, and epidemiologic studies. Arthritis Rheum. 1994;37(2):158-68.

2. Silverman S, Borenstein D, Espinoza L, Colin M. Preliminary operational criteria for systemic silicone related disease (SSRD)[Abstract]. Arthritis Rheum. 1996;39(Suppl 9):S51.

3. Fikuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A, and the International Chronic Fatigue Syndrome Study Group. The chronic fatigue syndrome: A comprehensive approach to its definition and study. Ann Intern Med. 1994;121:953-9.

4. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the multicenter criteria committee. Arthritis Rheum. 1990;33:160-72.

5. Cook RR, Harrison MC, LeVier RR. The breast implant controversy. Arthritis Rheum. 1994;37(2):153-7.

6. Silverman BG, Brown SL, Bright RA, Kaczmarek RG, Arrowsmith-Lowe JB, Kessler DA. Reported complications of silicone gel breast implants: An epidemiologic review. Ann Intern Med. 1996;124(8):744-56.

7. Gabriel SE, O'Fallon WM, Beard CM, Kurland LT, Woods JE, Melton LJ. Trends in the utilization of silicone breast implants, 1964-1991, and methodology for a population-based study of outcomes. J Clin Epidemiol. 1995;48:527-37.

8. Dumble LJ. Dismissing the evidence: The medical response to women with silicone implant-related disorders. Health Care for Women International 1996;17:515-25.

9. Bridges AJ. Silicone implant controversy continues [comment]. Lancet 1994;344:1451-2. Comment on: N Engl J Med. 1994;330:1697-1702.

10. Hochberg MC. Silicone breast implants and rheumatic disease [editorial]. Br J Rheumatol. 1994;33(7):601-2.

11. Romano TJ. Breast implants and connective tissue disease [comment]. JAMA 1996;276(2):102.

12. Litvak J. Breast implants and connective tissue disease [comment]. JAMA 1996;276(2):102-3.

13. Hennekens CH, Lee I-M, Cook NR, Hebert PR, Karlson EW, LaMotte F, et al. Breast implants and connective tissue disease [reply]. JAMA 1996;276(2):103.

14. Vasey FB, Aziz N. Breast implants and connective tissue diseases [comment]. N Engl J Med. 1995;333(21):1423. Comment on: N Engl J Med. 1995;332:1666-70.

15. Solomon G, Espinoza L, Silverman S. Breast implants and connective tissue disease [comment]. N Engl J Med. 1994;331(18):1231. Comment on: N Engl J Med 1994 Jun;330(24):1697-702.

16. Liang MH, Karlson EW, Sanchez-Guerrero J. Breast implants and connective tissue diseases [reply]. N Engl J Med. 1995;333(21):1424. Comment on: N Engl J Med 1995;332:1666-70.

17. Angell M. Shattuck Lecture - Evaluating the health risks of breast implants: The interplay of medical science, the law, and public opinion. N Engl J Med. 1996;334(23):1513-8.

18. Salit IE, and the Vancouver Chronic Fatigue Syndrome Consensus Group. The chronic fatigue syndrome: A position paper. J Rheumatol. 1996;23(3):540-4.

19. Rothman KJ. Modern Epidemiology. Toronto: Little Brown & Co; 1986; 2, Causal Inference in Epidemiology. p. 7-21.

20. Hill AB. The environment and disease: Association or causation? Proc R Soc Med. 1965;58:295-300.

21. Edelman DA, Grant S, van Os WAA. Autoimmune disease following the use of silicone gel-filled breast implants: A review of the clinical literature. Semin Arthritis Rheum. 1994;24(3):183-9.

22. Spiera RF, Gibofsky A, Spiera H. Silicone Gel Filled Breast Implants and Connective Tissue Disease: An Overview. J Rheumatol. 1994;21(2):239-45.

