Reliability and Reproducibility of Novel Methodology for Assessment of Pressure Pain Sensitivity in Pelvis

Denniz Zolnoun,*,z Eric Bair,y,x Greg Essick,z,x Richard Gracely,x Vinita Goyal,*,and William Maixnerz

*Department of Obstetrics and Gynecology, Pelvic Pain Research Unit, University of North Carolina, Chapel Hill, North Carolina.

yDepartment of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina.

zDepartment of Prosthodontics, and xCenter for Neurosensory Disorders, School of Dentistry, University of North Carolina, Chapel Hill, North Carolina.

{Department of Obstetrics and Gynecology, Brown University and Women & Infants Hospital of Rhode Island, Providence, Rhode Island.

Abstract:

Vestibulodynia, the most common type of chronic vulvovaginal pain, impairs the psychological, physical health of nearly 10% of women at some point in their lifetime. The aim of this investigation was to establish reliable standardized methodologies for assessment of pain sensitivity in vulvar mucosa and pelvic musculature. We enrolled 34 women with vestibulodynia and 21 pain free controls. The participants underwent a nuanced exam that consisted of palpation of precisely located vulvar mucosal and pelvic muscle sites. These measurements remained highly stable when participants were reexamined after 2 weeks, with high within-examiner correlation.

Vestibulodynia patients reported greater sensitivity than pain-free controls at the majority of examination sites, particularly at mucosal sites on the lower vestibule. The pain threshold measures at the lower mucosal sites were also associated with the participants' self-reported pain levels during intercourse. These mucosal pain threshold measurements were used to discriminate between vestibulodynia cases and controls with high sensitivity and specificity.

This data supports the feasibility of contemporaneous assessment of vulvar mucosa and underlying musculature in the pelvic region, offering the hope of a more precise case definition for vestibulodynia and related disorders.

Perspective: This study describes performance characteristics of novel methodologies for assessing pelvic muscle and mucosal sensitivity. These pain sensitivity measures were reproducible and associated with subjective pain reports and vestibulodynia case status and represent an important step toward a more precise case definition for vestibulodynia and related disorders.

Vestibulodynia, the most common type of chronic vulvovaginal pain (vulvodynia), impairs the physical, reproductive, and psychological health of nearly 1 in 10 women at some point in their lifetime.7
Because little is known about the etiology and optimal management of this disorder, most sufferers are left disillusioned. 13 Historically, vestibulodynia has been viewed as a localized disorder of the vulvar mucosa and surrounding region (specifically, underlying musculature).
However, over the past decade, a growing body of literature provides evidence in support of vestibulodynia as an idiopathic pain disorder (akin to fibromyalgia and temporomandibular pain disorder) with abnormalities in both peripheral and central nervous system processes. 2,19 The clinical diagnostic criteria for vestibulodynia dates back to Friedrich's case series of 86 patients published in 1987.5Heproposedthe following3 clinical criteria (known as Friedrich criteria) for establishing the diagnosis of the elusive condition he called ''VulvarVestibulitis Syndrome'':

1) entry dyspareunia (painful intercourse); 2) tenderness to pressure within the vestibule; and 3) physical findings limited to erythema without other obvious pathology.5 Over the past decade, the body of literature on vestibulodynia has expanded. However, revised criteria (based on elements of the history and an objective clinical exam) beyond Friedrich's initial description has not been formulated.

During clinical examination, the vulvar mucosal tissue is typically assessed by the application of a cotton swab, but the reliability and reproducibility of

the patient's response are generally poor. For example, intraexaminer correlations as high as .73 (upper vestibule, Fig 1) and as low as .38 (lower vestibule, Fig 1) have been reported.9 Over the past decade, many researchers have attempted to standardize the exam methodology.

12 Pukall et al10,11 was the first to standardize the pressure application upon recognizing the challenges of pain assessment in the vulvar region. Their group developed 7 spring-loaded syringes capable of applying a wide range of pressures (3–1,000 g). The examiner sequentially applies higher pressure (using the preloaded syringe) until the tactile perception or pain detection threshold is reached.10 However, the stability and reproducibility have not been evaluated, and the anatomical locations of tested sites are not well characterized.

Unlike mucosa, the spectrum of normal and abnormal tenderness in the pelvic floor and its constituent muscles (eg, puborectalis muscle) has not been described. For evaluation of pelvic muscle dysfunction, clinicians have utilized various subjective scales and surface electromyography (sEMG) to assess voluntary control as a proxy for normal tone and function. The usefulness and generalizability of various scales and sEMG are severely limited by poor methodological design and lack of normative data.8,16 Aside from these techniques, there is limited data on the assessment of pressure pain sensitivity in pelvic musculature using a mechanical algometer in pain-free women.17

Our overarching hypothesis is that vestibulodynia is a heterogeneous disorder and that mucosal and muscle pain sensitivities can be used to identify distinct subgroups of patients. The purpose of this study was threefold:

1) to develop instrumentation and methodology for the objective assessment of vulvar mucosal and pelvic muscle sensitivities; 2) to establish a standardized clinical approach for assessing pain sensitivity in vulvar mucosa and pelvic musculature; and 3) to relate these measures to patients' clinical pain reports. We began this work by developing and validating tools capable of standardizing the force applied to the mucosa and muscle, enabling a wide spectrum of pain sensitivities to be quantitatively and reproducibly assessed. We validated these instruments on women who experienced pain since coital debut or attempted tampon use during adolescence.

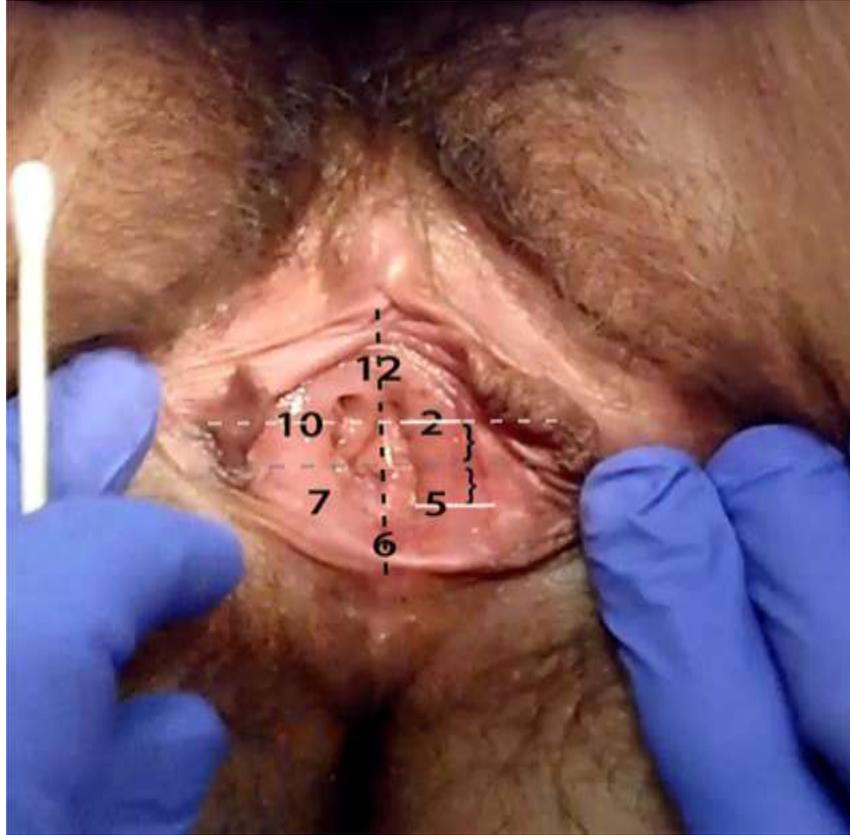1 These women constitute the most distressed subgroup of vestibulodynia patients.20

