



A randomized comparison of group cognitive–behavioral therapy, surface electromyographic biofeedback, and vestibulectomy in the treatment of dyspareunia resulting from vulvar vestibulitis

Sophie Bergeron^{a,*}, Yitzchak M. Binik^b, Samir Khalifé^c, Kelly Pagidas^c,
Howard I. Glazer^d, Marta Meana^e, Rhonda Amsel^a

^aDepartment of Psychology, McGill University, 1205 Dr. Penfield Ave., Montréal, Québec H3A 1B1, Canada

^bDepartments of Psychology, McGill University and Royal Victoria Hospital, 1205 Dr. Penfield Ave., Montréal, Québec H3A 1B1, Canada

^cDepartment of Obstetrics and Gynecology, McGill University and Jewish General Hospital, 687 Pine Ave. West, Montréal, Québec H3A 1A1, Canada

^dDepartment of Psychiatry, Cornell University Medical Center, New York Presbyterian Hospital, 340 East 63rd Street Suite #1A, New York, NY 10021, USA

^eDepartment of Psychology, University of Nevada Las Vegas, 4505 Maryland Parkway, Box 455030, Las Vegas, NV 89154-5030, USA

Received 3 April 2000; received in revised form 3 October 2000; accepted 12 October 2000

Abstract

This study compared group cognitive–behavioral therapy (12-week trial), surface electromyographic biofeedback (12-week trial), and vestibulectomy in the treatment of dyspareunia resulting from vulvar vestibulitis. Subjects were 78 women randomly assigned to one of three treatment conditions and assessed at pretreatment, posttreatment and 6-month follow-up via gynecological examinations, structured interviews and standard questionnaires pertaining to pain (Pain Rating Index and Sensory scale of the McGill Pain Questionnaire, vestibular pain index, pain during intercourse), sexual function (Sexual History Form, frequency of intercourse, Information subscale of the Derogatis Sexual Functioning Inventory), and psychological adjustment (Brief Symptom Inventory). As compared with pretreatment, study completers of all treatment groups reported statistically significant reductions on pain measures at posttreatment and 6-month follow-up, although the vestibulectomy group was significantly more successful than the two other groups. However, the apparent superiority of vestibulectomy needs to be interpreted with caution since seven women who had been assigned to this condition did not go ahead with the intervention. All three groups significantly improved on measures of psychological adjustment and sexual function from pretreatment to 6-month follow-up. Intent-to-treat analysis supported the general pattern of results of analysis by-treatment-received. Findings suggest that women with dyspareunia can benefit from both medical and behavioral interventions. © 2001 International Association for the Study of Pain. Published by Elsevier Science B.V. All rights reserved.

Keywords: Dyspareunia; Vulvodynia; Vulvar vestibulitis; Vestibulectomy; Cognitive–behavioral therapy; Biofeedback

1. Introduction

Chronic or recurrent pain involving the female reproductive system is a neglected, poorly understood, and costly women's health problem (Walling and Reiter, 1995). Dyspareunia, or painful intercourse, a recurrent acute pain which can be located anywhere from the vaginal introitus to the uterus and adnexae, affects 10–15% of women in North America (Laumann et al., 1994). Perhaps the most common type of premenopausal dyspareunia is vulvar vestibulitis

syndrome (Meana et al., 1997), a condition characterized by a sharp, burning pain located within and limited to the vulvar vestibule (vaginal entry) and elicited primarily via pressure applied to the area. This distressing syndrome has no clear etiological determinants, although it has been associated with repeated yeast infections and other urogenital inflammatory conditions (Goetsch, 1991; Bergeron et al., 1997b). Despite a prevalence rate of 15% in general gynecologic practice, there are no randomized trials evaluating treatments for vulvar vestibulitis.

Among the more commonly used therapeutic modalities are cognitive–behavioral therapy, biofeedback and vestibulectomy. Typical cognitive–behavioral interventions aim at reducing the pain and improving sexual function; they include Kegel exercises, vaginal dilatation, and relaxation (Meana and Binik, 1994). In terms of treatment outcome,

* Corresponding author. Present address: Department of Sexology, Université du Québec à Montréal, C.P. 8888, succursale Centre-Ville, Montréal, Québec H3C 3P8, Canada. Tel.: +1-514-987-3000 ext. 3031; fax: +1-514-987-6787.

E-mail address: bergeron.sophie@uqam.ca (S. Bergeron).

there have only been two uncontrolled studies conducted with a population of premenopausal women, and these show that 43–68% of women who undergo individual behavioral therapy treatments benefit from a significant improvement or complete relief of their pain (Abramov et al., 1994; Weijmar Shultz et al., 1996). However, these treatments were either unstandardized or multidisciplinary, which limits the conclusions that can be drawn regarding the efficacy of cognitive-behavioral therapy.

Surface electromyographic (sEMG) biofeedback has been used successfully in the treatment of various chronic pain syndromes (e.g. headaches) and has recently been adapted to dyspareunia. In a study conducted by Glazer et al. (1995), 33 women suffering from different types of vulvar pain underwent sEMG biofeedback training in order to reduce the instability and hypertonicity of their pelvic floor muscles. After an average of 16 weeks of practice, 22 of the 28 women who were abstaining from intercourse at the beginning of the study resumed this activity, and 52% of the women in the entire sample reported pain free intercourse.

Along with attempts to relieve the pain via various topical applications, gynecologists developed a surgery designed to excise the painful tissue in the vulvar vestibule. Although still a controversial procedure, vestibulectomy is becoming a frequently recommended surgical intervention and has been the most investigated treatment for vulvar vestibulitis, with over 20 retrospective studies evaluating its efficacy. Although the success rates of this minor day surgery range from 43–100%, with the majority of estimates surpassing the 60% mark, these conclusions are weakened by multiple methodological flaws (Bergeron et al., 1997b,c). There is thus no firm knowledge as to whether vestibulectomy meets optimal standards of care. Furthermore, its invasiveness may not appeal to many patients. It is conceivable that non-invasive behavioral treatment modalities could provide comparable outcomes without the inherent risks and disadvantages of a surgical procedure. As such, they would represent a much needed alternative to vestibulectomy.