23. Silman AJ. Epidemiology of scleroderma. Curr Opin Rheumatol. 1991;3(6):967-72.

24. Rodnan GP, Benedek TG, Medsger TAJ, Cammarata RJ. The association of progressive systemic sclerosis (scleroderma) with coal miners' pneumoconiosis and other forms of silicosis. Ann Intern Med. 1967;66(323):334

25. Veltman G, Lange C, Juhe S, Stein G, Bachner U. Clinical manifestations and course of vinyl chloride disease. Ann N Y Acad Sci. 1975;246:6-17.

26. Alonso-Ruiz A, Zea-Mendoza AC, Salazar-Vallinas JM, Rocamore-Ripoll A. Toxic oil syndrome: A syndrome with features overlapping those of various forms of scleroderma. Arthritis Rheum. 1986;15:200-12.

27. Peters W, Smith D, Fornasier V, Lugowski S, Ibanez D. An outcome analysis of 100 women after explantation of silicone gel breast implants. Ann Plast Surg. 1997;39:9-19. Comment in: Ann Plast Surg 1998;40(1):103-105.

28. Chandler PJ. Re: An outcome analysis of 100 women after explantation of silicone gel breast implants [Comment]. Ann Plast Surg. 1998;40(1):103-5. Comment on: Ann Plast Surg 1997;39:9-19.

29. Peters W. Re: An outcome analysis of 100 women after explantation of silicone gel breast implants [Reply]. Ann Plast Surg. 1998;40(1):105. Comment on: Ann Plast Surg 1997;39:9-19.

30. Wells AF, Daniels S, Gunasekaran S, Wells KE. Local increase in hyaluronic acid and interleukin-2 in the capsules surrounding silicone breast implants. Ann Plast Surg. 1994;33:1-5.

31. Blackburn Jr WD, Grotting JC, Everson MP. Lack of Evidence of Systemic Inflammatory Rheumatic Disorders in Symptomatic Women with Breast Implants. Plast Reconstr Surg. 1997;99(4):1054-60.

32. Young VL, Nemecek JR, Schwartz BD, Phelan DL, Schorr MW. HLA typing in women with breast implants. Plast Reconstr Surg. 1995;96(7):1497-520.

33. Meier LG, Barthel HR, Seidl C. Development of polyarthritis after insertionof silicone breast implants followed by remission after implant removal in 2 HLA-Identical Sisters bearing rheumatoid arthritis susceptibility genes. J Rheumatol. 1997;24(9):1838-41.

34. Hill SL, Landavere MG, Rose NR. The adjuvant effect of silicone gel and silicone elastomer particles in rats. Current Topics Microbiol Immunol. 1996;210:123-37.

35. Naim JO, Ippolito KM, Lanzafame RJ, van Oss CJ. Induction of type II collagen arthritis in the DA rat using silicone gels and oils as adjuvant. J Autoimmun. 1995;8:751-61.

36. Tenenbaum SA, Rice JC, Espinoza LR, Cuéllar ML, Plymale DR, Sander DM, Williamson LL, Haislip AM, Gluck OS, Tesser JRP, et al. Diagnostic and clinical criteria distinguishing silicone related disorders from known rheumatic diseases. Lancet 1997;349:449-54.

37. Marcus DM. An analytical review of silcone immunology. Arthritis Rheum. 1996;39(10):1619-26. Comment in: Arthritis Rheum 1997;40:1730-2.

38. Smalley DL, Shanklin DR. Analytic review of the scientific literature on silicone immune responses: Comment on the article by Marcus [Comment]. Arthritis Rheum. 1997;40:1730-1. Comment on: Arthritis Rheum 1996;39(10):1619-26.

39. Marcus DM. Analytic review of the scientific literature on silicone immune responses: Comment on the article by Marcus [reply]. Arthritis Rheum. 1996;40(10):1731-2.

40. Edlavitch SA. Antipolymer antibodies, silicone breast implants, and fibromyalgia [Comment]. Lancet 1997;349(9059):1170.

41. Lamm SH. Antipolymer antibodies, silicone breast implants, and fibromyalgia [Comment]. Lancet 1997;349(9059):1170-1. Comment on: Lancet 1997;349:449-54.

