THE JOURNAL OF
SEXUAL MEDICINE

PAIN

ORIGINAL RESEARCH

What Do Different Measures of Pain Tell Us? A Comparison in Sexually Active Women With Provoked Vestibulodynia

Frederikke Louise Wammen Rathenborg, MSc,[1] Bozena Zdaniuk, PhD,[2] and Lori A. Brotto, PhD[2]

ABSTRACT

Introduction: Studies of pain measurement in women with provoked vestibulodynia (PVD) use various methods of capturing pain intensity. The degree to which these different measures of pain correspond with one another is not known.

Aim: To compare 3 different measures of pain intensity in sexually active women with PVD participating in a clinical treatment study.

Methods: A total of 64 women (mean age 30.9 years) provided baseline measures of pain intensity using (i) a numeric rating scale that provided a self-report of pain during recalled vaginal penetration; (ii) the pain subscale of the female sexual function index; and (iii) pain elicited with a vulvalgesiometer, an objective method of eliciting pain.

Main Outcome Measure: Correlations among these 3 measures of pain were moderate in size (range $r = 0.39-0.61$). Moreover, the numeric rating scale of pain was more likely to be associated with self-reported measures of pain catastrophizing and pain hypervigilance than were scores on the pain subscale of the female sexual function index or scores from the vulvalgesiometer.

Clinical Implications: Overall, there was a moderate level of correlation between different often-used measures of pain in women with PVD. These findings suggest that, in addition to measuring a common dimension, these different measures tap into different aspects of women's experiences with vulvovaginal pain, and researchers should consider how the chosen measure addresses their primary research question when selecting pain measures in future PVD research.

Strengths & Limitations: A strength of this study was the large sample size (n = 64 sexually active women) who had received confirmed clinical diagnoses of PVD. 1 limitation of the findings is that our self-report outcome measures are based on retrospective ratings of pain over 4 weeks, and it is possible that other variables, such as mood, could have impacted scores on these measures.

Conclusion: This study showed statistically significant and moderate correlations among 3 different pain measures widely used in PVD research and treatment. In addition, only 1 pain measure showed a significant independent association with emotion function measures. These findings provide a rationale for including multiple measures of pain and emotional function in treatment outcome studies of PVD. **Wammen Rathenborg FL, Zdaniuk B, Brotto LA. What Do Different Measures of Pain Tell Us? A Comparison in Sexually Active Women With Provoked Vestibulodynia. J Sex Med 2019;16:278–288.**

Key Words: Pain Measurement; Genital Pain; Dyspareunia; Provoked Vestibulodynia; Vulvalgesiometer; Numeric Rating Scale

INTRODUCTION

Provoked vestibulodynia (PVD) is a chronic vulvar pain disorder affecting 8%–16% of women across ages.[1,2] The International Society for the Study of Vulvovaginal Disease characterizes PVD as sharp, stabbing, searing pain with contact to the vulva or vagina lasting ≥3 months that is not attributable to other physical causes for the pain, such as vulvar dermatoses, vulvar trauma, or skin

Received July 27, 2018. Accepted December 3, 2018.

[1]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark;

[2]Department of Obstetrics and Gynaecology, University of British Columbia, Vancouver, British Columbia, Canada

https://doi.org/10.1016/j.jsxm.2018.12.001

infections.[3] There has been considerable interest in the pathophysiology and treatment of PVD over the past decades, with dozens of empirical investigations[4,5] into the course of PVD symptoms over time, with and without treatment.

The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT)[6] is dedicated to investigating the best ways to measure chronic pain and this entity recommends using ≥2 different pain measures to assess changes in pain among those with chronic pain disorders. In addition to these measures of pain sensation and affect, IMMPACT recommends that emotional function, physical function, adverse events, overall treatment satisfaction, and participant disposition should be measured. In using the IMMPACT guidelines specifically through the lens of PVD, Pukall et al[7] recommended that measures of pain intensity (eg, numeric rating scale; NRS), pain temporality (eg, measures of spontaneous and unprovoked pain), and emotional function (eg, depression, pain catastrophizing) be standard in outcome studies.

Because PVD is considered a chronic pain disorder[8], the IMMPACT guidelines should be considered when designing clinical trials on PVD. However, studies of PVD have not systematically followed these IMMPACT guidelines nor used the same pain measures consistently, which makes cross-study comparisons difficult. The various measures used to quantify the pain of PVD have included both retrospective self-report scales, such as the NRS[9] and the pain subscale of the Female Sexual Function Index (FSFI),[10] and measures of pain in response to objective pain elicitation.[11] Among the objective measures of vulvar pain assessment, one innovative measure is the vulvalgesiometer, which consists of spring-loaded syringes for application of fixed amounts of pressure to the vulvar region whereas women are asked to rate the pain felt during palpation.[11–13]

It is unknown which measure of vulvovaginal pain best describes the pain that women experience in PVD. And whereas past studies have tended to examine the association between measures of pain and measures of psychological function,[14] sexual function,[15] and disease severity,[16] very few have examined how well these different measures of pain correlate (or correspond) with one another. One exception is a study by Pukall et al[12] that sought to validate the vulvalgesiometer in 15 women with and 15 women without PVD who underwent pressure pain threshold testing in the posterior portion of the vestibule during the late follicular phase of their cycles. Different amounts of pressure were applied to the vestibules while participants self-reported pressure or pain thresholds. They found a significant negative correlation between pressure pain thresholds (using the vulvalgesiometer) and pain rating during the cotton swab test ($r = -0.73$, $P < .01$). In contrast, another study found no correlation between the recalled intercourse-related pain on the McGill Pain Questionnaire and pain ratings during the cotton swab test.[13]

To date, we are aware of only 1 study that has tested the association between several different vulvovaginal pain measures in a sample of 98 women diagnosed with PVD.[15] Pain measures included intercourse-related pain, pressure pain threshold using the vulvalgesiometer, and the total FSFI score. Although the correlation between intercourse-related pain measured with the NRS and the total FSFI score was small ($r = -0.06$) and nonsignificant, intercourse-related pain was significantly related to objective measures of pain using a cotton swab test ($r = 0.23$) and using vestibular friction with a cotton swab rubbing back and forth over the vulva ($r = 0.23$). Of note, these correlations were small, indicating that the self-report of pain with intercourse and objective measures of pain intensity tap into different aspects of women's pain experiences.