The purpose of the present study was thus to prospectively evaluate and compare the differential efficacy of group cognitive-behavioral therapy (GCBT), sEMG biofeedback, and vestibulectomy in relieving dyspareunia as well as improving sexual function and psychological adjustment.

2. Materials and methods

2.1. Participants

Participants were 87 women suffering from vulvar vestibulitis. They were selected from a pool of 168 women suffering from different types of dyspareunia recruited between January and July 1996 through local media

announcements and professional referral. This study protocol was approved by our institution's ethics review board. Participants met the following inclusion criteria: (1) pain during intercourse which is (a) subjectively distressing, (b) occurs(ed) on most intercourse attempts, and (c) has lasted for at least 6 months; women who stopped attempting intercourse as a result of the pain were included if the pain could be confirmed during the gynecological examinations ($n = 1$); (2) pain limited to intercourse and other activities involving vestibular pressure (e.g. bicycling); (3) moderate to severe pain in one or more locations of the vestibule during the cotton-swab test (cf. Section 2.2); this was operationalized as a minimum average patient pain rating of 4 on a scale of 0–10. Exclusion criteria were the following: (1) pelvic or vulvar pain not clearly linked to intercourse; (2) presence of one of the following (a) major medical and/or psychiatric illness, (b) active infection, and (c) vaginismus; (3) ongoing treatment for dyspareunia; (4) pregnancy; (5) age less than 18 or greater than 50.

2.2. Procedure

On the first visit, each potential participant underwent two independent gynecological evaluations carried out according to the following standardized protocol. (1) A urine sample was obtained. (2) A brief interview about obstetrical/gynecological history, including painful intercourse, was conducted by the gynecologist performing the first examination. (3) Vaginal cultures were taken for *Candida*, *Gardnerella* and *Trichomonas*, as well as a Pap smear if the patient had not been tested in the past year. (4) A cotton-swab palpation of six vestibular sites (in a clockwise fashion: at 12 o'clock, then between 12–3 and 3–6, at 6 o'clock, then between 6–9 and 9–12); this is commonly referred to as the cotton-swab test and constitutes the main diagnostic tool for vulvar vestibulitis. Patients rated the pain at each site on a scale of 0 (no pain) to 10 (worst pain ever). (5) A standard gynecological examination was carried out. In addition, any other physical findings were noted, as were the gynecologists' final diagnoses. A structured interview and standardized questionnaires followed the gynecological examinations.

Potential participants were asked to remain untreated and to discontinue use of potential chemical allergens for a minimum of 6 weeks, at which point they were scheduled for two additional gynecological examinations, identical to the initial ones. The order of gynecologists carrying out the examination at time 1 was reversed at time 2. Based on the above procedure, we found substantial inter-rater reliability ($k = 0.68$) for the diagnosis of vulvar vestibulitis and moderate test-retest reliability ($k = 0.54$) (Bergeron et al., submitted).

Participants who did not meet our selection criteria ($n = 58$) were referred appropriately. Those meeting our criteria were given detailed explanations about the three treatments by the principal investigator. Twenty-three

women met the criteria but declined participation. These 23 women were not significantly different from the women who agreed to be randomized on any of the sociodemographic or pretreatment variables. The remaining 87 participants having provided written consent were randomized to one of the three treatments using blocked randomization. They were required to forgo receiving other interventions for the entire duration of the study. All treatments were free of charge.

2.3. Treatments

The vestibulectomy condition consisted of a minor day surgical procedure of 30 min performed under general anesthesia and involving the excision of the vestibular area to a depth of 2 mm and a width of 1 cm, all the way up to the urethra, with vaginal advancement when necessary. Participants were randomly assigned to one of the two participating gynecologists. Participants met with their gynecologist before surgery to receive information regarding the procedure and 6 weeks postsurgery to receive instructions concerning how to gradually resume intercourse.

Biofeedback participants were randomly assigned to one of two Ph.D. level clinical psychologists trained and supervised by H.I.G. Biofeedback training involved self-insertion of a small single-user sEMG sensor into the vagina. The automated protocol consisted of the following: (1) One 60-s prebaseline rest period; (2) six maximum-intensity rapid contractions or flicks (phasic contractions), each contraction being preceded by a 12-s rest period; (3) six maximum-intensity 12-s contractions (tonic contractions), each contraction being preceded by a 12-s rest period; (4) one maximum intensity 60-s contraction (endurance contraction) preceded by a 30-s rest; (5) one 60-s postbaseline rest period. Participants received eight 45-min sessions over a 12-week period. This treatment also included training in the use of a portable sEMG home trainer for daily practice sessions. Each of the two daily practice sessions consisted of 60 repetitions of a 10-s relaxation period alternated with a 10-s maximum contraction period. The following biofeedback instrumentation was employed: (a) a sEMG single-user vaginal sensor (Model T6050) and a portable sEMG biofeedback instrument (U-Control 60Hz Model T8825) manufactured by Thought Technology Ltd., Montréal, Canada; (b) computerized EMG data acquisition equipment consisting of the FlexiPlus sEMG signal processing hardware and the Glazer Pelvic Floor Muscle Rehabilitation Program, Version 2.2 (Biobehavioral Medical Rehabilitation, Jacksonville, FL) operating on a Pentium 166-Hz laptop computer.

GCBT was delivered by one of two Ph.D. level clinical psychologists in 2-h group sessions with seven to eight women per group. Participants received eight sessions over a 12-week period. They were randomly assigned to either therapist taking into account the language of the group (French or English). Therapists were trained and supervised via a treatment manual designed specifically for this purpose

by the first and second author (Bergeron and Binik, 1998). Adherence to the treatment manual was assessed by two independent clinical associates who viewed and coded a random sample of videotapes representing a quarter of all entire therapy sessions, with an inter-rater reliability of 0.87. Based on this coding of videotapes, therapists adhered to the treatment manual 89.6% of the time. The treatment package included the following: education and information about vulvar vestibulitis and how dyspareunia impacts on desire and arousal; education concerning a multifactorial view of pain; education about sexual anatomy; progressive muscle relaxation; abdominal breathing; Kegel exercises; vaginal dilatation; distraction techniques focusing on sexual imagery; rehearsal of coping self-statements; communication skills training, and cognitive restructuring. Such techniques aimed at reducing the fear of pain during intercourse and other maladaptive affective and cognitive responses, increasing sexual activity level, and reducing pain.