42. Everson MP, Blackburn WD. Antipolymer antibodies, silicone breast implants, and fibromyalgia [Comment]. Lancet 1997;349(9059):1171. Comment on: Lancet 1997;349:449-54.

43. Korn JH. Antipolymer antibodies, silicone breast implants, and fibromyalgia [Comment]. Lancet 1997;349(9059):1171. Comment on: Lancet 1997;349:449-54.

44. Angell M. Antipolymer antibodies, silicone breast implants, and fibromyalgia [Comment]. Lancet 1997;349(9059):1171-2. Comment on: Lancet 1997;349:449-54.

45. Tenenbaum SA, Rice JC, Espinoza LR, Garry RF. Antipolymer antibodies, silicone breast implants, and fibromyalgia [reply]. Lancet 1997;349(9059):1172-3. Comment on: Lancet 1997;349:449-54.

46. Ellis TM, Hardt NS, Atkinson MA. Antipolymer antibodies, silicone breast implants, and fibromyalgia [Comment]. Lancet 1997;349(9059):1173. Comment on: Lancet 1997;349:449-54.

47. Martin L, Edworthy SM, Barr SG, et al. Autoantibody profiles in patients with breast implants: the Alberta experience. [Abstract] Arthritis Rheum. 1995;38:(Suppl 9)S264.

48. Peters W, Keystone E, Snow K, Rubin L, Smith D. Is there a relationship between autoantibodies and silicone-gel implants? Ann Plast Surg. 1994;32(1):1-5.

49. Goldberg EP, Widenhouse C, Marotta J, Martin P. Failure of silicone gel breast implants: Analysis of literature data for 1652 explanted prostheses [Letter]. Plast Reconstr Surg. 1997;100(1):281-3. Comment in: Plast Reconstr Surg 1998;101(4):1162-4.

50. Cook RR, Hoshaw SJ, Perkins LL. Failure of silicone gel breast implants: Analysis of literature data for 1652 explanted prostheses [Comment]. Plast Reconstr Surg. 1998;101(4):1162. Comment on: Plast Reconstr Surg 1997;100:281-4.

51. Goldberg EP, Widenhouse C, Marotta J, Martin P. Failure of silicone gel breast implants: Analysis of literature data for 1652 explanted prostheses [Reply]. Plast Reconstr Surg. 1998;101(4):1163-4. Comment on: Plast Reconstr Surg 1997;100:281-4.

52. Dugowson C, Daling J, Kowpsell TD, Voigt L, Nelson JL. Silicone breast implants and risk for rheumatoid arthritis. Arthritis Rheum. 1992;35(Suppl 9):S66.

53. Englert HJ, Morris D, March L. Scleroderma and silicone gel breast prostheses - the Sydney study revisited. Aust N Z J Med. 1996;26:349-55.

54. Strom BL, Reidenberg MM, Freundlich B, Schinnar R. Breast silicone implants and risk of systemic lupus erythematosus. J Clin Epidemiol. 1994;47(10):1211-4.

55. Goldman JA, Greenblatt J, Joines R, White L, Aylward B, Lamm SH. Breast implants, rheumatoid arthritis, and connective tissue diseases in a clinical practice. J Clin Epidemiol. 1995;48(4):571-82.

56. Burns CJ, Laing TJ, Gillespie BW, Heeringa SG, Alcser KH, Mayes MD, Wasko MCM, Cooper BC, Garabrant DH, Schottenfeld D. The epidemiology of scleroderma among women: Assessment of risk from exposure to silicone and silica. J Rheumatol. 1996;23(11):1904-11.

57. Hochberg MC, Perlmutter DL, Medsger TA, Nguyen K, Steen V, Weisman MH, White B, Wigley FM. Lack of association between augmentation mammoplasty and systemic sclerosis (scleroderma). Arthritis Rheum. 1996;39(7):1125-31.

58. Wolfe F. Silicone breast implants and the risk of fibromyalgia and rheumatoid arthritis. [Abstract] Arthritis Rheum. 1995;38:(9)S264.