AIM

The aim of this study was to examine how 3 different pain measures, routinely used in samples of women with PVD correspond with one another—2 were retrospective self-reported measures: (i) intercourse-related pain using the numerical rating scale, (ii) the pain subscale of the FSFI, and (iii) responsive quantitative self-reported pain testing using the vulvalgesiometer. This was intended to test whether the findings of Aerts et al[15] could be replicated in a different sample. Based on the findings by Aerts et al,[15] we hypothesized that there would be significant but small to modest correlations among these 3 measures of pain. Second, we aimed to examine the association between measures of pain and measures of emotional function, including measures of chronic pain acceptance, pain catastrophizing, pain vigilance, and anxiety. The secondary goal stems from recommendations made by IMMPACT, which indicates that, in addition to measures of pain, clinical trials should also include measures of emotional function.[6] As indicated by Pukall et al,[7] who adapted these IMMPACT guidelines to the study of PVD, the extent to which pain measures and emotional function measures overlap in PVD is unknown. We were specifically examining each measure's unique association with emotional function above and beyond the other 2 pain measures.

MATERIALS AND METHODS

Participants

Data for this study were obtained from the baseline assessment of participants in a recently completed clinical trial of psychological treatments for women with PVD (Lori Brotto, unpublished data, 2018). Participants (n = 130) met the following inclusion criteria: (i) a diagnosis of PVD that was confirmed by both clinical history and a cotton-swab test carried out by a physician with expertise in sexual medicine and vulvovaginal conditions; (ii) a duration of PVD ≥6 months; (iii) ability to attend 8 weekly treatment sessions; (iv) age 19 years or older and premenopausal; (v) fluent in English; and (vi) a willingness not to begin any new treatments for PVD for the duration of the study until the 6-month follow-up point. Exclusion criteria were (i) generalized vulvodynia; (ii) deep pelvic pain; (iii) a vulvar skin condition (eg,

lichen sclerosis); and (iv) significant symptoms of dissociation (the study comprised of group psychological skills and participant dissociation would have made participation challenging). The latter was assessed using items 90–104 from the Borderline Personality Disorder of the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders.[17]

Procedure

All participants provided written consent, and all procedures were approved by the Clinical Research Ethics Board at the University of British Columbia and the hospital research ethics board of Vancouver Coastal Health Research Institute.

Each woman took part in a physical examination and objective assessment of vulvovaginal pain in a private office by a physician investigator on the study. With the woman lying on an examination table with her legs supported by stirrups, the physician applied a pressure of 30 g using the vulvalgesiometer at the 1, 3, 4, 6, 8, 9, and 11 o'clock positions (in random order) around the vestibule and recorded the woman's rating of pain using an NRS.[12] An average pain rating across these 7 sites was then computed for analyses. The order of touch was randomized to reduce the occurrence of sensitization if different areas around the vulva were sequentially touched.[11] All physicians were trained in the use of the vulvalgesiometer before the study began.

After the physical examination, participants were emailed a link to an online questionnaire to be filled out at home at their convenience, which included retrospective pain measures. Women were asked to recall and rate their average pain during intercourse in the past 4 weeks on a classic numerical rating scale[9,18–20] and the pain subscale of the FSFI.[10] The NRS is the pain measure most often used in a clinical setting to evaluate both acute and chronic pain,[21] and the validity of both these self-report measures of pain has been independently established.

Also, the women completed a select number of psychological questionnaires—among these the Chronic Pain Acceptance Questionnaire,[22] the Pain Catastrophizing Scale (PCS),[23] the Pain Vigilance and Awareness Questionnaire,[24] and the Anxiety Sensitivity Index (ASI).[25] After the NRS, all questionnaires were given in random order.

MEASUREMENTS

Objective Pain Measurement Using the Vulvalgesiometer

The vulvalgesiometer makes quantitative sensory testing more standardized by allowing different clinicians to apply a fixed amount of pressure to the vulvar area. Unlike the commonly used cotton swab test, which is a standard clinical tool for making the diagnosis of PVD, the vulvalgesiometer is largely reserved for research and serves as a validated and controlled measure for determining pressure pain thresholds in women with PVD. The vulvalgesiometer consists of a set of 5 syringes containing spring loads calibrated to exert precise pressure between 3 and 1000 g.[12] At the end of each syringe is a disposable cotton swab tip applying the actual touch.

For this study, only 1 syringe calibrated to apply a total pressure of 30 g was used. This was based on the recommendation of the developer of the vulvalgesiometer (Pukall, personal communication, 2012), because women were self-reporting pain of 5 of 10 on the NRS with a light touch during the cotton swab test, so we did not want to exert pressure that induced too much pain. The vulvalgesiometer has been shown to be able to discriminate accurately between women with PVD and healthy control subjects and has also been shown to have excellent interrater reliability.[12,13,26] In addition to this, it has been shown to mimic better the pain felt during intercourse than the original cotton swab test.[26]

Self-Report Pain Measurements

NRS

The NRS is a patient-reported outcome measure often used in assessing pain in clinical settings and specially to assess the potential change in pain.[9] It asked participants to rate the "intensity of pain during vaginal penetration attempts with sexual intercourse or penetration over the past 4 weeks." This question was rated on an 11-point Likert scale ranging from 0 (no pain at all) to 10 (worst possible pain). Women who did not engage in vaginal penetration over the previous 4 weeks received a not applicable score for this question, and their data were not analyzed.