2.4. Dependent measures

With the exception of the vestibular pain index, which was part of the gynecological examinations, the following outcome measures were administered by an independent clinical associate at the first visit of the participant selection process (pretreatment), at posttreatment, and at 6-month follow-up. Data for three of the pain measures (vestibular pain index, the Pain Rating Index and the Sensory scale of the McGill Pain Questionnaire) were also collected at the second selection visit following the 6-week baseline period.

Pain dependent measures included the following. (a) A vestibular pain index, derived from the two independent gynecological examinations conducted at each assessment. The participant pain ratings taken during the cotton-swab test at six different points in the vulvar vestibule were averaged across the two gynecological examinations (per assessment) to form one single index of vestibular pain. Vestibular participant pain ratings have been found to correlate significantly between gynecologists for each palpation site, with correlation coefficients ranging from 0.42 to 0.64, $P < 0.001$ (Bergeron et al., submitted). (b) A self-report measure of the intensity of painful intercourse on a scale of 0–10, taken during the structured interview. (c) The Pain Rating Index (PRI) of the McGill Pain Questionnaire (MPQ) (Melzack, 1975). (d) The Sensory scale of the MPQ. For these last two measures, participants were asked to provide global ratings of the pain they had experienced in the last 3–6 months, depending on the assessment.

Sexual function dependent measures included: (a) The Global Sexual Functioning score of the Sexual History Form (Nowinski and LoPiccolo, 1979), which evaluates desire, arousal, orgasm, frequency of sexual activities, and overall sexual satisfaction and has demonstrated good reliability and validity (Creti et al., 1998); (b) The Sexual Information scale of the Derogatis Sexual Functioning Inventory (Derogatis and Melisaratos, 1979), A reliable and valid

measure of sexual knowledge; (c) A self-report measure of frequency of intercourse per month, taken during the structured interview. Psychological adjustment was assessed using the Global Severity Index of the Brief Symptom Inventory (BSI-GSI) (Derogatis, 1992), a 53-item self-report inventory of psychological symptom patterns.

Treatment credibility was assessed at the first treatment session or during the presurgery appointment using two questions rated on a scale of 0 (not at all) to 10 (completely): (1) 'Up to what point do you think the treatment you are receiving is logical in terms of its efficacy in alleviating vulvar vestibulitis syndrome?' and (2) 'How confident are you that the present treatment will improve your pain condition?' Treatment adherence for sEMG biofeedback and GCBT was measured via frequency ratings of weekly practice of exercises. Participant treatment evaluations involved two questions about subjective improvement (scale of 0 (worse) to 5 (complete cure)) and treatment satisfaction (scale of 0 (completely dissatisfied) to 10 (completely satisfied)). These were part of the posttreatment and 6-month follow-up structured interviews.

2.5. Data analytic strategy

Data were analyzed using a repeated measures multivariate analysis of variance (MANOVA) approach with time as the within-subjects variable and treatment as the between-subjects variable. Outcome measures were clustered per conceptual domain (treatment credibility, pain, sexual function, patient treatment evaluations). When multivariate results were significant, univariate analyses were conducted. If significant, these were followed by planned contrasts or post hoc comparisons with Bonferroni corrections. For variables not significantly correlated with any other (psychological adjustment), a repeated measures ANOVA approach was used. Greenhouse–Geisser adjustment was applied to compensate for violations of homogeneity of covariances. MANCOVA analyses on posttreatment and 6-month follow-up measures using pretreatment measures as covariates were also conducted. We report all the results of the MANOVAs because these analyses involve unadjusted means; results of the MANCOVAs are reported only when they are different from those of the MANOVAs. Results of analysis by treatment actually received are presented first, followed by an intention to treat analysis. Correlational and chi-square analyses were used to investigate the relationship between sociodemographic variables, pretreatment dependent measures and treatment outcome on pain measures at 6-month follow-up in order to determine relevant covariates.

3. Results

3.1. Final sample size

Nine women (seven assigned to vestibulectomy, one

assigned to biofeedback, and one assigned to GCBT) who had agreed to participate dropped out of the study before receiving treatment. They were not different from the women who completed treatment on any of the sociodemographic or pretreatment dependent measures except for the Global Severity Index of the Brief Symptom Inventory ($t = -4.76$, $P < 0.0001$), on which they showed significantly more psychological distress. No further data were collected concerning these women. There were significantly more pretreatment drop-outs in the vestibulectomy condition, $\chi^2(2, n = 87) = 8.92$, $P < 0.01$ than in the two other conditions. This higher level of attrition for the vestibulectomy condition is consistent with clinical practice.

In the biofeedback condition, there were two drop-outs at the posttreatment assessment as well as eight more at the 6-month follow-up. In the vestibulectomy condition, there were three drop-outs at the 6-month follow-up assessment. There were significantly more 6-month follow-up drop-outs in the biofeedback condition, $\chi^2(2, n = 78) = 13.06$, $P < 0.001$. The 13 drop-outs were not significantly different from the women who completed all three assessments on any of the sociodemographic or pretreatment variables. They were included in the analyses by using imputations for missing values (carrying values forward) and reducing the error degrees of freedom by the number of estimated values in order to minimize the risk of Type I error (Nich and Carroll, 1997). The final sample size thus included the 78 participants who actually received treatment. Detailed sociodemographic characteristics of the sample are shown in Table 1. There were no significant differences between treatment groups on any of the sociodemographic or pretreatment variables. In addition, none of the sociodemographic variables were significantly correlated with the pretreatment dependent measures.