59. Williams HJ, Weisman MH, Berry CC. Breast implants in patients with differentiated and undifferentiated connective tissue disease. Arthritis Rheum. 1997;40(3):437-40.

60. Schusterman MA, Kroll SS, Reece GP, Miller MJ, Ainslie N, Halabi S, Balch CM. Incidence of autoimmune disease in patients after breast reconstruction with silicone gel implants versus autogenous tissue: a preliminary report. Ann Plast Surg. 1993;31(1):1-6.

61. Gabriel SE, O'Fallon WM, Kurland LT, Beard CM, Woods JE, Melton LJ. Risk of connective tissue diseases and other disorders after breast implantation. N Engl J Med. 1994;330:1697-702. Comment in: Lancet 1994 Nov;344:1994; N Engl J Med 1994 Nov;331(18):1231-5.

62. Giltay EJ, Moens HJB, Riley AH, Tan RG. Silicone breast prostheses and rheumatic symptoms: a retrospective followup study. Ann Rheum Dis. 1994;53:194-6.

63. Wells KE, Cruse CW, Baker JL, Daniels SM, Stern RA, Newman C, Seleznick MJ, Vasey FB, Brozena S, Albers SE, et al. The health status of women following cosmetic surgery. Plast Reconstr Surg. 1994;93(5):907-12.

64. Sánchez-Guerrero J, Colditz GA, Karlson EW, Hunter DJ, Speizer FE, Liang MH. Silicone breast implants and the risk of connective tissue diseases and symptoms. N Engl J Med. 1995;332(25):1666-70. Comment in: N Engl J Med 1995 Nov;333(21):1423-1424.

65. Hennekens CH, Lee I-M, Cook NR, Hebert PR, Karlson EW, LaMotte F, Manson JE, Buring J. Self-reported breast implants and connective tissue diseases in female health professionals: A retrospective cohort study. JAMA 1996;275(8):616-21. Comment in: JAMA 1996 Jul;276(2):101-3.

66. Nyrén O, Yin L, Josefsson S, McLaughlin JK, Blot WJ, Engqvist M, Hakelius L, Boice Jr JD, Adami H-O. Risk of connective tissue disease and related disorders among women with breast implants: a nation-wide retrospective cohort study in Sweden. BMJ. 1998;316:417-22.

67. Friis S, Mellemkjær L, McLaughlin JK, Breiting V, Kruger-Kjær S, Blot WJ, Olsen JH. Connective Tissue Disease and Other Rheumatic Conditions Following Breast Implants in Denmark. Ann Plast Surg. 1997;39:1-8.

68. Edworthy SM, Martin L, Barr SG, Birdsell DC, Brant RF, Fritzler MJ. A clinical study of the relationship between silicone breast implants and connective tissue disease. J Rheumatol. 1998;25(2):254-60.

69. Park AJ, Black RJ, Sarhadi NS, Chetty U, Watson ACH. Silicone Gel-Filled Breast Implants and Connective Tissue Diseases. Plast Reconstr Surg. 1998;101(2):261-8.

70. Whorton D, Wong O. Scleroderma and silicone breast implants. West J Med. 1997;167(3):159-65.

71. O'Toole BI, Battistutta D, Long A, Crouch K. A comparison of costs and data quality of three health survey methods: mail, telephone and personal home interview. Am J Epidemiol 1986;124(2):317-28.

72. Rasooly I, Papageorgiou AC, Badley EM. Comparison of clinical and self-reported diagnosis for rheumatology outpatients. Ann Rheum Dis. 1995;54:850-2.

73. Bolumar F, Ruiz MT, Hernandez I, Pascual E. Reliability of the diagnosis of rheumatic conditions at the primary health care level. J Rheumatol. 1994;21(12):2344-8.

74. Colditz GA, Martin P, Stampfer MJ, Willett WC, Sampson L, Rosner B, Hennekens CH, Speizer FE. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. Am J Epidemiol 1986;123(5):894-900.