Figure 1. Anatomical location of the tested mucosal sites.

Methods
Participants

Thirty-four women with vestibulodynia and 21 painfree controls were recruited through an institutional review board-approved protocol at the University of North Carolina, Chapel Hill between March 2006 and August 2009.This study was nested within a larger study (HD053631, PI D. Zolnoun) aimed at identifying and characterizing clinical phenotypes of women with complaints of provoked pain in the mucosa of the vulvar vestibule. A preliminary eligibility assessment of interested participants was conducted via a telephone interview. A positive response to any of the following criteria lead to exclusion: 1) age <21 or age >45; 2) breastfeeding/ pregnant, or menopausal; 3) significant medical conditions (eg, seizure disorder, diabetes, thyroid disorde) and 4) known diagnosis of comorbid urogenital pain conditions (eg, nonmenstrual daily pain localized to the pelvic region, interstitial cystitis). Upon confirmation of general

eligibility (the exclusions above did not apply) potential subjects were queried regarding vulvovaginal symptoms. Women who endorsed unprovoked pain/discomfort (such as burning, itching, knifelike shooting pain) with or without concomitant provoked pain upon contact were then excluded. Only women without vulvovaginal complaints (potential controls) or isolated complaint of provoked pain with intercourse and/or tampon use (potential cases) were scheduled for the clinical eligibility assessment.

On the day of the clinical eligibility assessment, the research coordinator confirmed the accuracy of the participant's responses during the initial phone interview, followed by a description of the study and informed consent. Participants were asked to complete a pain questionnaire (the modified Gracely Pain Scale) after which they underwent a screening gynecological exam. Using the Gracely Pain Scale, women rated the lowest, average, and maximal pain with intercourse on a scale of 0 to100.[6,20] Modified versions of this questionnaire are used by our research program for assessing pain in patients with other idiopathic pain disorders (eg, temporomandibular disorder and fibromyalgia) in addition to vestibulodynia. Vestibulodynia was diagnosed based on the subjective report of pain during intercourse (or tampon use) and tenderness to touch upon palpation of the vestibule with a cotton swab during the gynecological exam.[10] This diagnosis was rendered only after the examining clinician failed to identify any infectious or dermatological etiologies for the positive responses. In the course of the screening examination, participants with suspected dermatological disorders (eg. lichen sclerosis, contact dermatitis), and vaginismus were excluded (n = 11 for the entire cohort).

Study Design

Once eligibility as a case or control was established, participants underwent a nuanced vulvar mucosal and pelvic muscle pain sensitivity assessment. For some subjects the exam was repeated during the same session by a second examiner, or by 1 or 2 of the same examiners in a separate session approximately 2 weeks later. The order of the second examination was reversed such that the second examiner in session 1 was the first examiner in session 2, etc. Within each session, the second examination (when conducted) was performed following a 5- to 10-minute rest period.

Our nuanced exam consisted of a standardized approach to stimulation of precisely located vulvar mucosal and pelvic muscle sites (described below). This approach was adapted from existing constructs on structured clinical exams in the orofacial region for which a standardized methodology for mucocutaneous and muscle pain sensitivity assessment has long been utilized.3 Six mucosal sites and 3 muscle sites were examined in a predetermined order. Pain perception thresholds were collected at each of the 6 mucosal sites. Pain tolerance measures were not collected at the mucosal sites due to excessive sensitivity of the vestibule in patients with vestibulodynia. Both pain perception threshold and pain tolerance measures were collected from each of the 3 muscle sites. A prototypical experimental session consisted of assessment of the vestibula r mucosal pressure pain perception threshold at each site, followed by the pelvic muscle pressure pain perception threshold and pain tolerance testing at each site, which were conducted in a predetermined order described below.

We first examined the upper vestibular sites (Fig 1), which are least sensitive clinically, in order to decrease the likelihood of obtaining lower threshold values if the more sensitive lower vestibule were examined first.10,12 Thus, the examination set yielded a minimum of 18 data points (3 at each of the 6 vestibular mucosa sites) in the course of the mucosal pain perception assessment, and 9 data points (3 per each of the 3 muscle sites), in the course of the muscle pain perception and tolerance testing each, respectively.

Our data include mucosal threshold measurements from 34 vestibulodynia patients and 21 pain-free controls, muscle threshold measurements from 33 vestibulodynia patients and 22 pain-free controls, and muscle tolerance measurements from 31 vestibulodynia patients and 20 pain-free controls. The data from the second visit of 2 subjects was excluded due to calibration problems with the equipment on the day of these visits. Additionally, the maximal mucosal threshold was set at 4 N. Mucosal measurements of greater than 4 N were very rare, and often represented extreme outliers that could bias our analysis.

Mucosal Pressure Pain Detection Threshold Measurement

  The 6 anatomical sites (3 on the upper vestibule and 3 on the lower vestibule) were determined in reference to the conventional clock face. Using an imaginary clock, a total of 12 equidistant sites on the vestibule can be described circumferentially, with positions 12 and 6 corresponding to the

anterior and posterior position on the midline (in dorsal lithotomy position). Thus, sites 10, 12, and 2 are located in the upper vestibule and sites 5, 6, and 7 are located in the lower vestibule (Fig 1). Furthermore, we used anatomical landmarks in order to standardize the location of these equidistant vestibular sites between patients. Sites 2 and 10 correspond to the mid-positions on the vestibular mucosa located on an imaginary line through the uretheral meatus and parallel to the horizontal plane. Sites 2 and 10 were examined in sequence, followed by Site 12, which was located equidistant from the glan clitoris and the uretheral meatus (measured vertically). The lower 3 sites were selected to be the mirror image of the upper 3 vestibular sites.