Finally, in terms of the use of ongoing treatments during the course of the study, five participants out of 76 (7%) reported having used other means to alleviate their pain at post-treatment, and five participants out of 65 (8%) reported having done so at the 6-month follow-up. Other means included mild remedies such as sitz baths, relaxation, massage, etc. The number of participants having tried to alleviate their pain in other ways did not differ as a function of treatment condition at either posttreatment, $\chi^2(2, n = 76) = 0.74$, $P = 0.69$, or 6-month follow-up, $\chi^2(2, n = 65) = 3.63$, $P = 0.16$.

3.2. Outcome: analysis by-treatment-received

(1) Treatment credibility ratings were high for all three groups, as shown in Table 2. Results from the MANOVA indicated a significant treatment main effect, $F(4, 136) = 4.41$, $P < 0.002$. Univariate analyses demonstrated that GCBT participants' ratings were significantly lower than those of the vestibulectomy participants for logic of treatment, $F(2, 69) = 4.32$, $P < 0.02$, and for confidence in treatment, $F(2, 69) = 9.21$, $P < 0.0001$. Planned

Table 1
Sociodemographic characteristics of the sample^a

Variable	Vestibulectomy	sEMG biofeedback	GCBT	Total
<i>Age (years)</i>				
Mean	26.2	27.0	27.1	26.8
SD	4.8	6.3	5.0	5.4
<i>Pain duration (months)</i>				
Mean	56.4	63.4	52.3	57.4
SD	35.9	65.2	41.0	49.5
<i>Education (years)</i>				
Mean	15.5	16.0	16.3	16.0
SD	3.3	2.0	1.8	2.4
<i>Religion</i>				
Catholic	14	16	17	47
Protestant	1	2	1	4
Jewish	1	0	2	3
Other	0	2	0	2
None	6	8	8	22
<i>Place of birth</i>				
North America	22	25	24	71
Europe	0	3	2	5
Latin/South America	0	0	1	1
Other	0	0	1	1
<i>Marital status</i>				
No partner	5	4	5	14
Dating	5	8	6	19
Living with partner	10	11	12	33
Married	2	5	5	12
<i>Language of interview</i>				
French	18	21	19	58
English	4	7	9	20
<i>Annual income (\$Canadian)</i>				
0–19 999	8	11	15	34
20 000–39 999	6	3	3	12
\$40 000–59 999	4	8	4	16
>\$60 000	4	6	6	16
<i>Ever experienced childbirth</i>				
Yes	1	3	1	5
No	21	25	27	73

^a sEMG, surface electromyographic; GCBT, group cognitive–behavioral therapy.

Table 2
Credibility ratings by treatment condition^a

Variable	Vestibulectomy	sEMG biofeedback	GCBT
<i>Logic</i>			
Mean	8.31	8.14	7.32
SD	1.40	1.30	1.12
<i>Confidence</i>			
Mean	8.31	7.96	6.68
SD	1.49	1.35	1.36

^a sEMG, surface electromyographic; GCBT, group cognitive–behavioral therapy.

comparisons showed that GCBT participants' ratings were also significantly lower than those of the biofeedback participants for confidence in treatment, $F(2, 69) = 12.05$, $P < 0.001$. Correlational analyses with Bonferroni correction revealed that logic of treatment was inversely related to self-reported pain during intercourse at 6-month follow-up, $r = -0.35$, $P < 0.01$. No other relationship was found between treatment credibility and other dependent measures. We nonetheless conducted a separate set of outcome analyses using logic of treatment as a covariate. Results of these analyses are reported only when they differ from the results of the regular set of analyses.

(2) Treatment adherence was defined as complying with at least 70% of the homework exercises. Based on this definition, 65% of GCBT participants complied with treatment, as compared to 57% of biofeedback participants. Chi-square analyses revealed no significant difference in adherence between treatment conditions. For the biofeedback and the GCBT groups either analyzed separately or collapsed

together, there were no significant correlations between degree of adherence to treatment and 6-month follow-up pain measures.

(3) The means and standard deviations for the pain and sexual function measures, the BSI-GSI, and the participant evaluations by treatment and time of assessment are shown in Table 3. There were no significant differences between

Table 3
Dependent measures by time of assessment and treatment condition^a

Measure and group	Pretreatment		Posttreatment		Six-month follow-up	
	Mean	SD	Mean	SD	Mean	SD
<i>Pain</i>						
Vestibular pain index						
Vestibulectomy	6.34	1.85	1.89	1.68	1.90	2.24
sEMG feedback	5.79	1.59	4.55	2.36	4.42	2.63
GCBT	5.45	1.88	5.26	2.00	3.89	2.09
Pain intensity during intercourse						
Vestibulectomy	7.18	1.62	3.93	3.25	3.41	3.17
sEMG feedback	6.93	1.80	5.43	2.36	4.50	2.63
GCBT	7.14	1.53	6.00	2.13	4.46	2.47
MPQ-PRI						
Vestibulectomy	26.82	14.68	15.86	16.18	14.27	13.06
sEMG feedback	26.46	15.99	23.79	17.23	20.43	18.10
GCBT	28.93	12.29	27.75	15.09	20.93	14.18
MPQ – Sensory scale						
Vestibulectomy	17.86	8.40	10.82	9.74	9.45	8.19
sEMG feedback	17.07	8.34	15.57	10.18	13.82	10.66
GCBT	18.61	7.28	18.68	8.69	14.75	8.87
<i>Sexual function</i>						
Sexual history form						
Vestibulectomy	0.47	0.11	0.49	0.14	0.45	0.15
sEMG feedback	0.51	0.11	0.51	0.08	0.48	0.08
GCBT	0.51	0.13	0.49	0.12	0.48	0.11
Frequency of intercourse						
Vestibulectomy	4.61	4.30	1.44	2.85	5.74	5.47
sEMG feedback	3.38	2.91	3.43	3.04	4.04	4.56
GCBT	3.69	3.22	3.25	3.84	3.92	3.77
DSFI, Information subscale						
Vestibulectomy	21.68	1.91	22.41	1.74	22.46	1.90
sEMG feedback	21.46	2.33	22.18	1.61	23.36	1.81
GCBT	21.82	2.31	21.75	2.15	22.25	1.84
<i>Psychological adjustment</i>						
BSI-GSI						
Vestibulectomy	53.32	9.62	52.00	8.25	50.09	10.49
sEMG feedback	54.11	8.78	51.29	8.93	50.79	9.39
GCBT	56.36	8.11	52.89	7.21	51.79	7.61
<i>Participant evaluations</i>						
Satisfaction						
Vestibulectomy	–	–	7.11	2.85	7.73	2.69
sEMG feedback	–	–	6.54	2.55	5.62	3.03
GCBT	–	–	6.91	1.42	7.07	2.09
Improvement						
Vestibulectomy	–	–	3.27	1.49	3.32	1.46
sEMG feedback	–	–	2.46	1.24	2.69	1.46
GCBT	–	–	2.43	1.07	3.00	1.09