75. Garbers S, Terry MB, Toniolo P. Accuracy of self-report of breast implants. Plast Reconstr Surg. 1998;101(3):695-8.

76. Rose NR. The silicone breast implant controversy: the other courtroom [editorial]. Arthritis Rheum. 1996;39:1615-8.

77. Borenstein D. Siliconosis: A Spectrum of Illness. Semin Arthritis Rheum. 1994;24 (Suppl 1):1-7.

78. Kovarsky J. Science in the courtroom [comment]. N Engl J Med. 1995;332(19):1099

79. vanVollenhoven R. Breast implants and connective tissue diseases [comment]. N Engl J Med. 1995;333(21):1424

80. Croft P, Rigby AS, Boswell R, Schollum J, Silman A. The prevalence of chronic widespread pain in the general population. J Rheumatol. 1993;20(4):710-3.

81. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum. 1995;38(1):19-28.

82. Meltzer J, Hochstim JR. Reliability and validity of survey data on physical health. Public Health Reports 1970;85(12):1075-86.

83. Pincus T, Summey JA, Soraci Jr SA, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum. 1983;26(11):1346-53.

84. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988;31:315-24.

85. Tan EM, Cohen AS, Fries JF. The 1982 revised criteria for the classification of systemic lupus erythematosus (SLE). Arthritis Rheum. 1982;25:1271-7.

86. Anonymous. Preliminary criteria for the classification of systemic sclerosis (scleroderma): Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum. 1980;23:581-90.

87. International Ethical Guidelines for Ethical Review of Epidemiological Studies. Geneva. CIOMS. 1993.

88. Kuczmarski RJ, Carroll MD, Flegel KM, Troiano RP. Varying body mass index cutoff points to describe overweight prevalence among U.S. adults: NHANES III (1988 to 1994). Obes Res. 1997;5(6):542-8.

89. Statistics Canada. National occupational classification conversion tables: CCDO to NOC, NOC to CCDO. Ottawa: Employment and Immigration Canada. 1993.

90. Cook LS, Daling JR, Voigt LF, deHart MP, Malone KE, Stanford JL, Weiss NS, Brinton LA, Gammon MD, Brogan D. Characteristics of women with and without breast augmentation. JAMA 1997;277(20):1612-7.

91. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. Med Care 1981;19(8):787-805.

92. Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, Shepherd R, Battista RN, Catchlove BR. Measuring the quality of life of cancer patients: A concise QL-Index for use by physicians. J Chron Dis. 1981;34:585-97.

93. Fries JF, Spitz PW, Young DY. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23:137-45.

94. Ramey DR, Raynauld J-P, Fries JF. The Health Assessment Questionnaire 1992: Status and review. Arthritis Care Res. 1992;5(3):119-29.

95. Devins GM. Illness intrusiveness and the psychosocial impact of lifestyle disruptions in chronic life-threatening disease. Adv Ren Replace Ther. 1994;1(3):251-63.

96. Last JM. A Dictionary of Epidemiology. 3rd ed. Toronto: Oxford University Press; 1995.

97. Fisher LD, Van Belle G. Biostatistics: A methodology for the health sciences. 1 ed. Toronto: John Wiley & Sons, Inc.; 1993; 8, Nonparametric, distribution-free and permutation models: Robust procedures. p. 304-44.

98. Altman DG. Practical Statistics for Medical Research. 1st ed. New York: Chapman & Hall; 1991; 14, Some common problems in medical research. p. 396-439.

99. Johnson RA, Wichern DW, editors.Applied Multivariate Statistical Analysis. 1st ed. New Jersey: Prentice-Hall, Inc.; 1982; 11, Clustering. p. 532-78.

100. Everitt, B. Cluster Analysis. Toronto: Heinemann Educational Books Ltd. 1974; p.1-122.

101. Blashfield RK, Aldenderfer MS. The literature on cluster analysis. Multivariate Behavioral Research 1978;13:271-95.

102. Aldenderfer, M.S. and Blashfield, R.K. Cluster Analysis. Newbury Park, CA: SAGE Publications, Inc. 1984; 44. Quantitative Applications in the Social Sciences. p.7-87.