Thus, 6 o'clock is the most posterior position on the vestibule and equidistant from 5 o'clock position on the right (subject's left) and 7 o'clock position on the left (subject's right). Following the examination of the upper vestibular sites (2, 10, and 12), the lower vestibular sites were examined in the following sequence: 5, 7, 6. In collaboration with the biomedical engineering core at the Center for Neurosensory Disorders at the University of North Carolina in Chapel Hill, we developed an algometer for assessment of pressure pain sensitivity in the vestibule (Fig 2A) and the underlying pelvic musculature (Fig 2B). For the former, we used a digital algometer (Wegner Instruments, Greenwich, CT) affixed to a disposable cotton swab and a custom-built computer interface for real time data acquisition (Fig 2A). Upon initiation of contact with the vestibular mucosa, the examiner(s) began application of pressure at an approximate rate of 1 N per second. Subjects were instructed to click a computer mouse upon the first sensation of pain, at which point the pressure was immediately terminated and automatically recorded as the ''mucosal pressure pain threshold.'' Each mucosal site was examined 3 times with an interstimulus interval (ISI) of 2 seconds. A verbal command via a computer interface announced the order of sites, number of repetitions (3 per site), and the interval between the palpations of individual sites and repetitions within each site. For example, a typical exam consisted of a verbal command to start at site 2, and after obtaining 3 threshold values (with an ISI of 2 seconds), the examiner was prompted to move to sites 10, 12, 5, 7, and 6 sequentially. Technical specifications are described in detail at the following secure website: (https://www. unc.edu/_ogmonbur/ea/wa-docs/wa-info1.html).
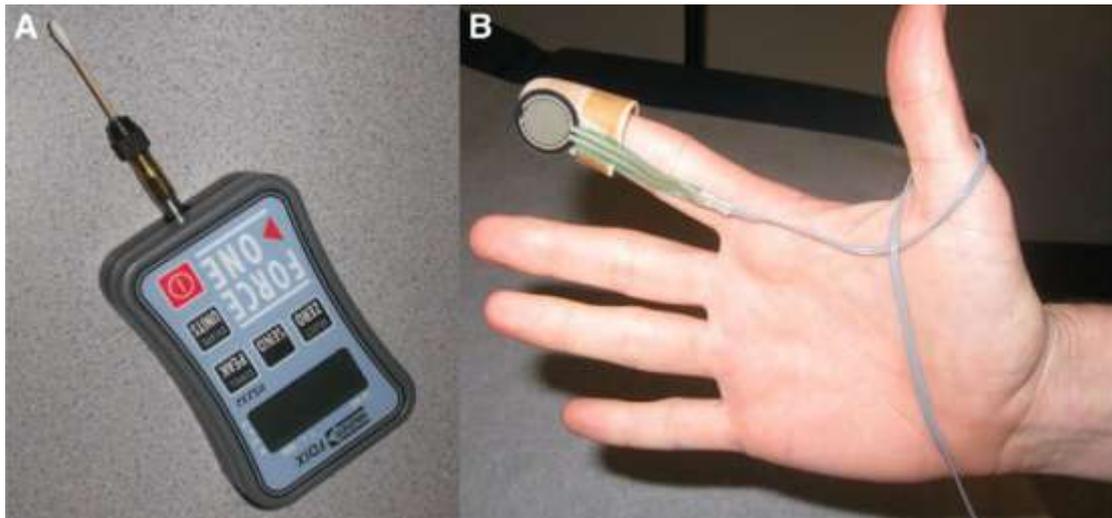
Figure 2. Mucosal and muscle algometer. Real time data acquisition was conducted via a computer interface calibrated to: (A)
Wegner Digital Algometer custom fitted with a cotton swab for the purposes of mucosal sensitivity assessment; and (B) Force Sensing
Resistor (FSR) affixed on plastic thimble to examine pelvic muscles

Muscle Pressure Pain Detection
Threshold and Tolerance Measurements

Our pelvic muscle algometer (Fig 2B) uses a pressure sensor capable of measuring a wide spectrum of force (<100 g to >10 kg or <1N to >98 N) and allows direct and isolated palpation of the levator muscles (specifically right and left puborectalis) and perineal muscle complex.A computer program similar to the one described for the sensory assessment of the mucosa allows real-time data acquisition while standardizing the rate of application, inter- and intra-stimulus intervals, and the order of examination. Technical specifications are described in detail at the following secure website: (https://www.unc. edu/_ogmonbur/ea-docs/ea-info1B.htm).

The pressure sensor is affixed to a plastic thimble that is worn over the examiner's right index finger (Fig 2B). The exam is conducted transvaginally in accordance with conventional clinical practice. However, the manner in which the examination was conducted (specifically subject/ examiner interaction, verbal cue, and anatomical landmarks) was standardized and scripted for the purposes of this protocol. The examination started with the dorsal surface of the right index finger facing the subject's right pubic rami. With a bent index finger (approximately 30–60 degrees), the examiner palpates mid-pubic rami, which is the bony landmark for

identifying the mid-segment (belly) of puborectalis muscle. After confirming the mid-position placement of the index finger, the examiner palpated the puborectalis muscle, which was immediately adjacent/behind the pubic rami (<.5 cm). Subsequently, the examiner rotated the index finger downwards (approximate length of 1–1.5 cm) parallel to the length of puborectalis, until the transition point/insertion to superficial muscles (perineal muscle complex) was palpated. This was the point of pressure application for the assessment of pain detection threshold and tolerance on the right and left puborectalis muscle.
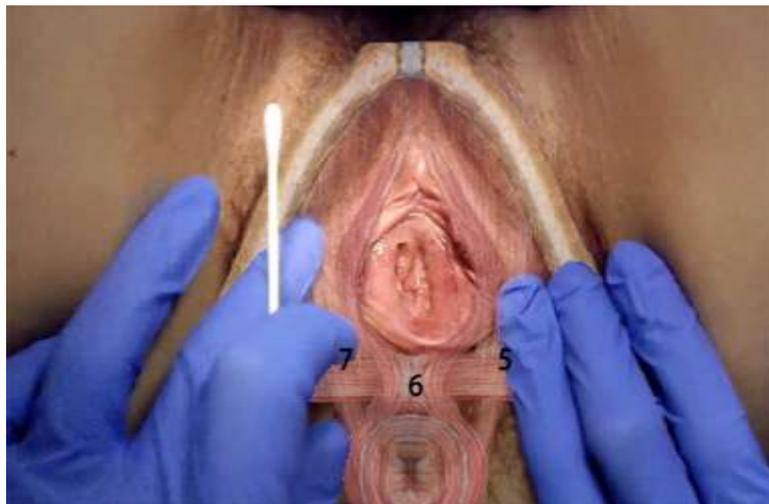

Figure 3. Anatomical location of the tested muscle sites.