^a GCBT, group cognitive-behavioral therapy; sEMG, surface electromyographic; MPQ-PRI, McGill Pain Questionnaire – Pain Rating Index; DSFI, Derogatis Sexual Functioning Inventory; BSI-GSI, Brief symptom Inventory – Global Severity Index.

the pretreatment measures and those taken immediately after the 6-week baseline period for the vestibular pain index, the PRI and the Sensory scale of the MPQ.

(4) Results from the MANOVA conducted on the pain measures indicated a significant time main effect, $F(8, 68) = 18.44$, $P < 0.01$ and a significant time \times treatment interaction, $F(16, 136) = 3.26$, $P < 0.01$. Univariate analyses indicated the following. (a) For the vestibular pain index, there was a significant time main effect, $F(2, 74) = 53.68$, $P < 0.01$, a significant treatment main effect, $F(2, 75) = 6.24$, $P < 0.01$, and a significant time \times treatment interaction effect, $F(4, 148) = 13.24$, $P < 0.01$. Analysis of simple effects and planned comparisons revealed that at posttreatment, vestibulectomy participants had significantly lower pain levels than both GCBT, $F(2, 75) = 17.75$, $P < 0.01$, and biofeedback participants, $F(2, 75) = 20.60$, $P < 0.01$. The same pattern held true at the 6-month follow-up, $F(2, 75) = 7.72$, $P < 0.01$, and $F(2, 75) = 8.99$, $P < 0.01$. Participants from all three treatments improved significantly from pretreatment to 6-month follow-up: vestibulectomy, $F(2, 74) = 59.66$, $P < 0.01$, biofeedback, $F(2, 74) = 6.59$, $P < 0.01$, and GCBT, $F(2, 74) = 10.26$, $P < 0.01$. Planned comparisons showed that both vestibulectomy, $F(2, 74) = 9.86$, $P < 0.01$ and biofeedback, $F(2, 74) = 99.95$, $P < 0.01$ participants significantly improved from pre- to posttreatment. (b) For the self-reported pain intensity during intercourse, there was a significant time main effect, $F(2, 74) = 45.93$, $P < 0.01$. Post hoc analyses revealed that participants significantly improved from pre- to posttreatment, $F(2, 74) = 45.45$, $P < 0.01$, and from posttreatment to 6-month follow-up, $F(2, 74) = 11.71$, $P < 0.01$. MANCOVA analyses yielded the following: a treatment main effect was found, $F(2, 74) = 3.74$, $P < 0.05$, indicating that for posttreatment and 6-month follow-up taken together, vestibulectomy participants were significantly more improved than those from GCBT. Planned comparisons indicated that they were also significantly more improved than those from biofeedback, $F(2, 74) = 5.17$, $P < 0.01$. (c) For the MPQ-PRI, there was a significant time main effect, $F(2, 74) = 12.22$, $P < 0.01$. Post hoc analyses revealed that participants significantly improved from pre- to posttreatment, $F(2, 74) = 6.66$, $P < 0.01$, and from posttreatment to 6-month follow-up, $F(2, 74) = 11.09$, $P < 0.01$. (d) For the Sensory scale of the MPQ, there was a significant time main effect, $F(2, 74) = 9.77$, $P < 0.01$. Post hoc analyses demonstrated

that participants significantly improved from pre- to posttreatment, $F(2, 74) = 6.17$, $P < 0.01$, and from posttreatment to 6-month follow-up, $F(2, 74) = 8.20$, $P < 0.01$. MANCOVA analyses yielded the following: a treatment main effect was found, $F(2, 74) = 3.79$, $P < 0.05$, indicating that for posttreatment and 6-month follow-up taken together, vestibulectomy participants were significantly more improved than those from GCBT. Planned comparisons indicated that they were also significantly more improved than those from biofeedback, $F(2, 74) = 4.23$, $P < 0.05$. Average percentages of pain reduction by dependent measure and treatment condition are shown in Table 4.

(5) Results from the MANOVA conducted on sexual function measures indicated a time main effect, $F(8, 68) = 4.00$, $P < 0.01$, but no interaction effect. Univariate analyses indicated the following. (a) For the Sexual History Form, there was a significant time main effect, $F(2, 74) = 5.60$, $P < 0.01$. Planned comparisons demonstrated that participants significantly improved from posttreatment to 6-month follow-up, $F(2, 74) = 10.53$, $P < 0.01$. The mean for normal women aged 21–46 is 0.46 (Creti et al., 1998). (b) For the Information subscale of the DSFI, there was a significant time main effect, $F(2, 74) = 3.91$, $P < 0.05$. Planned comparisons revealed that participants significantly improved from pre- to posttreatment, $F(2, 74) = 4.63$, $P < 0.05$. However, this effect did not hold when logic of treatment was used as a covariate, $F(2, 68) = 2.70$, $P = 0.07$. Means at all three assessment times are above the mean for female non-patient normals (21.31) (Derogatis and Melisaratos, 1979). (c) For frequency of intercourse, there was a significant time main effect, $F(2, 74) = 10.80$, $P < 0.01$, showing that participants significantly improved from posttreatment to 6-month follow-up. Planned comparisons did not demonstrate a significant difference between pretreatment and 6-month follow-up. When logic of treatment was used as a covariate, results also yielded a significant time \times treatment interaction effect, $F(4, 136) = 3.68$, $P < 0.01$. Analysis of simple effects revealed that vestibulectomy participants significantly improved from posttreatment to 6-month follow-up, $F(2, 68) = 13.25$, $P < 0.01$. Means at all three assessment times are below the mean frequency of intercourse for women aged 25–29 (7.5 times/month) (Laumann et al., 1994). Results from the ANOVA conducted on the BSI-GSI indicated a time main effect, $F(2, 74) = 7.29$, $P < 0.01$. Planned comparisons revealed that participants