103. Altman DG. Practical Statistics for Medical Research. 1st ed. New York: Chapman & Hall USA; 1991; 11, Relation between two continuous variables. p. 277-324.

104. Freundlich B, Altman C, Sandorfi N, Greenberg M, Tomaszewski J. A profile of symptomatic patients with silicone breast implants: A Sjögren's-like syndrome. Semin Arthritis Rheum. 1994;24(Suppl 1):44-53.

105. Blashfield RK. Propositions Regarding the Use of Cluster Analysis in Clinical Research. J Consult Clin Psych. 1980;48(4):456-9.

106. Goldenberg DL. Fibromyalgia syndrome. JAMA 1987;257:2782-7.

107. Goldenberg DL, Simms RW, Geiger A, Komaroff AL. High frequency of fibromyalgia in patients with chronic fatigue seen in a primary care practice. Arthritis Rheum. 1990;33(3):381-7.

108. Komaroff A, Fagioli LR, Geiger AM, Doolittle TH, Lee J, Kornish RJ, Gleit MA, Guerriero RT. An examination of the working case definition of chronic fatigue syndrome. Am J Med 1996;100:56-64.

109. The Iowa Persian Gulf Study Group. Self-reported illness and health status among Gulf War Veterans: A population-based study. JAMA 1998;277(3):238-45.

110. Fiedler N, Kipen HM, eLuca J, elly-McNeil K, atelson B. A controlled comparison of multiple chemical sensitivities and chronic fatigue syndrome. Psychosomatic Medicine 1996;58:38-49.

111. Ford CV. Somatization and fashionable diagnoses: Illness as a way of life. Scand J Work Environ Health 1997;23(Suppl 3):7-16.

112. Edworthy SM, Martin L, Barr SG, Skrla P. Local connective tissue effects of saline and silicone breast implants: the Alberta experience. Arthritis Rheum. 1995;38(Suppl 9):S325.

113. Barbuto JP. Breast implants and autoimmunity [comment]. West.J.Med. 1994;161(1):90-1. Comment on: West J Med 1994;160:225-228.

114. Chow HY, Calabrese LH, Wilke WS, et al. Is silicone associated illness really chronic fatigue syndrome? [Abstract] Arthritis Rheum. 1995;38:(9Suppl)S264.

115. Solomon G. Clinical and serologic features of 639 symptomatic women with silicone gel implants: Evidence for a novel disease siliconosis. [Abstract] Arthritis Rheum. 1998;37(9):S423.

116. Abeles M, Waterman J. An evaluation of silicone breast implant patients for silicone associated disease. [Abstract] Arthritis Rheum. 1995;38(9):S325.

117. Romano TJ, Stiller JW. The development of fibromyalgia after silicone breast implants: The Korean experience. [Abstract] Arthritis Rheum. 1995;38(9):S325.

118. Silver D, Silverman SL, Mendoza M. Chest wall syndrome in patients with silicone breast implants. [Abstract] Arthritis Rheum. 1994;37(9):S270.

119. Romano TJ. Clinical characteristics of silicone breast implant patients. AJPM. 1996;6(1):13-6.

120. Hubert HB, Fries JF. Predictors of physical disability after age 50. Six-year longitudinal study in a runners club and a university population. Ann Epidemiol. 1994;4(4):285-94.

121. Fitzcharles M-A, Esdaile JM. Nonphysician practitioner treatments and fibromyalgia syndrome. J Rheumatol. 1997;24(5):937-40.

122. Devins GM, Edworthy SM, Guthrie NG, Martin L. Illness intrusiveness in rheumatoid arthritis: Differential impact on depressive symptoms over the adult lifespan. J Rheumatol. 1992;19(5):709-15.