FSFI, Pain Subscale

The FSFI includes 19 questions distributed across 6 subscales: desire, arousal, lubrication, orgasm, satisfaction, and pain; only the latter was analyzed in this study. The pain subscale is comprised of 3 items: (i) "Over the past 4 weeks, how often did you experience discomfort or pain *during* vaginal penetration?"; (ii) "Over the past 4 weeks, how often did you experience discomfort or pain *following* vaginal penetration?," both of which have response options ranging from almost always/always to almost never/never. Participants were instructed to take an average of their pain levels over the previous 4 weeks. Item 3 was (iii) "Over the past 4 weeks, how would you rate your level (degree) of discomfort or pain during or following vaginal penetration?," with possible answers ranging from very high to none at all. For all 3 items, the answer "did not attempt intercourse" was available. Scores range between 0–6 on the pain subscale, with a score of 0 indicating that intercourse was not attempted in the previous 4 weeks, whereas a score of 6 indicated no pain or discomfort. Lower scores indicate higher levels of pain and discomfort. We omitted from analysis any woman who indicated that she did not engage in sexual activity over the previous 4 weeks. The scale showed satisfactory reliability with standardized Cronbach's alpha = 0.78 (standardized alpha is reported due to the items having different answer options).

Emotional Function Measurements

4 different domains of emotional and psychological function were measured. We administered the (i) the Chronic Pain Acceptance Questionnaire,[22] which measures the extent to which one tries to avoid or control pain and the extent to which one participates in valued activities despite living with pain. 2 subscales of the questionnaire were used in this study, the Activities Engagement subscale, which measures pursuit of life activities regardless of pain, and the Pain Willingness subscale, which recognizes that avoidance and control are not adaptive means of coping with pain. The 11 items of the Activities Engagement subscale had a total score range from 0−66, and the 9 items of the Pain Willingness subscale had a total score range from 0−54, with higher scores on both domains indicating higher levels of pain acceptance. Cronbach's alpha was 0.85 for the Activities Engagement subscale and 0.84 for the Pain Willingness subscale at the pre-treatment assessment; the (ii) PCS,[23] a 13-item self-report measure that asks participants to indicate the degree to which they have certain thoughts or feelings when experiencing pain and that includes the following 3 subscales: rumination (eg, inability to keep pain out of mind) with Cronbach's alpha = 0.95, magnification (eg, fear pain will worsen) with Cronbach's alpha equal to 0.77, and helplessness (eg, feeling overwhelmed by pain) with Cronbach's alpha = 0.94.[23] We specifically asked participants to complete the PCS in relation to their vestibular pain. Items were rated on a scale from 0 (not at all) to 4 (all the time), with higher scores indicating higher levels of catastrophizing; the (iii) Pain Vigilance and Awareness Questionnaire (PVAQ)[24,27] was used to measure awareness of and attention to vulvar pain. The PVAQ is a self-reported measure of attention to pain that assesses pain awareness, consciousness, vigilance, and observation. Respondents are asked to consider their pain experiences, if applicable, over the previous 2 weeks and to indicate the frequency with which each item describes their response to pain. The PVAQ contains 16 items rated on a 0 (never) to 5 (always) scale, with higher scores indicating higher levels of attention to pain. The PVAQ demonstrates good internal consistency and evidence of validity.[27] In the current sample, Cronbach's alpha = 0.92 at the pre-treatment assessment; and (4) ASI is a self-report questionnaire which measures fear of the negative consequences of anxiety.[28] The 16 items are rated on a 5 point-Likert scale ranging from 0 (very little) to 4 (very much). Scores for the ASI are summed, with higher scores indicating higher sensitivity to anxiety. There are 3 subscales of the ASI: Physical Concerns (Cronbach's alpha in current sample = 0.90), Mental Incapacitation Concerns (Cronbach's alpha in current sample = 0.92), and Social Concerns (Cronbach's alpha in current sample = 0.84).

Statistical Analysis

Data from this study were analyzed using IBM SPSS 25.0. We provide descriptive characteristics of the sample using means and proportions. Baseline scores from the FSFI pain subscale, the average intercourse-related pain and the pain rating from the vulvalgesiometer test were collected and a bivariate correlation analysis was done between the different pain measures and scores obtained on the psychological measures. Assumptions underlying correlation analysis were tested for all 3 pain measures. Besides the 2 rating scale pain measures traditionally treated as continuous, the FSFI-Pain measure is a mean of 3 Likert scale questions and as such can also be treated as continuous.[29,30] All measures were tested for presence of outliers, and 1 outlier was removed from FSFI-Pain variable.

All variables were skewed as would be expected in a clinical sample. Pearson's r correlation is still used because it produces a robust estimator of the relationship in larger samples. The non-normality affects the tests of statistical significance, which is not the focus here because, in this study, the strength of pain measures relationships is of primary interest. The examination of scatter plots indicates linearity of the bivariate associations and does not indicate any considerable deviation from homoskedasticity. Additionally, 2 tests of homoskedasticity (Breusch-Pagan and Koenker) were performed for each bivariate association, and none was significant in providing additional support for lack of violation of homoskedasticity. For bivariate correlations, a sample of 64 participants provided us with sufficient (0.80) power to find significant correlations $\geq$0.3 and similar level of power to find medium-size impact of pain measures explaining emotional function variables in hierarchical multivariate regression.

RESULTS

Sample Characteristics

Among the 130 participants, only 64 women reported having engaged in sexual intercourse in the previous 4 weeks before their inclusion in the study and baseline assessment. Because 2 of the 3 pain measures analyzed are based on sexual activity in the previous 4 weeks, the present analyses were based on this subset of 64 women. Where possible, we compared responses between the sexually active vs inactive women. The most common reasons provided for not engaging in vaginal penetration over the previous 4 weeks were fear of pain (45%), advised not to by a clinician (22%), and no current partner (19%).

The 64 sexually active women were slightly younger than the 66 sexually inactive women, and they had higher scores on the cognitive concerns of the Anxiety Sensitivity Index: sexually active women (M = 5.8, SD = 6.12) and inactive women (M = 3.8, SD = 4.8), F(118.9, 127) = 6.7, P = .04. There were no other significant differences on sociodemographic (Table 1) or pain-related characteristics (Table 2) between the 2 groups.

Pain Intensity

Among the sexually active women, the level of intercourse-related pain during the previous 4 weeks was M = 6.31, SD = 2.04, on the 11-point 0 to 10 NRS, M = 2.48 , SD = 1.13 on the FSFI pain subscale, and M = 6.54, SD = 2.00 (0 to 10) during testing with the vulvalgesiometer.