Following the assessment of the right side, the examiner rotated his/her right hand in order to palpate the left puborectalis as described above. The perineal muscle complex, which is immediately underneath the 6 o'clock mucosal position, was palpated last. These muscle sites are adjacent to the respective lower vestibular mucosal sites (5, 6, and 7).

Participants were instructed to click a mouse at the first sensation of pain (pain perception threshold). However, the examiner terminated the pressure when she was no longer able to apply additional force to reach the participant's threshold. Both the amount of force and the person who terminated the application (examiner versus subject) were recorded by the computer. Following the completion of threshold testing, participants were given the option to either proceed or opt out of the tolerance testing. Tolerance testing was performed (47/51 = 92% of participants) after a 5-minute rest period. Participantsl were instructed to click the mouse when they were no longer able or willing to tolerate the pressure.

Statistical Analysis

Analysis of the mucosal/muscle algometer pain sensitivity measures presents special challenges due to the nature of the data. At least 3 replicate measurements were collected from each patient at each site, but for some patients, data was also collected on 2 different visits and/or from 2 different examiners. Also, in some cases, the examiner was unable to exert sufficient force to reach the subject's threshold and terminated the measurement procedure before the patient experienced pain, resulting in a measurement that is right-censored. Nearly 50% of the muscle tolerance measures in the control group were terminated by the examiner in this manner.

Failure to account for censoring mechanism has the potential to bias the results of any comparison between cases and controls. Thus, to test the null hypothesis of no association between these measurements and vestibulodynia case status, we fit a Cox proportional hazards model to predict each algometer measurement based on a 0/1 dummy variable for case status. We treated all measurements that were terminated by the examiner as censored. To account for repeated measures for each patient, we treated the effect of each patient as a Gaussian random effect (frailty). See Ripatti and Palmgren14 or Therneau et al15 for a more detailed description of this procedure. Each visit was also modeled as a random effect within the random effect for the subjects. To test the null hypothesis of no association between a given measure and vestibulodynia case status, we used a score test to test the null hypothesis that the coefficient for the case status dummy variable was equal to 0. The effect size of case status (and other covariates) was expressed as a hazard ratio, which is the relative likelihood that a subject will terminate the procedure at any given force level compared to the baseline risk group. In other words, if the hazard ratio for vestibulodynia case status is 2, then a given patient with vestibulodynia is twice as likely as a pain-free participant to terminate the procedure at any given force level.

Results
Demographics

Our cohort (n = 55) consisted primarily of educated, white women in their late 20s. Demographic characteristics of women with vestibulodynia did not differ significantly from the pain-free comparison group on: age (26.6 versus 27.7 years), ethnicity (80 versus 66.7% non-Hispanic white), education (77.1 versus 50% college educated), and marital status (42.9 versus 22.2% married).

Reproducibility

To establish reproducibility of the threshold measurement, the first measurement collected for each subject at each site was compared to the third measurement collected by the same experimenter at the same site (which was collected approximately 4 seconds later).

All subjects in the study (34 vestibulodynia patients and 21 healthy controls) were included in this analysis. To assess reproducibility of the pain threshold over time, the results of the first visit were compared to those of the second visit among subjects with 2 visits. Sixteen women participated in 2 different visits, including 6 pain-free and 10 women with vestibulodynia. Similarly, when 2 examiners collected measurements on the same subject at the same visit, we compared the measurements of the 2 examiners to assess interexaminer reproducibility. There were 16 participants for whom we collected measurements from 2 examiners at the same visit, including 4 painfree and 12 women with vestibulodynia. The mean difference (and associated standard errors) and the Pearson correlation coefficients between each pair of measurements are shown in Table 1. The correlation between the first and third measurement is very high for all 3 types of pain threshold measurements at all sites; all such correlations except for 1 were .8 or greater. The mean differences between the first and third measurements are small (ranging from .018 N and .151 N) for the mucosal measurements. However, the third muscle measurement (for both threshold and tolerance) was consistently lower than the first measurement (the mean differences ranged between .55 N and 1.82 N), suggesting that sensitivity increased with repetitive stimulation. The visit 1/visit 2 correlation is very high for both mucosal and muscle measurements at all examination sites: The correlations ranged from .547 to .871 for the mucosal data and from .748 to .85 for the muscle data. The mean visit 1/visit 2 differences were modest for both mucosal and muscle measurements (ranging between .008 N and .15 N for the mucosal measurements and .03 N and 1.68N for the muscle measurements). The experimenter 1/experimenter 2 correlations were lower (ranging from .107 to .577), and the mean experimenter 1/experimenter 2 differences were larger (ranging from .071 N to .43 N for the mucosal measurements and .17 N to 3.9 N for the muscle measurements). This is not surprising given the challenges of standardizing the actual conduct of the exam (see Discussion).

Group Differences in Mucosal and Muscle Pain Sensitivity **Measures**

Having established the reproducibility of the pain thresholds obtained by our instruments, we evaluated their ability to discriminate between vestibulodynia patients (cases) and pain-free controls. Fig 4 shows Kaplan-Meier plots of the set of mucosal detection thresholds for pain-free controls and patients with vestibulodynia. There is a large difference in the distribution of the mucosal perception thresholds between patients with vestibulodynia and the pain-free comparison group. In all 6 of the measured mucosal sites, women with vestibulodynia showed significantly lower pressure pain detection thresholds compared to their pain-free counterparts. This was particularly robust at the lower vestibule, corresponding to sites 5, 6, and 7. Similarly, compared to pain-free participants, patients with vestibulodynia showed lower pelvic muscle pressure

Table 1. Correlation Coefficient and Mean Difference (SE) for Mucosal and Muscle Measurements at Specified Anatomical Sites

ANATOMICAL SITES*	SITE 12	SITE2	SITE5	SITE6	SITE7	SITE 10
Mucosal threshold						
Within examiner (r)	.84 .	.82	.84 .	. 86 .	.85	.83
Mean difference (SE)	_.05 (.05)	.15 (.05)	.02 (.05)	.08 (.07)	_.08 (.04)	_.06 (.05)
Between examiner	.27	.11	.11	.58	.31	.11
	.43 (.16)	.20 (.15)	.07 (.13)	_.25 (.12)	.07 (.13)	.31 (.15)
Between visit	.66	.55	.82	.86	.87	.58
	_.08 (.07)	_.15 (.09)	_.00 (.08)	_.23 (.11)	_.11 (.07)	.02 (.08)
Muscle threshold						
Within examiner		.84	.87	.81		
		.86 (.34)	1.46 (.36)	.76 (.40)		
Between examiner		.62	.58	.85		
		_2.38 (.84)	_2.98 (.85)	_1.17 (.67)		
Between visit		.67	.67	.61		
		_.96 (.61)	_1.85 (.82)	_.03 (.70)		
Muscle tolerance						
Within examiner		.84	.86	.76		
		.951 (.35)	.545 (.39)	1.82 (.48)		
Between examiner		_.11	_.28	.20		
		_3.90 (.92)	_2.97 (1.03)	_3.37 (.85)		
Between visit		.56	.67	.45		
		_.9 (.82)	_1.68 (.84)	_.52 (.99)		