Table 4
Average percentages of pain reduction by dependent measure and treatment condition^a

Variable	Vestibulectomy	sEMG biofeedback	GCBT
Vestibular pain index	70.0	23.7	28.6
Pain intensity during intercourse	52.5	35.0	37.5
MPQ-PRI	46.8	22.8	27.7
MPQ – Sensory scale	47.1	19.0	20.7

^a sEMG, surface electromyographic; GCBT, group cognitive-behavioral therapy.

displayed significantly better psychological adjustment at post- as compared to pretreatment, $F(2, 74) = 6.72$, $P < 0.01$. Means at all three assessments are below the clinical cut-off point (T -score = 63) (Derogatis and Melisaratos, 1983).

(6) Results from the MANOVA conducted on participant evaluations indicated a significant time \times treatment interaction, $F(4, 144) = 3.85$, $P < 0.01$. Univariate analyses indicated the following. (a) For subjective improvement, there was a significant time main effect, $F(1, 73) = 4.96$, $P < 0.05$, showing that participants reported a significant improvement from posttreatment to 6-month follow-up. However, this effect did not hold when logic of treatment was used as a covariate, $F(1, 67) = 3.61$, $P = 0.06$. (b) For treatment satisfaction, there was a significant time \times treatment interaction, $F(2, 73) = 4.55$, $P < 0.05$. Analyses of simple effects revealed that biofeedback participants were significantly less satisfied at 6-month follow-up as compared to posttreatment, $F(1, 73) = 6.57$, $P < 0.05$, and that they were significantly less satisfied than vestibulectomy participants at 6-month follow-up, $F(2, 73) = 4.18$, $P < 0.05$. Planned comparisons indicated that they were also significantly less satisfied than GCBT participants at 6-month follow-up, $F(2, 73) = 4.17$, $P < 0.05$.

(7) Treatment success was defined as self-reported great improvement or complete relief of pain on the subjective improvement measure of the participant treatment evaluations (4 or 5 on a scale of 0–5). At 6-month follow-up, 68.2% of vestibulectomy participants, 34.6% of biofeedback participants, and 39.3% of GCBT participants can be said to have a successful outcome. However, 9.1% of vestibulectomy participants ($n = 2$) reported being worse at posttreatment as compared to pretreatment. These two participants' subjective impression was confirmed by all their pain measures except the vestibular pain index, which showed no change.

3.3. Outcome: intent-to-treat analysis

(1) Intent-to-treat analysis was conducted by using imputations for missing values (carrying values forward) of the nine pre-treatment drop-out participants. Results from the MANOVA conducted on the pain measures indicated a significant time main effect, $F(8, 77) = 12.06$, $P < 0.01$ and a significant time \times treatment interaction, $F(16, 154) = 2.10$, $P < 0.01$. Univariate analyses indicated the following. (a) For the vestibular pain index, there was a significant time main effect, $F(2, 83) = 34.66$, $P < 0.01$, a significant treatment main effect, $F(2, 84) = 3.72$, $P < 0.05$, and a significant time \times treatment interaction effect, $F(4, 166) = 8.06$, $P < 0.01$. Analysis of simple effects and planned comparisons revealed that at posttreatment, vestibulectomy participants had significantly lower pain levels than both GCBT, $F(2, 84) = 10.73$, $P < 0.01$, and biofeedback participants, $F(2, 84) = 10.59$, $P < 0.01$. At the 6-month follow-up, vestibulectomy participants had

significantly lower pain levels than biofeedback participants only, $F(2, 84) = 4.10$, $P < 0.05$. Participants from all three treatments improved significantly from pretreatment to 6-month follow-up: vestibulectomy, $F(2, 83) = 36.32$, $P < 0.01$, biofeedback, $F(2, 83) = 5.13$, $P < 0.01$, and GCBT, $F(2, 83) = 10.19$, $P < 0.01$. (b) For the self-reported pain intensity during intercourse, there was a significant time main effect, $F(2, 83) = 37.89$, $P < 0.01$, indicating that participants as a whole improved from pretreatment to 6-month follow-up. (c) For the MPQ-PRI, there was a significant time main effect, $F(2, 83) = 12.55$, $P < 0.01$, indicating that participants as a whole improved from pre-treatment to 6-month follow-up. (d) For the Sensory scale of the MPQ, there was a significant time main effect, $F(2, 83) = 9.59$, $P < 0.01$, indicating that participants as a whole improved from pre-treatment to 6-month follow-up.

(2) Results from the MANOVA conducted on sexual function measures indicated a time main effect, $F(6, 79) = 4.05$, $P < 0.01$, but no interaction effect. Univariate analyses indicated the following. (a) For the Sexual History Form, there was a significant time main effect, $F(2, 83) = 5.19$, $P < 0.01$, indicating that participants improved from pre-treatment to 6-month follow-up. (b) For the Information subscale of the DSFI, there was a significant time main effect, $F(2, 83) = 3.90$, $P < 0.05$, indicating that participants improved from pre-treatment to 6-month follow-up. (c) For frequency of intercourse, there was a significant time main effect, $F(2, 83) = 7.89$, $P < 0.01$, showing that participants improved from post-treatment to 6-month follow-up. Results from the ANOVA conducted on the BSI-GSI revealed a significant time main effect, $F(2, 83) = 7.46$, $P < 0.01$, indicating that participants displayed better psychological adjustment at 6-month follow-up as compared to pretreatment.