Table 1. Differences in sociodemographic characteristics of women with provoked vestibulodynia who were sexually active in the previous 4 weeks (n = 64) and women who were not (n = 66)

	Sexually active	Sexually inactive	*P*
Age (y), mean (SD)	30.9 (7.3)	34.1 (9.0)	.03
Sexual orientation			.19
Heterosexual	87.5%	90.6%	
Homosexual	0%	3.1%	
Bisexual	12.5%	6.3%	
Relationship status			.28
Single	11.1%	16.9%	
Dating	23.8%	13.8%	
Married/common law	65.1%	69.2%	
Length of relationship (y), mean (SD)	6.6 (5.9)	8.8 (6.9)	.07
Satisfaction with relationship closeness (0-10), mean (SD)	7.7 (2.2)	7.4 (2.1)	.48
Number of children, mean (SD)	0.8 (0.4)	0.8 (0.4)	.97
Ethnicity			.16
Euro-Canadian	74.2%	58.7%	
South/East Asian	11.3%	22.2%	
Other	14.5%	19.0%	
Highest education			.67
High school	3.5%	1.8%	
College/technical/trade school	22.8%	24.6%	
Undergraduate degree	52.6%	42.1%	
Master's degree	17.5%	24.6%	
Doctoral degree or MD	3.5%	7.0%	

MD = medical doctor degree.

Correlation of Pain Measures

Bivariate correlations between pain measures were conducted first. Next, to examine whether the severity of PVD condition affects the strength of bivariate pain measure associations, partial correlations were performed controlling for the length of PVD and taking medication for PVD. There was a significant negative correlation between the vulvalgesiometer rating and the FSFI pain subscale score ($r = -0.390$, $P = .001$; partial $r = -0.373$, $P = .006$), indicating that as vulvalgesiometer pain increased, FSFI pain scores decreased (ie, more pain). Similarly, intercourse-related pain rated with the NRS and scores on the FSFI pain subscale were significantly negatively related ($r = -0.605$, $P < .001$; partial $r = -0.592$, $P < .001$) such that increases in self-reported pain with intercourse were associated with lower scores (ie, more pain) on the FSFI pain domain. The correlation between the vulvalgesiometer rating and the numeric rating scale of pain was positive ($r = 0.504$, $P < .001$; partial $r = 0.578$, $P < .001$), indicating that increases in NRS were associated with increases in vulvalgesiometer ratings.

Correlations ranged from moderate (between FSFI-Pain and vulvalgesiometer) to large (between NSR and the other 2 measures). Steiger's Z test for comparing dependent correlations [31,32] was used to compare the coefficients and found that the only significant difference was between the FSFI-Pain—vulvalgesiometer rating correlation and FSFI-Pain—NRS rating correlation ($r = 0.390$ vs $r = 0.605$, $Z = 2.06$, $P = .039$). Partial correlation comparisons revealed that the FSFI-Pain—vulvalgesiometer rating correlation is significantly smaller than the other 2 correlations ($r = 0.373$ vs $r = 0.578$ and vs $r = -0.592$, $Z = 2.11$, $P = .035$ and $Z = 2.23$, $P = .025$, respectively).

Association of Pain and Psychological Domains

Unique association of each pain measure with measures of emotional function was examined by a series of hierarchical linear regression models. Each pain measure was added as a predictor in a separate step allowing for evaluation of its unique exploratory contribution for each emotional function outcome (NRS was entered first, followed by FSFI Pain and vulvalgesiometer). 3 control variables (age, ethnicity, and PVD duration) were also entered. The unstandardized beta coefficients for all regression models including changes in the amount of explained variance are presented in Table 3. The results indicated that of the 3 pain measures, the NRS measure is the best predictor of emotional function outcomes. It was a significant predictor of pain catastrophizing and pain hypervigilance. None of the pain measures predicted chronic pain acceptance or anxiety sensitivity scores. Adding FSFI Pain and vulvalgesiometer scores to the models after NRS was already entered did not increase the amount of explained variance for any of the outcomes indicating that these 2 measures of pain do not have any unique association with emotional function above and beyond that of NRS.

DISCUSSION

Here we attempted to replicate the findings of Aerts et al,[15] who examined the correlations between multiple measures of pain. We found statistically significant, moderate to large correlations for all pairwise combinations of the 3 pain measures, with the highest correlation between intercourse-related pain on the NRS and the FSFI pain subscale ($r = -0.605$). Moderate to large correlations between vulvalgesiometer ratings and self-report measures mirror the findings of another study[12] that found large correlations between pressure-pain thresholds and intensity, and NRS pain ratings. This study adds to this literature by reporting correlations (moderate) between vulvalgesiometer and another averaging self-report measure—FSFI-Pain.

In contrast to our findings, Aerts et al[15] did not find any significant correlations between vulvalgesiometer ratings of the pressure-pain threshold and women's self-reported pain or their FSFI total scores. It is possible that differences in study design account for the contradictory findings. For example, in this study, we used the vulvalgesiometer to exert a fixed pressure of 30 g at all locations around the vulvar vestibule in random order, whereas Aerts et al[15] assessed participants' pressure pain threshold only at the 3 o'clock location. Furthermore, Aerts et al[15] asked women to rate the experience at the threshold on 2 NRSs—1 rating the pain

Table 2. Participant characteristics and pain levels for the women who were sexually active in the past 4 weeks (n = 64) and women who were not (n = 66)

	Sexually active	Sexually inactive	*P*
PVD history			.33
Primary (first penetration)	57.8%	66.2%	
Secondary (acquired)	42.2%	33.8%	
Previous attempts to treat PVD (Yes)	54.7%	46.2%	.33
Current medication to treat PVD (Yes)	20.3%	9.2%	.08
Years since diagnosis	8.00 ± 6.46	8.48 ± 6.92	.69
PVD pain levels mean (SD)			
Usual pain level (0-10)	5.55 (1.89)	6.32 ± 1.90	.02
Level of worst pain (0-10)	8.17 (1.15)	8.29 ± 1.27	.57
Baseline Pain NRS	6.31 (2.04)	—	
Baseline VVG score	6.54 (2.00)	6.72 ± 2.34	.64
Baseline FSFI pain score	2.48 (1.13)	—	

FSFI = female sexual function index; NRS = numeric rating scale; PVD = provoked vestibulodynia; VVG = vulvalgesiometer.
Averages shown in percentages or mean (SD).
Sexually inactive women over the previous 4 weeks were coded as such in response to the answer they provided on the numeric rating scale.

on a scale from 0−10 (similar to our study) and 1 rating the unpleasantness on a scale from 0−10. Because the correlation between these 2 NRSs was >0.6, they were combined to create a single score. As such, their NRS was based on a composite pain/unpleasantness rating as opposed to pain alone. When Aerts et al[15] focused their analysis on the FSFI pain domain (not total score), they did find a significant correlation with their measure of intercourse-related pain, which is consistent with our study findings, although their associations were smaller than in this study.