NOTE. Pearson correlation (r) with mean measured force difference (Newton) and standard error in parentheses.
*Anatomical sites are ordered with reference to clock's face (see Figs 1 and 3).

pain threshold and tolerance measurements. Fig 5 shows the corresponding Kaplan-Meier plots for the muscle threshold and tolerance measurements at each of the three muscle sites.

To formally test the null hypothesis of no difference between the mucosal detection thresholds between the 2 groups, we fit a Cox proportional hazards model to predict a subject's mucosal threshold based on an indicator variable for case status (see Methods). We also fit a second model that adjusted for several covariates:

namely, a dummy variable for set (reflecting which examiner went first and second), and an interaction term between set and case status. The effects of other covariates were not significant so they were not included in the final model. These results are summarized in Table 2.

The hazard ratio for vulvodynia case status was observed to be approximately 1.41 times the hazard ratio for controls. This difference remained basically unchange 1.29after adjustment for possible confounders. The hazard ratio varied from site to site. The ratio was largest at sites 5, 6, and 7 and smallest at site 12(Table 2). This suggests that sites 5, 6, and 7 may be the most informative sites for discriminating between women with vestibulodynia and pain-free controls. Interestingly, when both examiners collected a set of measurements at the same visit, pain-free participants had a lower hazard ratio during the second set of measurements whereas vestibulodynia patients had a slightly higher hazard ratio. In other words, women without pain were less likely to terminate the procedure at a given force level during the second set than during the first set of measurements; no difference was observed in vestibulodynia patients. A similar approach tested the null hypothesis that the vestibulodynia patients and pain-free comparison group had equal hazard ratios for the muscle measurements (both pain threshold and tolerance). Once again, we fit 2 Cox proportional hazards model to predict a subject's muscle threshold based on case status as vestibulodynia (Table 2).We also fit a second set of Cox models that adjusted for covariates, but none of the coefficients of the covariates were significantly different from 0, so these results are not reported in Table 2.

We observed a significant association between vestibulodynia case status and muscle threshold measures when measurements from all 3 sites were combined (HR = 2.9, P = .047). Examining each site separately, there is a significant association between vestibulodynia case status and measurements at 2 of the 3 measured sites, namely sites 5 (HR = 4.89, P = .003) and 7 (HR = 2.23, P = .0001). There is a similar pattern of association with respect to the muscle tolerance measures: the measurements were associated with vestibulodynia case status across all 3 sites combined (HR

= 3.09, P = .005), as well as at sites 5 (HR = 6.31, P = .034) and 6 (HR = 4.83, P = .049) individually.

Clinical Relevance (Validity) of the Mucosal Pain Threshold as it Relates to Subjective Pain Reports For the purpose of validating our instrument, the association between a patient's self-reported pain during intercourse and the patient's mucosal and muscle measurements was evaluated. Each participant was asked to rate her average pain during intercourse on a scale from 0 to 100. To assess the association between mucosal threshold and intercourse-related pain, we fit a Cox proportional hazards model to predict each woman's mucosal threshold based on her intercourse-related pain level, as described previously. Two models were fit: 1 with only intercourse-related pain as a predictor, and 1 with intercourse-related pain adjusted for examiner and measurement set. We observed a statistically significant association between intercourse-related pain and mucosal threshold. The strength of this association varied greatly from site to site. The association was strongest at sites 6 (P = .009), 5 (P = .004), and 7 (P = .028) and much weaker and nonsignificant at the upper vestibular sites 2 (P = .75), 10 (P = .32), and 12 (P = .38).
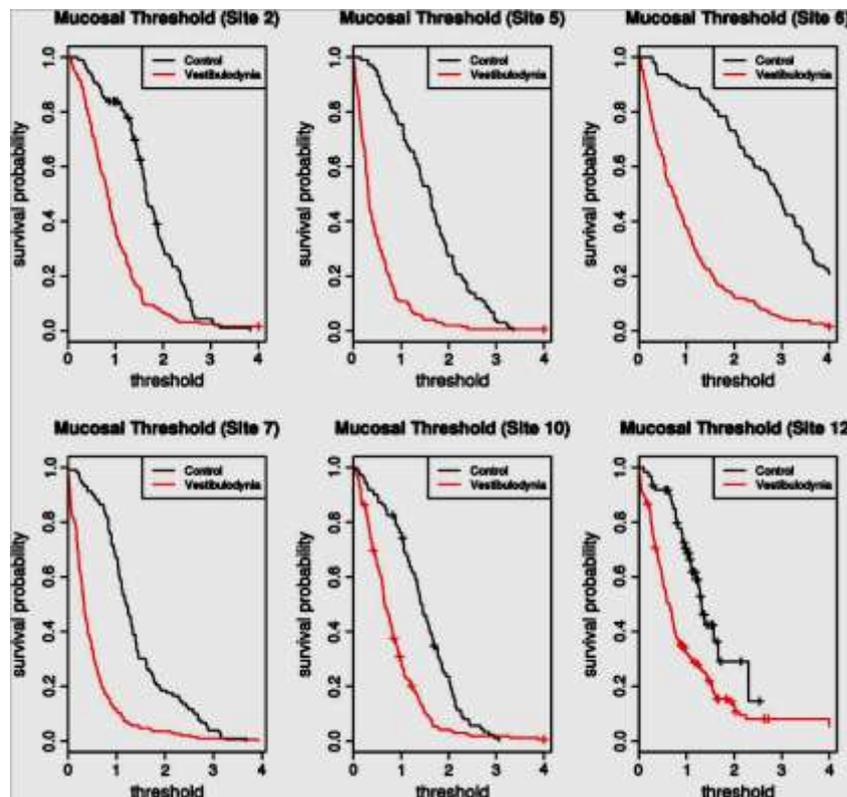

Figure 4. Kaplan-Meier plots for vulvar mucosal pressure pain detection threshold. This figure shows the estimated ''survival'' probabilities (ie, the probability of terminating the exam prior to a given amount of force) for controls and vestibulodynia patients. The

termination pressure threshold (N) is shown on the x-axis, and the probability of surviving to given threshold is shown on the y-axis. For example, the probability that a control will not terminate the examination at site 2 prior to receiving 1 N of force is approximately .8. Vestibulodynia patients were more likely than controls to terminate the force at a given pressure at all 6 mucosal sites. The difference was more robust (as noted by a larger area between the 2 curves) at lower vestibule sites

For example,

at mucosal site 6, a respondent with a mild pain report (VAS = 30) has a hazard ratio of 2.46 compared to a respondent with no pain (VAS = 0), and a respondent with a moderate pain report (VAS = 60) has a hazard ratio of 6.05 compared to a respondent with no pain. There was no statistically significant association between intercourse-related pain and the muscle threshold or muscle tolerance measurements.