123. Tuttleman M, Pillemer SR, Tilley BC, Fowler SE, Buckley LM, Alarcon GS, Trentham DE, Neuner R, Clegg DO, Leisen JC, et al. A cross sectional assessment of health status instruments in patients with rheumatoid arthritis participating in a clinical trial. Minocycline in Rheumatoid Arthritis Trial Group. J Rheumatol. 1997;24(10):1910-5.

124. Wolfe F. Fibromyalgia: The clinical syndrome. Rheum Dis Clin N Am. 1989;15:1-18.

125. Hagen KB, Kvien TK, Bjorndal A. Musculoskeletal pain and quality of life in patients with noninflammatory joint pain compared to rheumatoid arthritis: A population survey. J Rheumatol. 1997;24:1703-9.

126. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol. 1991;18(5):728-33.

127. Buchwald D, Pearlman T, Umali J, Schmaling K, Katon W. Functional status in patients with chronic fatigue syndrome, other fatiguing illnesses, and healthy individuals. Am J Med 1996;101(4):364-70.

APPENDIX A: Self-Administered Questionnaire (Relevant Sections Only)

Physical Symptoms

In this section we are interested in learning what symptoms, conditions, or side effects you have had.

Please check any symptoms you have experienced and indicate if the symptoms first occurred **BEFORE**, **AFTER**, or **BEFORE AND AFTER** your cosmetic surgery by checking the appropriate box(es) beside each symptom. If you did not experience a symptom, do not check anything.

BEFORE	AFTER	GENERAL
		Enlargement of lymph nodes

BEFORE	AFTER	EYES/NOSE/MOUTH
		Dry eyes
		Dry mouth
		Mouth sores (ulcers)
		Nose sores (ulcers)

BEFORE	AFTER	HAIR/SKIN
		Loss of hair (on head)
		Rash on body
		Rash on face
		Rash over cheeks (butterfly rash)
		Sun sensitivity (unusual skin reaction, not sunburn)
		Red, white and blue skin color change in fingers on exposure to cold or with emotional upset
		Skin ulcers (fingers, toes or legs)

BEFORE	AFTER	STOMACH/BOWELS
		Difficulty swallowing or feeling of food getting stuck
		Heart burn, indigestion, or belching

BEFORE	AFTER	CHEST/LUNGS/HEART
		Chest pain on taking a deep breath

APPENDIX A (continued)

BEFORE	AFTER	MUSCLES/JOINTS
		Muscle pain
		Joint Swelling
		Hand joint pain
		Hand deformities
		If you are stiff in the morning, how long does it last? (hour/min)

BEFORE	AFTER	NERVOUS SYSTEM
		Headaches
		Seizures or convulsions
		Numbness or tingling in hands or feet
		Waking at night (insomnia)
		Trouble thinking and remembering

BEFORE	AFTER	GYNECOLOGICAL HISTORY
		Abnormal vaginal bleeding
		Menstrual problems
		Cessation of menstrual periods (periods stopped)
		Miscarriages

General Opinions on Health and Cosmetic Surgery

1. **Rate your health:**
   Considering all aspects of your **health**, rate how you are doing on the scale below by placing an **X** on the line at the place which you feel best represents your overall health.

```
|------|------|------|------|------|------|------|------|------|------|
VERY                WELL        FAIR        POOR                VERY
WELL                                                            POOR
```

APPENDIX A (continued)

2. Rate your quality of life:

Please mark with an **X** the appropriate place on the line to indicate how you would rate your QUALITY OF LIFE DURING THE PAST WEEK.

LOWEST QUALITY applies to someone completely dependent physically on others, seriously troubled mentally, unaware of surroundings and in a hopeless position.