What can be made of the findings that our observed correlations between measures were moderate to large? Although each of these have independently been used in treatment outcome studies of PVD,[26,33] only 1 study, as far as we are aware, used all 3 of these measures,[14] and this study was carried out before IMMPACT recommendations about pain measurement were made. 1 interpretation of the current correlations is that, in addition to measuring similar dimension of pain, each of these measures also taps into a different aspect of women's pain experience with PVD and does not measure the same construct.

The finding that the 2 self-reported measures of pain yielded higher correlations than the correlation of either of them with the vulvalgesiometer deserves mention. Because the numeric rating scale and the FSFI pain scale both tap into pain with penetrative sexual activity and require women to recall actual sexual encounters that transpired at home, this may explain why those 2 measures had a higher degree of overlap. On the other hand, measurement of pain using the vulvalgesiometer took place in a clinical setting in a different environment and a different interpersonal context where the women may have perceived less control, potentially affecting their pain experience; in the privacy of their own relationship and home, women might have a very different experience of their vulvar pain than they do in a clinical setting where testing is being done by physicians. This suggests that the vulvalgesiometer pain ratings may reflect a different aspect of the pain experience than do women's self-reported pain intensity with sexual activity. It may be that using the vulvalgesiometer as a proxy for sexual activity in women may offer only an imperfect approximation, given that the NRS and the FSFI pain scale rely on sexual context and sexual arousal.[34,35] Compared with the cotton swab test, however, the vulvalgesiometer is a better objective measure of vulvar pain that is correlated with self-reported pain with intercourse, whereas the cotton swab was not found to be correlated with intercourse pain[15] or with FSFI-Pain.[36] Both the cotton swab and the vulvalgesiometer, however, involve administering a discrete stimulus that does not mimic the friction and movement associated with sexual penetration and thus likely creates a different pain response.

Measurement of the sensory aspects of pain has typically relied on measures of static mechanical allodynia, where a fixed touch is applied to the vulva and the participant self-reports a rating of pain. However, there is a dynamic aspect of pain perception that may not be captured by either a vulvalgesiometer or a cotton swab test. In an experimental comparison of the psychophysical properties of static vs dynamic touch to the vulva (using a brush), Farmer et al[37] found that the static vs dynamic measures differed in their sensitivity to the temporal aspects of pain, and these differences may relate to activation of different ß fibers.[38]

In terms of the representativeness of our sample, we found that the levels of pain experienced by the current sample of women were similar using the NRS (6.31 ± 2.04) and the FSFI pain subscale (2.48 ± 1.13) to other studies of PVD. This suggests that our sample is comparable to other published studies of sexually active women with PVD.[14−16,18−20] These pain levels are not atypical for women with PVD but instead reflect the fact that women have at least moderate levels of pain with vulvar contact.

Not surprisingly, pain is a complex biopsychosocial construct and is based entirely on self-report because there has yet to be a validated signature of pain.[39] The experience of pain depends on the type and intensity of the stimulus, the context in which it is

Table 3. Hierarchical regression models predicting emotion function outcomes by the 3 pain measures

Measure	CPAQ Will	CPAQ Act	PCS Rum	PCS Magn	PCS Helpls	PCS Total	PVAQ	ASI Phys	ASI Cogn	ASI Soc	ASI Total
Step 1—Controls	$R^2 = 0.09$	$R^2 = 0.03$	$R^2 = 0.12$	$R^2 = 0.14$	$R^2 = 0.12$	$R^2 = 0.12$	$R^2 = 0.04$	$R^2 = 0.15$	$R^2 = 0.16$	$R^2 = 0.16$	$R^2 = 0.18$*
Age	0.150	0.130	0.002	−0.097	−0.034	−0.163	−0.087	−0.090	0.063	−0.042	−0.058
Ethnicity											
South-EA	4.669	4.398	−2.174	−2.083	−2.264	−8.786	−7.902	0.703	0.243	0.669	1.610
Other	−1.923	0.031	4.247*	1.056	2.312	9.930	1.683	4.095	5.938*	6.018†	16.134*
PVD Months	0.021	0.010	0.002	−0.005	0.002	0.001	−0.013	−0.017	−0.022	−0.010	−0.050
Step 2	$\Delta R^2 = 0.01$	$\Delta R^2 = 0.01$	$\Delta R^2 = 0.16$*	$\Delta R^2 = 0.13$†	$\Delta R^2 = 0.23$‡	$\Delta R^2 = 0.22$‡	$\Delta R^2 = 0.13$†	$\Delta R^2 = 0.05$	$\Delta R^2 = 0.06$	$\Delta R^2 = 0.02$	$\Delta R^2 = 0.05$
NRS	−0.338	−0.605	0.955†	0.559†	0.812‡	3.138‡	2.313†	0.639	0.788	0.370	1.805
Step 3	$\Delta R^2 = 0.002$	$\Delta R^2 = 0.08$*	$\Delta R^2 = 0.003$	$\Delta R^2 = 0.002$	$\Delta R^2 = 0.003$	$\Delta R^2 = 0.000$	$\Delta R^2 = 0.004$	$\Delta R^2 = 0.003$	$\Delta R^2 = 0.002$	$\Delta R^2 = 0.005$	$\Delta R^2 = 0.000$
NRS	−0.180	−1.722	1.053†	0.605*	0.741†	3.138†	2.003	0.529	0.698	0.513	1.760
FSFI-Pain	0.476	−3.357*	0.311	0.144	−0.225	0.002	−0.984	−0.351	−0.284	0.457	−0.141
Step 4	$\Delta R^2 = 0.003$	$\Delta R^2 = 0.05$	$\Delta R^2 = 0.05$	$\Delta R^2 = 0.001$	$\Delta R^2 = 0.02$	$\Delta R^2 = 0.03$	$\Delta R^2 = 0.004$	$\Delta R^2 = 0.001$	$\Delta R^2 = 0.003$	$\Delta R^2 = 0.02$	$\Delta R^2 = 0.004$
NRS	0.016	−0.859	1.452†	0.633*	0.920†	3.925†	1.725	0.591	0.571	0.218	1.397
FSFI-Pain	0.485	−3.315	0.294	0.142	−0.232	−0.031	−0.972	−0.353	−0.279	0.469	−0.126
Vulvalges	−0.326	−1.437	−0.715	−0.052	−0.321	−1.410	0.499	−0.111	0.228	0.528	0.651