Case-Control Classification

We have shown that vestibulodynia patients have lower average mucosal thresholds than pain-free controls. This raises the question of whether this result can be used to more accurately diagnose patients with vestibulodynia in a clinical setting. We attempted to answer this question by using a simple procedure to develop a classifier to discriminate between vestibulodynia patients and pain-free controls. Since we previously observed that mucosal measurements at sites 5, 6, and 7 are most strongly associated with intercourse-related pain and case status, we built our classifier using only measurements from these 3 sites. For each subject, we averaged her first 3 mucosal threshold measurements at sites 5, 6, and 7. We then averaged these 3 averages to construct our final classifier. We only considered measurements collected in the first measurement set of the first visit for each patient. Also, we only considered measurements collected from examiner 1, since we observed considerable variation in measurements collected by different examiners, and examiner 1 collected more data than the other examiner.

The ROC curve for predicting case status based on this classifier is shown in Fig 6. Note that we can obtain both high sensitivity and specificity by choosing an appropriate cutoff for discriminating between cases and controls. In particular, if we classify subjects with an average threshold force of less than 1 N as cases and subjects with an average threshold of greater than 1 N as controls, then we obtain a sensitivity of 85.7% and a specificity of 82.4%. This suggests that our methodology has the potential to help diagnose patients with provoked vestibulodynia.
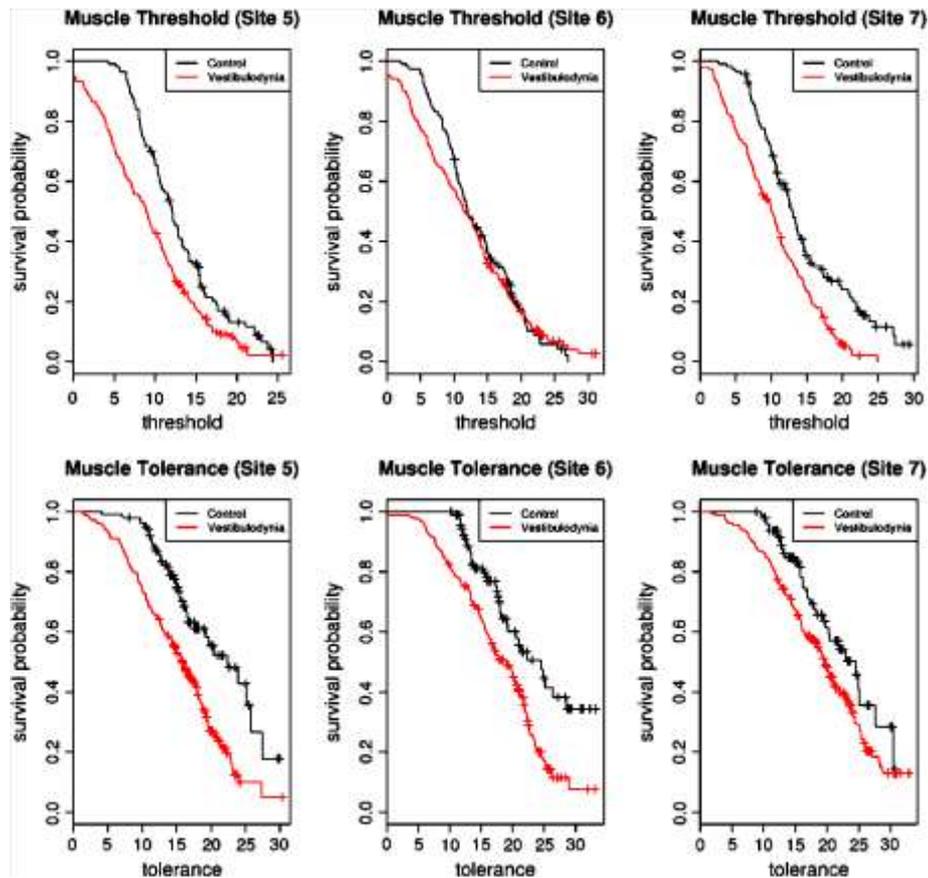
Figure 5. Kaplan-Meier plots for pelvic muscle pressure pain threshold and tolerance. This figure shows the estimated "survival" probabilities (ie, the probability of terminating the exam prior to a given amount of force) for controls and vestibulodynia patients.

The termination pressure threshold/tolerance (N) is shown on the x-axis, and the probability of surviving to given threshold is shown on the y-axis. For example, the probability that a control will not terminate the muscle threshold examination at site 5 prior to receiving

10N of force is approximately .7. Compared to pain-free participants, patients with vestibulodynia show lower pelvic muscle pressure

pain threshold (upper panel) and tolerance (lower panel) measurements.

Discussion

Provision for a reliable and valid system of pain assessment is a necessary first step toward describing the prevalence, natural history, and treatment outcomes in vestibulodynia. This is particularly important in light of significant interindividual variability of sensory thresholds, quintessential to human perception in both diseased and nondiseased states.4

In this study, we demonstrate that standardized methodologies based on psychophysical approaches provide reliable indices of pelvic muscle and

mucosal pressure pain sensitivity, and that select indices relate to patients' clinical pain reports.

The usefulness of examination methods requires both reliability (reproducibility) and validity (accuracy) of the outcome measures. The reliability score is the conceptual sum of 4 dynamic domains: intra- and inter-examiner reliability, stability of phenomenon being measured over time, and reliability of the participant's pain report.3 To date, limited work has been conducted on these domains

Table 2. Hazard Ratio for Termination of Pressure Application at Specified Anatomical Sites

....

Anatomical Sites	Site 12	Site2	Site5	Site6	Site7	Site 10
Mucosal threshold	5.59 (.03)	7.68*	27.57*	44.04*	14.11*	6.72 (.0006)
Muscle threshold			4.89 (.003)	1.80 (.29)	9.29*	
Muscle tolerance			6.31 (.03)	4.83 (.05)	3.02 (.09)	

....