4. Discussion

Four main conclusions can be drawn from the results of this study: (1) there are potentially efficacious medical and psychosocial treatments for vulvar vestibulitis; (2) based on results of the intent-to-treat analysis for the vestibular pain index measure, vestibulectomy is significantly more successful than sEMG biofeedback; (3) the three treatments provide equally positive sexual function and psychological adjustment outcomes; (4) gains are maintained at 6-month follow-up for participants in all treatment conditions.

However, these results need to be interpreted with caution since there were significantly more participants in the vestibulectomy condition who refused to undergo the treatment they had been randomized to, as compared to participants in the two other treatment conditions. Such differential failure to accept randomization clouds interpretation of the data. Indeed, the final sample of women who received vestibulectomy may have been biased in favor of this procedure

and thus may not be representative of the population of women with vulvar vestibulitis encountered in clinical settings. We have attempted to correct this problem by conducting an a posteriori intent-to-treat analysis which included the vestibulectomy pretreatment drop-outs; this analysis confirmed the general pattern of results found with analysis by-treatment-received. One of the functions of an intent-to-treat strategy is to preserve the comparability of groups allowed by randomization (Newell, 1992).

Differential treatment credibility constitutes yet another factor which complicates interpretation of the findings. Considering that participants were significantly less confident in the efficacy of GCBT as compared to vestibulectomy and biofeedback, one can question whether this study represents a fair comparison of treatments, especially when evaluating interventions as disparate as CBT and surgery. We have tried to limit the impact of this problem by conducting additional analyses using the two treatment credibility items as covariates. Results of these analyses cancelled the effects of two variables, namely sexual knowledge (the Information subscale of the DSFI) and subjective improvement from posttreatment to 6-month follow-up. Future research comparing interventions with different underlying mechanisms will need to examine more thoroughly the issue of treatment credibility.

Findings of the present study cannot be easily accounted for by a placebo or attention effect for the following reasons: (1) important differential treatment effects were found; (2) there were no significant changes in pain during the 6-week baseline period, despite multiple gynecological examinations, an extensive psychosexual evaluation, and the expectation of entering a treatment study; (3) GCBT participants did not experience a significant change in pain on the vestibular pain index from pre- to posttreatment, even though this was the treatment condition in which the women received the most clinical attention. It is impossible, however, to totally discount the possibility of a placebo effect since credible placebo controls in non-pharmacological studies are difficult if not almost impossible to design (Turner et al., 1994; Schwartz et al., 1997). In addition, findings do not appear to be a reflection of the mere passage of time since the mean duration of vulvar vestibulitis prior to study entry was close to 5 years and 39% of women reported trying a number of different medical treatments before entering the study. Nonetheless, only a wait-list control group equal in duration to the length of treatment and follow-up could allow us to discount a passage of time interpretation. However, such a design often increases drop-out rates considerably (e.g. Peters and Large, 1990).

Results of this study show that vestibulectomy participants had significantly lower levels of pain on the vestibular pain index than participants from the GCBT and biofeedback conditions at posttreatment, and than participants from the biofeedback condition at 6-month follow-up. When posttreatment and 6-month follow-up assessments were taken together in the analysis by-treatment-received, vesti-

bulectomy participants also had significantly lower levels of pain during intercourse and lower MPQ Sensory scores. Moreover, vestibulectomy is characterized by a high success rate and by elevated percentages of pain reduction (from 46.8% to 70.0%). This successful outcome needs nonetheless to be considered within a larger perspective. Two out of 22 surgery participants reported being worse after the intervention, and seven women who had initially been randomized to vestibulectomy refused to go ahead with the treatment. This high pretreatment drop-out rate suggests that a significant percentage of women are reticent to undergo such an invasive procedure. Moreover, adequate evaluation of vestibulectomy was limited by the absence of systematic follow-up of pretreatment drop-outs and thus an intent-to-treat analysis conducted on estimated rather than collected data.

Average decreases in pain for GCBT participants are encouraging in light of findings from a meta-analysis conducted by Flor et al. (1992), showing that the average reduction in pain intensity for multidisciplinary pain clinic patients across 65 studies was 37%. In addition, no participants dropped-out of GCBT once the treatment had begun. GCBT may reduce anxiety by giving participants more control over their pain and by changing the meaning of the situation for them, thereby affecting cognitive and emotional factors. The group aspect of the treatment may produce change in social expectations by normalizing dyspareunia for participants. The significantly lower initial treatment credibility of GCBT may have affected outcome; the negative correlation between the logic of treatment and self-reported pain during intercourse at 6-month follow-up supports this hypothesis. Such findings highlight the need to improve the presentation of psychological pain treatments to patients (Turk and Rudy, 1990).

Success rate as well as decreases in pain for biofeedback participants are somewhat lower than those reported in the first published retrospective study of biofeedback (Glazer et al., 1995). This might be linked to the way the treatment was delivered in this study as opposed to the Glazer et al. (1995) study, in which participants were (1) thought to be 100% compliant with homework exercises, (2) suffering from various types of vulvar pain as opposed to only vulvar vestibulitis, (3) receiving other concomitant treatments. Furthermore, the two therapists delivering the treatment in the present study were less experienced than the therapist in the original Glazer et al. (1995) study. The significantly higher 6-month follow-up drop-out rate and the significantly lower satisfaction rate in the biofeedback show that a substantial number of participants experienced difficulties in following through with this intervention. This appears to be independent of the outcome because participants in the GCBT did not drop-out and had similar outcomes. It is possible that the large time investment and the repetitive exercises may be responsible for this negative effect.