ASI Cogn = Anxiety Sensitivity Index Cognitive Concerns Scale; ASI Phys = Anxiety Sensitivity Index Physical Concerns Scale; ASI Soc = Anxiety Sensitivity Index Social Concerns Scale; ASI Total = Anxiety Sensitivity Index Total Scale; CPAQ Act = CPAQ Activities Engagement; CPAQ Will = CPAQ Pain Willingness; FSFI-Pain = Female Sexual Function Index—Pain; NRS = numeric rating scale; PCS Helpls = Pain Catastrophizing Helplessness Scale; PCS Rum = Pain Catastrophizing Rumination Scale; PCS Magn = Pain Catastrophizing Magnification Scale; PCS Total = Pain Catastrophizing Total Scale; PVAQ = Pain Vigilance and Awareness Questionnaire; PVD = provoked vestibulodynia; Vulvalges = vulvalgesiometer.

For Ethnicity, white is a reference group.

*$P < .05$.

†$P < .01$.

‡$P < 001$.

experienced, and the emotional state of the individual.[40] The results of regression analyses indicate that neither FSFI-Pain nor vulvalgesiometer have any unique association with emotion function measures. NRS is the only pain indicator that predicts some of the emotion outcomes (pain catastrophizing and pain vigilance). This may have some implications. More research is needed to determine what psychological, physical, emotional, or social function vulvalgesiometer rating is predictive of or associated with to efficiently use that measure in future research on PVD. Second, overall, pain measures explained only a small amount of variance in emotion function variables, and NRS was the only variable significantly associated with emotion function among the 3 pain measures. This finding stresses the importance of including emotion function measures in the studies of PVD in addition to direct pain measures because they appear to tap into distinct aspects of pain experience. Because pain does not appear to be directly associated with overall sexual function,[41] the present findings lend support for the role of other psychological factors such as fear of pain, hypervigilance, catastrophizing, and low levels of self-efficacy may serve as predictors of distinct pain aspects and sexual functioning.[41–44]

We might conclude that the NRS is an acceptable proxy for intercourse-related pain based on the present findings. The NRS and FSFI Pain, 2 pain-averaging retrospective measures, have, not surprisingly, the highest correlation whereas the responsive vulvalgesiometer measure has a moderate correlation with NRS and the lowest correlation with FSFI Pain. It would position NRS as the most comprehensive measure—it shares a lot of variance with FSFI Pain but also a moderate amount of variance with VVG and should be recommended when only 1 measure of pain is to be administered. An additional advantage of using the NRS measure is that a clinically meaningful change has been established for it (2 point[6]). Because the IMMPACT recommends using $\geq$1 anchor-based method in addition to distribution-based information, NRS can provide an estimate of anchor-based amount of change in response to treatment. However, even the NRS has been criticized, given that not all steps on the scale have the same importance or value, eg, a change in pain on the scale from 3 to 1 on the NRS is of greater importance than a change in pain rating from 6 to 4.[45] Also, consistency in the description of the endpoint is important for the way women respond. Endpoints such as "worst possible pain," "worst pain imaginable," and "worst pain ever felt" may seem like interchangeable descriptors; however, it is known that these varying anchors can impact how people rate scores on numerical scales,[46] and other studies evaluating the use of Likert-based scales show that people are less likely to use the values at either end of the scale.[47]

Strengths and Limitations of the Study

A strength of this study was the adequate sample size (n = 64 sexually active women) of patients who had received confirmed clinical diagnoses of PVD. Although we had to exclude a rather large group of women who had not been sexually active in the preceding 4 weeks (n = 66), we also found that these 2 groups differed only in that the inactive women reported higher levels of typical vulvar pain compared with the sexually active groups. No other demographic or clinical variable differed between these 2 groups. Although the sexually active and inactive women appeared not to differ with regard to baseline pain and other characteristics, we cannot rule out the possibility that they may have different correlations between pain measures or between pain and emotional function variables. For example, levels of fear of pain may differ between groups in a way that impacts both their self-report ratings and subsequently the correlations between measures. We did not measure fear of pain in this study to test this hypothesis.

1 study limitation is the fact that both the NRS and the FSFI pain subscale included retrospective ratings and averages of the intercourse-related pain for the past 4 weeks. Even though both measures have been validated,[10,12,21] it may be that other variables, such as mood, could have impacted scores on these measures by operating through a mood congruency bias.[48] A study by Weinfurt et al[49] showed that a recall period of 1 month is acceptable and fairly reliable when it comes to variables such as sexual functioning and vaginal discomfort. Unfortunately, our participants were not tested at the same time in their menstrual cycle, and this may have introduced variability into pain ratings, given the known effects of menstrual cycle phase on self-reports of pain.[50,51]

This study did not differentiate between whether the women suffered from primary (ie, when pain started with the first vaginal penetration, first tampon use, first intercourse, first invasive gynecologic examination) or secondary PVD (ie, when women had a period of pain-free penetration before the onset of PVD).[16] Future studies should examine whether the overlap between pain measures depends on PVD subtype. Our study, however, has examined potential impact of the severity of PVD by partialing out 2 severity-related variables—the length of PVD and taking medication for PVD from the pain measure correlations—and we found that the impact of the PVD severity on the relationships between the pain measures was negligible.