NOTE. Hazard Ratio is calculated with pain-free controls as the reference group. Numerical value under each anatomical site represents the relative likelihood of terminating the pressure application by clicking the computer mouse. Since reference group consists of pain-free controls, numerical values in individual cell represent the magnitude of increase in likelihood of pressure termination by vestibulodynia patient at a given anatomical site.
*Unless otherwise indicated in parentheses, the P-value for relative likelihood of early pressure termination by vestibuloydnia (compared to pain free controls) was less than .0001.

with respect to the clinical measurements of pelvic muscle and mucosal sensitivity. The methods we employed enabled investigation of the joint and independent contribution of these domains (eg, subject versus experimenter versus time) on the reliability of measures of mucosal (threshold) and muscle (threshold and tolerance) pressure pain sensitivity obtained by 2 examiners as much as 2 weeks apart.

The measures were found to be temporally stable, as demonstrated by the high correlation between values obtained from subsequent repetitions in the same visit and between values obtained 2 weeks apart. The correlation between subsequent repetitions in a single exam was excellent (r$.8), indicating high reliability within the course of an exam. While the between-visit correlations for the muscle and mucosal measurements were within conventional norms (r $ .6),3 the mucosal correlations were not consistently over .6, being as low as .54 at some sites)Table 1). However, upon closer inspection, the relatively lower between-visit correlations (r = .5–.6) were consistently associated with upper vestibular sites (Fig 1) whereas between-

visit correlations for the lower 3 mucosal sites were on the order of .8 and higher, similar in magnitude to those of the muscle measurements. Meshab et al9 observed a similar discrepancy, which can be attributed to anatomical challenges. Because of immediate proximity to the pubic bone, urethralmeatus, and glan clitoris all within a radius of 1cmfromthe examined sites slight variation in the angle of cotton swab application likely leads to variation in the measured threshold (Fig 1).
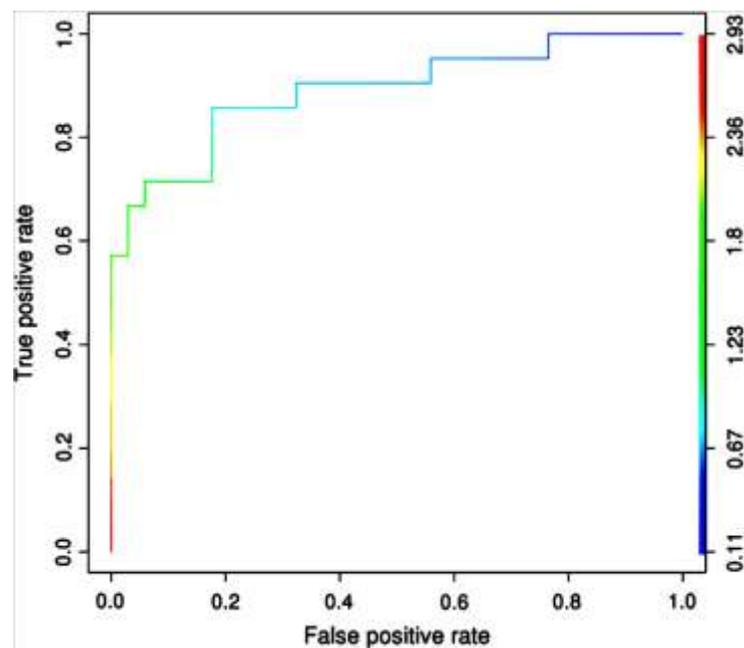

Figure 6. Receiver operator curve for predicting case status.
This figure shows the receiver operating characteristics (ROC)curve demonstrating the predictive capability of the average of the 3 lower vestibular mucosal sites to predict vestibulodyniacase status. If we classify subjects with an average threshold of less than 1 N as cases and subjects with an average threshold of greater than 1 N as controls, then we obtain a sensitivity of 85.7% and a specificity of 82.4%.

Interexaminer Reliability of Pain Measures

We observed low interexaminer correlations, particularly for mucosal sites. Only 1 mucosal site (6 o'clock) reached an acceptable interexaminer correlation (r = .577). The 6 o'clock site was the only site tested for which the perpendicular application of pressure using a cotton swab could be conducted under direct visualization with a high degree of manual accuracy. On average the interexaminer correlations for muscle were higher than those for the mucosa; the highest correlation for muscle was seen at the 7 o'clock position. This is attributed to the examiner's ability to apply pressure in an anatomically congruent manner (right-handed examination of

the patient's right pelvic muscle). The examiner used the right index finger to sequentially examine the right, left, and mid position. Because muscle tolerance measures are primarily a reflection of examiner's maximal force, lower correlation between the 2 examiners was observed for pain tolerance than for pain detection thresholds. The lower interexaminer reliability, particularly for mucosal assessment, highlights the importance of training and calibrating the examiners.

Validity of Pain Measures

As eloquently discussed by Dworkin and LeResche,3 ''Clinical measure may be extremely reliable or repeatable, but if a measure does not accurately portray the entity examined, then it provides no useful information for diagnosis.'' The construct of validity requires either a ''gold standard,'' in the form of an ''accepted truth'' based on empirical data, or a validated instrument that can be compared with the new approach. In our analysis, we used participants' self-reported pain during intercourse, a clinically meaningful construct, as a proxy for a ''gold standard.'' Of the tested vestibular sites (6), only 3 lower vestibular sites (5, 6, and 7) were associated with intercourse-related pain. Measures of muscle pain sensitivity were not associated with participants' selfreported pain during intercourse, despite being associated with case status as described below.

Usefulness of Pain Measures in Vestibulodynia Identification

Compared to pain-free controls, vestibulodynia cases reported lower mucosal and muscle thresholds, suggesting higher sensitivity. This was particularly evident for the lower vestibular sites (5, 6, and 7), suggesting that these sites robustly discriminated cases from controls. ROC analysis further demonstrated that an average threshold of less than 1N provided a cut-off for discriminating between cases and controls with a sensitivity of 85.7% and a specificity of 82.4%. This suggests that our methodology has the potential to help diagnose patients with provoked vestibulodynia.
With respect to the muscle threshold measures, only 2 (sites 5 and 7) of the 3 tested sites were significantly associated with case status. The group differences at sites 5 and 7 were notably smaller for muscle tolerance measures.
In fact, the hazard ratio for muscle pain tolerance at site 7 did not reach statistical significance (P = .09), despite being remarkably robust (P < .0001) for the muscle pain threshold measure.