All three treatments have an equally positive effect on sexual functioning and psychological adjustment. The find-

ing that women assigned to GCBT do not improve their sexual functioning significantly more than the others remains puzzling. This may be due to the fact that improved sexual function is dependent on the degree of pain that one experiences. However, the results also show that a significant reduction in pain, as found in the vestibulectomy condition, does not necessarily bring about increased frequency of intercourse or better overall sexual functioning. These conflicting results suggest that multimodal treatment approaches may be essential to achieve significant improvement in all aspects of the disorder (Bergeron et al., 1997a).

Indeed, medical and psychosocial treatments are not mutually exclusive and can be combined in an effort to provide women suffering from dyspareunia with the best possible outcomes. Within a multidisciplinary framework, GCBT and sEMG biofeedback represent promising alternatives to vestibulectomy because they do not involve significant physical risks. Future studies should consider combining behavioral interventions to evaluate whether their effects are additive and can equal those of vestibulectomy.

Acknowledgements

This article stands in partial fulfillment of S.B.'s Ph.D. requirements. The initial sample of women from which we recruited participants for this study was also used to conduct a reliability study on the diagnosis of vulvar vestibulitis syndrome (manuscript submitted for publication). This research was supported by a Social Sciences and Humanities Research Council of Canada Fellowship to S.B. and by a Health Canada (NHRDP 6605-4381-503) research grant to Y.M.B. We thank Maria Amore, Geneviève Bédard, Janet Bradley, Patricia Costantino, Dina Giannopoulos, Louise Labelle, Terry Peled, Caroline Pukall, Dionne Rodrigues, and Vivienne Zhao for their help in conducting this study.

References

- Abramov L, Wolman I, David MP. Vaginismus: An important factor in the evaluation and management of vulvar vestibulitis syndrome. *Gynecol Obstet Invest* 1994;38:194–197.
- Bergeron S, Binik YM. Treatment manual for cognitive-behavioral group therapy with women suffering from vulvar vestibulitis syndrome. Unpublished treatment manual. Montreal, Canada: McGill University, 1998.
- Bergeron S, Binik YM, Khalifé S, Meana M, Berkley KJ, Pagidas K. The treatment of vulvar vestibulitis syndrome: towards a multimodal approach. *Sex Marital Ther* 1997a;12:305–311.
- Bergeron S, Binik YM, Khalifé S, Pagidas K. Vulvar vestibulitis syndrome: a critical review. *Clin J Pain* 1997b;13:27–42.
- Bergeron S, Binik YM, Khalifé S, Pagidas K, Glazer HI. Vulvar vestibulitis syndrome: reliability of diagnosis and validity of current diagnostic criteria. Montreal: McGill University, Manuscript submitted for publication.
- Bergeron S, Bouchard C, Fortier M, Binik YM, Khalifé S. The surgical treatment of vulvar vestibulitis syndrome: a follow-up study. *J Sex Marital Ther* 1997c;23:317–325.
- Creti L, Fichten CS, Amsel R, Brender W, Schover LR, Kalogeropoulos D, Libman E. Global sexual functioning: A single summary score for Nowinski and LoPiccolo's Sexual History Form (SHF). In: Davis CM, Yarber WL, Bauserman R, Schreer G, Davis SL, editors. *Handbook of sexuality-related measures*, Thousand Oaks, CA: Sage Publications, 1998, pp. 261–267.
- Derogatis LR. BSI: administration, scoring and procedures manual, II. Baltimore, MD: Clinical Psychometric Research, 1992.
- Derogatis LR, Melisaratos N. The DSFI: A multidimensional measure of sexual functioning. *J Sex Marital Ther* 1979;5:244–281.
- Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med* 1983;13:595–605.
- Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain* 1992;49:221–230.
- Glazer HI, Rodke G, Swencionis C, Hertz R, Young AW. The treatment of vulvar vestibulitis syndrome by electromyographic biofeedback of pelvic floor musculature. *J Reprod Med* 1995;40:283–290.
- Goetsch MF. Vulvar vestibulitis: Prevalence and historic features in a general gynecologic practice population. *Am J Obstet Gynecol* 1991;161:1609–1617.
- Laumann EO, Gagnon JH, Michael RT, Michaels S. *The social organization of sexuality*, Chicago, IL: University of Chicago Press, 1994.
- Meana M, Binik YM. Painful coitus: a review of female dyspareunia. *J Nerv Ment Dis* 1994;182:264–272.
- Meana M, Binik YM, Khalifé S, Cohen D. Biopsychosocial profile of women with dyspareunia. *Obstet Gynecol* 1997;90:583–589.
- Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277–299.
- Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol* 1992;21:837–841.
- Nich C, Carroll K. Now you see it, now you don't: a comparison of traditional versus random-effects regression models in the analysis of longitudinal follow-up data from a clinical trial. *J Consult Clin Psychol* 1997;65:252–261.
- Nowinski JK, LoPiccolo J. Assessing sexual behaviors in couples. *J Sex Marital Ther* 1979;5:225–243.
- Peters JL, Large RG. A randomised control trial evaluating in- and outpatient pain management programmes. *Pain* 1990;41:283–293.
- Schwartz CE, Chesney MA, Irvine MJ, Keefe FJ. The control group dilemma in clinical research: applications for psychosocial and behavioral medicine trials. *Psychosom Med* 1997;59:362–371.
- Turk DC, Rudy TE. Neglected factors in chronic pain treatment outcome studies – referral patterns, failure to enter treatment, and attrition. *Pain* 1990;43:7–25.
- Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. *J Am Med Assoc* 1994;271:1609–1614.
- Walling MK, Reiter RC. Chronic pelvic pain. In: O'Hara MW, Reiter RC, Johnson SR, Milburn A, Engeldinger J, editors. *Psychological Aspects of women's reproductive health*, New York: Springer, 1995, pp. 65–80.
- Weijmar Shultz WCM, Gianotten WL, van der Meijden WI, van de Miel HBM, Blindeman B, Chadha S, Drogendijk AC. Behavioural approach with or without surgical intervention for the vulvar vestibulitis syndrome: a prospective randomized and non-randomized study. *J Psychosom Obstet Gynecol* 1996;17:143–148.