In line with IMMPACT ratings, we continue to recommend the numeric rating scale as 1 endpoint in treatment outcome studies of PVD, but it is important for researchers to know that this may be influenced by memory and does not account for differences in pain intensity across a number of penetration attempts over the past year. On the other hand, in-lab assessment of pain using a static measure, such as the vulvalgesiometer, is not the same as a dynamic measure and may represent the more sensory and not affective aspects of pain. The implications of the findings for clinical practice are that not all pain measures are created equal and that not all available pain measures map exactly onto emotional function and well-being. It thus remains important, as recommended by IMMPACT, that emotional function measures be included as separate from pain sensation measures.

CONCLUSION

This study showed statistically significant, moderate to large correlations among 3 different pain measures widely used in PVD pain assessment. It also demonstrated that only 1 pain measure was uniquely associated with emotion function measures associated with pain. Such findings indicate that clinically tested pain does not necessarily correspond with the pain sexually active women feel in the safety of their own home and relationship and highlight the importance of using both self-reported measures and quantitative sensory measuring, as well as measures of emotion function. Because the study sample was comparatively large and pain ratings were similar to earlier studies of sexually active women, the generalizability of the findings of this study may be regarded as adequate.

ACKNOWLEDGMENTS

A special thank you to Bent Husum for editorial assistance with this paper and to Adrienne Marsh for managing all aspects of data collection.

Corresponding Author: Lori A. Brotto, 2775 Laurel Street, Vancouver, BC, Canada, V5Z 1M9. Tel: (604) 875-4111 ext. 68898; Fax: 604-875-4869; E-mail: Lori.Brotto@vch.ca

Conflict of interest: None.

Funding: Funding for the larger study from which these data were extracted was provided by the Canadian Institutes of Health Research to the senior author.

STATEMENT OF AUTHORSHIP

Category 1

(a) Conception and Design
Bozena Zdaniuk; Lori A. Brotto

(b) Acquisition of Data
Lori A. Brotto

(c) Analysis and Interpretation of Data
Frederikke Louise Wammen Rathenborg; Bozena Zdaniuk; Lori A. Brotto

Category 2

(a) Drafting the Article
Frederikke Louise Wammen Rathenborg; Bozena Zdaniuk; Lori A. Brotto

(b) Revising the Article for Intellectual Content
Frederikke Louise Wammen Rathenborg; Bozena Zdaniuk; Lori A. Brotto

Category 3

(a) Final Approval of the Completed Article
Frederikke Louise Wammen Rathenborg; Bozena Zdaniuk; Lori A. Brotto

REFERENCES

1. Harlow BL, Wise LA, Stewart EG. Prevalence and predictors of chronic lower genital tract discomfort. **Am J Obstet Gynecol 2001;185:545-550.**
2. Reed BD, Harlow SD, Sen A, Legocki LJ, Edwards RM, Arato N, Haefner HK. Prevalence and demographic characteristics of vulvodynia in a population-based sample. **Am J Obstet Gynecol 2012;206:170.e1-170.e9.**
3. Bornstein J, Goldstein AT, Stockdale CK, et al. 2015 ISSVD, ISSWSH and IPPS consensus terminology and classification of persistent vulvar pain and vulvodynia. **Obstet Gynecol 2016;127:745-751.**
4. Henzell H, Berzins K, Langford J. Provoked vestibulodynia: Current perspectives. **Int J Womens Health 2017;9:631-642.**
5. De Andres J, Sanchis-Lopez N, Asensio-Samper M. Vulvodynia—An evidence based literature review and proposed treatment algorithm. **Pain Pract 2016;16:204-236.**
6. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. **J Pain 2008;9:105-121.**
7. Pukall CF, Bergeron S, Brown C, et al. Recommendations for self-report outcome measures in vulvodynia clinical trials. **Clin J Pain 2017;33:756-765.**
8. Lester RA, Brotto LA, Sadownik LA. Provoked vestibulodynia and the health care implications of comorbid pain conditions. **J Obstet Gynaecol Canada 2015;37:995-1005.**
9. Farrar JT, Young JP, LaMoreaux L, et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. **Pain 2001;94:149-158.**
10. Rosen R, Brown C, Heiman J, et al. The Female Sexual Function Index (FSFI): A multidimensional self-report instrument for the assessment of female sexual function. **J Sex Marital Ther 2000;26:191-208.**
11. Pukall CF, Payne KA, Binik YM, et al. Pain measurement in vulvodynia. **J Sex Marital Ther 2003;29:111-120.**
12. Pukall CF, Young RA, Roberts MJ, et al. The vulvalgesiometer as a device to measure genital pressure-pain threshold. **Physiol Meas 2007;28:1543-1550.**
13. Pukall CF, Binik YM, Khalifé S, et al. Vestibular tactile and pain thresholds in women with vulvar vestibulitis syndrome. **Pain 2002;96:163-175.**
14. Sutton KS, Pukall CF, Chamberlain SM. Pain ratings, sensory thresholds, and psychosocial functioning in women with provoked vestibulodynia. **J Sex Marital Ther 2009;35:262-281.**
15. Aerts L, Bergeron S, Pukall CF, et al. Provoked vestibulodynia: Does pain intensity correlate with sexual dysfunction and dissatisfaction? **J Sex Med 2016;13:955-962.**
16. Lowenstein L, Vardi Y, Deutsch M, et al. Vulvar vestibulitis severity—Assessment by sensory and pain testing modalities. **Pain 2004;107:47-53.**