Limitations

The extent to which increased sensitivity of the mucosa affected the pain threshold measured for the pelvic musculature remains unclear. Because of the thinness of mucosa 1>)mm), our methodology could have led to termination of pressure during the muscle examination due to the experience of mucosal pain secondary to soft tissue traction. Because gentle pressure (enough to indent the mucosa approximately 1 N) was sufficient for discriminating cases from controls, the higher force used for the muscle examination in cases would be expected to exceed the mucosal pain threshold of these patients.

   Regardless of this possibility, our results suggest that transvaginal assessment of muscle without irritating the overlying mucosa is possible. First, all of our participants with vestibulodynia and mucosal pain thresholds of 1 to 4N completed the muscle threshold and tolerance assessment where pressures as high as 30N were sometimes delivered to reach the pain threshold. Second, the 6 o'clock mucosal site was one of the most sensitive sites where pressures as low as 1 N led to termination by subjects with vestibulodynia. However, an association between case status and muscle pressure pain threshold (P = .29) and tolerance (P = .06) at this position was not observed, suggesting that the 2 methodologies (mucosal versus muscle) captured different, though arguably related constructs.
Taken together, these observations highlight the feasibility of transvaginal assessment of pelvic muscles without irritating the overlying vulvar mucosa.

   We strived to standardize various facets of subject/examiner interaction and clinical assessment. The research coordinator and the examiners followed a scripted protocol for interacting with study participants. The coordinator explained the different aspects of the exam and prompted the transition from 1 segment to another.
However, training and calibration of examiners was quite limited in scope. In fact, it was during the conduct of the present study that we identified elements of the exam that needed to be standardized. For example, it was only after we began our study that we noticed that a slight deviation of the wrist during mucosal assessment can change the intended location and angle of application and that the anatomical distribution of thevulvar mucosa and mucocutaneous junction differs among subgroups of women.
The extent to which increased pain sensitivity of the pelvic musculature affected the pain threshold measured at the vulvarmucosa remains unclear.

The etiology ofmuscle involvement in vestibulodynia is commonly viewed as reactive and secondary to overlying mucosal sensitivity.

However, wecannot compare the relativemerit and directionality of the relationship between muscle andmucosal sensitivity based on the magnitude of HR ratios alone.

Unlike mucosa, which has a high fidelity input, the musculoskeletal input is characterized by diffuse, poorly localized, low fidelity input.18 Therefore, it is plausible that muscle may be the primary source of pain and functional aberration in some women, with referral to the overlying mucosa. In other cases, the mucosal input is primary with secondary muscle involvement. Our study was not designed to address this question, but this methodology will enable the design and implementation of future studies in order to evaluate the directionality of this relationship (P01 NS045685, currently underway).

To our knowledge, the work reported in this paper is the first research aimed at operationalizing a standardized and validated clinical pain examination of the pelvic region. Research into the pathophysiology of chronic pelvic pain in women is in its infancy and is in urgent need of collaborative effortsamongcliniciansandscientists across a broad range of expertise. Success will be dependent on standardized approaches, and we hope that they are adopted and further developed by other investigators.

Acknowledgments

References

.1Bornstein J, Maman M, Abramovici H: ''Primary'' versus "secondary" vulvar vestibulitis: One disease, two variants.Am J Obstet Gynecol 184:28-31, 2001

.2Diatchenko L, Nackley AG, Slade GD, Fillingim RB, Maixner W: Idiopathic pain disorders–pathways of vulnerability.Pain 123:226-230, 2006

.3Dworkin SF, LeResche L: Research diagnostic criteria for temporomandibular disorders: Review, criteria, examinations and specifications, critique. J Craniomandib Disord 6: 1992 ,355-301

.4Fagius J, Wahren LK: Variability of sensory threshold determination in clinical use. J Neurol Sci 51:11-27, 1981

5. Friedrich EG Jr: Vulvar vestibulitis syndrome. J Reprod Med 32:110-114, 1987

6. Gracely RH: Evaluation of multi-dimensional pain scales.Pain 48:297-300, 1992

7. Harlow BL, Stewart EG: A population-based assessment of chronic unexplained vulvar pain: Have we underestimated the prevalence of vulvodynia? J Am Med Womens Assoc 58:82-88, 2003

8. Lund JP, Widmer CG: Evaluation of the use of surface electromyography in the diagnosis, documentation, and treatment of dental patients. J Craniomandib Disord 3:125-137, 1989

9. Masheb RM, Lozano C, Richman S, Minkin MJ, Kerns RD: On the reliability and validity of physician ratings for vulvodynia and the discriminant validity of its subtypes. Pain Med 5:349-358, 2004

10. Pukall CF, Binik YM, Khalife S: A new instrument for pain assessment in vulvar vestibulitis syndrome. J Sex Marital Ther 30:69-78, 2004

11. Pukall CF, Binik YM, Khalife S, Amsel R, Abbott FV: Vestibular tactile and pain thresholds in women with vulvar vestibulitis syndrome. Pain 96:163-175, 2002

12. Pukall CF, Payne KA, Binik YM, Khalife S: Pain measurement in vulvodynia. J Sex Marital Ther 29(Suppl 1):111-120, 2003

13. Reed BD, Crawford S, Couper M, Cave C, Haefner HK: Pain at the vulvar vestibule: Aweb-based survey. J Low Genit Tract Dis 8:48-57, 2004

14. Ripatti S, Palmgren J: Estimation of multivariate frailty models using penalized partial likelihood. Biometrics 56: 1016-1022, 2000

15. Therneau TM, Grambsch PM, Pankratz VS: Penalized survival models and frailty. J Comput Graph Statis 12:156-175, 2003

16. Tu FF, As-Sanie S, Steege JF: Musculoskeletal causes of chronic pelvic pain: A systematic review of diagnosis: Part I. Obstet Gynecol Surv 60:379-385, 2005

17. Tu FF, Fitzgerald CM, Kuiken T, Farrell T, Norman Harden R: Vaginal pressure-pain thresholds: Initial validation and reliability assessment in healthy women. Clin J Pain 24:45-50, 2008

18. Walder RY, Sluka KA: Bilateral, long lasting hyperalgesia due to repeated excitation of muscle nociceptors: A role of ASIC3, in Graven-Nielsen T, Arendt-Nielsen L, Mense S (eds): Fundamentals of Musculosketal Pain. Seattle, WA, International Association for the Study of Pain, 2008, pp. 75-86

19. Zolnoun D, Hartmann K, Lamvu G, As-Sanie S, MaixnerW, Steege J: A conceptual model for the pathophysiology of vulvar vestibulitis syndrome. Obstet Gynecol Surv 61:395-401, 2006. quiz 423

20. Zolnoun D, Park EM, Moore CG, Liebert CA, Tu FF, As- Sanie S: Somatization and psychological distress among women with vulvar vestibulitis syndrome. Int J Gynaecol Obstet 103:38-43, 2008

....