HIGHEST QUALITY applies to someone physically and mentally independent, communicating well with others, able to do most of the things enjoyed, pulling own weight, with a hopeful yet realistic attitude.

|------|------|------|------|------|------|------|------|------|------|

LOWEST QUALITY — HIGHEST QUALITY

3. Rate your disability:

Please check (✓) the ONE best answer to the questions below:

AT THIS MOMENT, are you able to:	Without **ANY** Difficulty	With **SOME** Difficulty	With **MUCH** Difficulty	**UNABLE** To Do
a. Dress yourself, including tying shoelaces and doing buttons?	_____	_____	_____	_____
b. Get in and out of bed?	_____	_____	_____	_____
c. Lift a ful cup or glass to your mouth?	_____	_____	_____	_____
d. Walk outdoors on flat ground?	_____	_____	_____	_____
e. Wash and dry your entire body?	_____	_____	_____	_____
f. Bend down to pick up clothing from the floor?	_____	_____	_____	_____
g. Turn regular faucets on and off?	_____	_____	_____	_____
h. Get in and out of a car?	_____	_____	_____	_____

APPENDIX A (continued)

Possible intrusiveness of Cosmetic Surgery

The following questions ask if **complications from your cosmetic surgery interfere with different aspects of your life. PLEASE CIRCLE THE NUMBER THAT BEST DESCRIBES YOUR CURRENT LIFE SITUATION. If an item is not applicable,** please circle the number one (1) to indicate that this aspect of your life is not affected. If you have no complications, check the box and go to the next page. ❑

1. How much do complications from your cosmetic surgery interfere with your **HEALTH** (i.e., how well you feel physically)?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

2. How much do complications from your cosmetic surgery interfere with your **DIET** (i.e., the things you eat and drink)?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

3. How much do complications from your cosmetic surgery interfere with your **WORK**?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

4. How much do complications from your cosmetic surgery interfere with your **ACTIVE RECREATION** (i.e., sports)?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

5. How much do complications from your cosmetic surgery interfere with your **PASSIVE RECREATION** (e.g., reading, listening to music)?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

6. How much do complications from your cosmetic surgery interfere with your **FINANCIAL SITUATION**?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

7. How much do complications from your cosmetic surgery interfere with your **RELATIONSHIP WITH YOUR HUSBAND** (or with your boyfriend, if you are not married)?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

APPENDIX A (continued)

8. How much do complications from your cosmetic surgery interfere with your **SEX LIFE**?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

9. How much do complications from your cosmetic surgery interfere with your **FAMILY RELATIONS**?

   NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

10. How much do complications from your cosmetic surgery interfere with your **OTHER SOCIAL RELATIONS**?

    NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

11. How much do complications from your cosmetic surgery interfere with your **SELF-EXPRESSION / SELF-IMPROVEMENT**?

    NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

12. How much do complications from your cosmetic surgery interfere with your **RELIGIOUS EXPRESSION**?

    NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

13. How much do complications from your cosmetic surgery interfere with your **COMMUNITY AND CIVIC INVOLVEMENT**?

    NOT VERY MUCH 1 2 3 4 5 6 7 VERY MUCH

APPENDIX B: PHYSICIAN ASSESSMENT FORM (relevant parts only)

Other History	No	Yes
Weight loss*	❑	❑
Loss of appetite	❑	❑
Fatigue	❑	❑
Sleep disturbances	❑	❑
Headaches	❑	❑
Seizures	❑	❑
Hallucinations	❑	❑
Numbness	❑	❑
Weakness	❑	❑
Dry eyes	❑	❑
Dry mouth	❑	❑
Mouth/nose ulcers	❑	❑
Body rash	❑	❑
Malar rash	❑	❑
Photosensitive rash	❑	❑
Alopecia	❑	❑
Raynaud's syndrome	❑	❑
Skin tightening	❑	❑

If "Yes" ____ #lbs/____ #months

	No	Yes
Joint swelling	❑	❑
Morning stiffness**	❑	❑
Arthralgia	❑	❑
Myalgia	❑	❑
Back pain	❑	❑
Proximal muscle weakness	❑	❑
Cough	❑	❑
Dyspnea	❑	❑
Pleuritic chest pain	❑	❑
Hemoptysis	❑	❑
Edema	❑	❑
Dysphagia	❑	❑
Diarrhea	❑	❑
Constipation	❑	❑
Abdominal pain	❑	❑
Dysuria / Frequency	❑	❑
Nocturia	❑	❑

**If "Yes", duration _____# minutes