17. First MB, Spitzer RL, Gibbon M, et al. Structured clinical interview for DSM-IV Axis II personality disorders, (SCID-II). Washington, D.C.: American Psychiatric Press, Inc.; 1997.
18. Sutton KS, Pukall CF, Chamberlain SM. Pain, psychosocial, sexual, and psychophysical characteristics of women with primary vs. secondary provoked vestibulodynia. **J Sex Med 2009;6:205-214.**
19. Smith KB, Pukall CF, Chamberlain SM. Sexual and relationship satisfaction and vestibular pain sensitivity among women with provoked vestibulodynia. **J Sex Med 2013;10:2009-2023.**
20. Goldfinger C, Pukall CF, Gentilcore-Saulnier E, et al. A prospective study of pelvic floor physical therapy: Pain and psychosexual outcomes in provoked vestibulodynia. **J Sex Med 2009;6:1955-1968.**
21. Hawker GA, Mian S, Kendzerska T, et al. Measures of Adult Pain. Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). **Arthritis Care Res 2011;63:240-252.**
22. McCracken LM, Carson JW, Eccleston C, et al. Acceptance and change in the context of chronic pain. **Pain 2004;109:4-7.**
23. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. **Psychol Assess 1995; 7:524-532.**
24. McCracken LM, Klock PA, Mingay DJ, et al. Assessment of satisfaction with treatment for chronic pain. **J Pain Symptom Manage 1997;14:292-299.**
25. Peterson RA, Heilbronner RL. The anxiety sensitivity index: Construct validity and factor analytic structure. **J Anxiety Disord 1987;1:117-121.**
26. Pukall CF, Binik YM, Khalifé S. A new instrument for pain assessment in vulvar vestibulitis syndrome. **J Sex Marital Ther 2004;30:69-78.**
27. McCracken LM. "Attention" to pain in persons with chronic pain: A behavioral approach. **Behav Ther 1997;28:271-284.**
28. Taylor S, Zvolensky MJ, Cox BJ, et al. Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. **Psychol Assessment 2007; 19:176-188.**
29. Sullivan GM, Artino AR Jr. Analyzing and interpreting data from Likert-type scales. **J Grad Med Educ 2013;5:541-542.**
30. Zumbo BD, Zimmerman DW. Is the selection of statistical methods governed by level of measurement? **Can Psychol 1993;34:390-400.**
31. Steiger JH. Tests for comparing elements of a correlation matrix. **Psychol Bull 1980;87:245-251.**
32. Hoerger M. Z_H: An updated version of Steiger's Z and web-based calculator for testing the statistical significance of the difference between dependent correlations. 2013. Retrieved from, http://www.psychmike.com/dependent_correlations.php. Accessed November 2, 2018.
33. Sutton K, Pukall C, Wild C, et al. Cognitive, psychophysical, and neural correlates of vulvar pain in primary and secondary provoked vestibulodynia: A pilot study. **J Sex Med 2015; 12:1283-1297.**
34. Pukall CF, Meana M, Sutton KS. Psychological evaluation and measurement of dyspareunia. In: Goldstein AT, Pukall CF, Goldstein I, eds. Female Sexual Pain Disorders. Hoboken, NJ: Blackwell Publishing Ltd; 2009. p. 21-26.
35. Rosen NO, Bergeron S, Leclerc B, et al. Woman and partner-perceived partner responses predict pain and sexual satisfaction in provoked vestibulodynia (PVD) couples. **J Sex Med 2010;7:3715-3724.**
36. Masheb RM, Kerns RD, Lozano C, et al. A randomized clinical trial for women with vulvodynia: Cognitive-behavioral therapy vs supportive psychotherapy. **Pain 2009;141:31-40.**
37. Farmer MA, Maykut CA, Huberman JS, et al. Psychophysical properties of female genital sensation. **Pain 2013;154:2277-2286.**
38. Landerholm ÅH, Hansson PT. The perception threshold counterpart to dynamic and static mechanical allodynia assessed using von Frey filaments in peripheral neuropathic pain patients. **Scand J Pain 2011;2:9-16.**
39. Cowen R, Stasiowska MK, Laycock H, et al. Assessing pain objectively: The use of physiological markers. **Anaesthesia 2015;70:828-847.**
40. Lumley MA, Cohen JL, Borszcz GS, et al. Pain and emotion: A biopsychosocial review of recent research. **J Clin Psychol 2011; 67:942-968.**
41. Bergeron S, Rosen NO, Morin M. Genital pain in women: Beyond interference with intercourse. **Pain 2011;152:1223-1225.**
42. Desrochers G, Bergeron S, Khalifé S, et al. Fear avoidance and self-efficacy in relation to pain and sexual impairment in women with provoked vestibulodynia. **Clin J Pain 2009; 25:520-527.**
43. Desrochers G, Bergeron S, Khalifé S, et al. Provoked vestibulodynia: Psychological predictors of topical and cognitive-behavioral treatment outcome. **Behav Res Ther 2010; 48:106-115.**
44. Lemieux AJ, Bergeron S, Steben M, et al. Do romantic partners' responses to entry dyspareunia affect women's experience of pain? The roles of catastrophizing and self-efficacy. **J Sex Med 2013;10:2274-2284.**
45. Beilin Y, Hossain S, Bodian CA. The numeric rating scale and labor epidural analgesia. **Anesth Analg 2003;96:1794-1798.**
46. Schwarz N, Knauper B, Hippler H-J, et al. Rating scales: Numeric values may change the meaning of scale labels. **Public Opin Q 1991;55:570-582.**
47. Ogden J, Lo J. How meaningful are data from Likert scales? An evaluation of how ratings are made and the role of the response shift in the socially disadvantaged. **J Health Psychol 2012;17:350-361.**
48. Sato H, Kawahara JI. Selective bias in retrospective self-reports of negative mood states. **Anxiety Stress Coping 2011;24:359-367.**

49. Weinfurt KP, Lin L, Dombeck CB, et al. Accuracy of 30-day recall for components of sexual function and the moderating effects of gender and mood. **J Sex Med 2014;11:678-696.**

50. Besson M, Piguet V, Desmeules J, et al. Influence of menstrual cycle on experimental pain threshold [Influence du cycle menstruel sur le seuil de la douleur expérimentale]. **Pain and Pain Relief [Douleur et Analgésie] 2004;17:73-76 [in French].**

51. Anderberg UM, Marteinsdottir I, Hallman J, et al. Variability in cyclicity affects pain and other symptoms in female fibromyalgia syndrome patients. **J Musculoskelet Pain 1998;6:5-22.**