Neurourology and Urodynamics 26:858–863 (2007)

Test-Retest Reliability of an Instrumented Speculum for Measuring Vaginal Closure Force

J.M. Miller,[1,4*†] J.A. Ashton-Miller,[2‡] D. Perruchini,[3] and J.O.L. DeLancey[4¶]
[1]School of Nursing, University of Michigan, Ann Arbor, Michigan
[2]Department of Mechanical Engineering, Department of Biomedical Engineering, Institute of Gerontology, University of Michigan, Ann Arbor, Michigan
[3]Division of Gynecology, Department of Obstetrics and Gynecology, University Hospital of Zurich, Zurich, Switzerland
[4]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, Michigan

Aims: The study aimed to: a) determine reliability of an instrumented speculum designed for measuring intravaginal closure pressure, and b) compare findings with a comparable device reported in the literature. The goal of these new devices is to reduce subjectivity, improve precision, and acknowledge reliability issues in quantifying levator ani closure force acting on the vagina. **Methods:** The instrumented speculum consisted of two parallel aluminum bills, similar in size to a Peterson speculum. Strain gages located near the root of each bill measure the magnitude of force exerted in the distal vagina. A contraction of the "U-shaped" levator ani muscle closes the levator hiatus with resultant reaction force measured by the speculum in the mid-sagittal plane. We tested the device in twelve nulliparous women making repeated measures within and across 3 different visit days. All measures were made by the same investigator. **Results:** Same day measures were repeatable within ±3.8 N by the third visit, with lesser repeatability on the 1st and 2nd visit days. Across days, repeatability was improved by Visits 2 and 3 with a coefficient of repeatability between those days of ±5.5 N. Better repeatability was obtained using averaged scores rather than 'best effort' ; but average scores can underestimate best effort. **Conclusion:** Reasonable within-visit repeatability was found. Across-visit repeatability is consistent with the known difficulty that women have in maximally isolating and activating their levator ani muscles. The results corroborate the repeatability results of Dumoulin et al. [2004] using a similar type of dynamometer. *Neurourol. Urodynam. 26:858–863, 2007.* © 2007 Wiley-Liss, Inc.

Key words: Kegel strength; levator ani; pelvic floor muscles

INTRODUCTION

Measurement of maximum voluntary pelvic muscle strength is an important part of the assessment of the female pelvic floor muscle function. A primary component of the pelvic floor musculature is the levator ani muscles. The levator ani muscles play a critical role in supporting the bladder neck, important to maintenance of continence.[1,2] This muscle set also is responsible for automatically adjusting to the loads placed upon them[3,4] so as to maintain constriction of the urogenital hiatus and thereby relieving the connective tissue supports of the constant stress that might lead to their failure and consequent prolapse.[5] In either instance, measurement of the muscle strength is critical in quantifying this important element of the pelvic floor support system.

A number of methods have been used to measure maximum voluntary levator ani muscle strength in vivo.[6] Subjective scoring systems involving digital palpation have been used to quantify the character of contraction; for example, the digital measure[7] and the Oxford scale.[8]

Objective measures include the measurement of intravaginal closure pressure via balloon perineometers or manometry to quantify maximum voluntary intravaginal squeeze pressure.[9–13] However, the compliance of balloon-type devices allows the pelvic floor muscles to shorten as they develop increasing force, because there is nothing to prevent a reduction in balloon radius and its elongation along the vagina, even if it operates isovolumically. Since muscle force decreases with shortening velocity, and also generally with decreasing length, readings made with the balloon are susceptible to systematic artifacts due to the unknown pelvic floor muscle length and shortening velocity. Intravaginal squeeze pressure has been found to vary along vaginal length, from intra-abdominal pressure in the proximal vagina to the greatest squeeze pressures (50 ± 6 mm Hg) being recorded in the distal 4–5 cm of the vagina.[14] The correlation between subjective scoring and vaginal squeeze pressure has been studied; however, digital palpation was not found to be a reliable measure of maximum pelvic muscle strength.[15,13]

Levator muscle recruitment have been quantified in various activities using surface or needle myoelectric activity measurements,[16] but the difficulty of knowing exactly which structure is being recorded from, along with the possibility of cross-talk from the obturator muscles, can complicate interpretation of results. In addition, the problem of localized atrophy in parts of the pubic portion of the levator[17] means that no myoelectric signals can be expected to be recorded from muscle that is frankly missing.

Magnetic resonance imaging now allows objective assessment of levator ani muscles morphology in vivo,[18,19] as well as their cross- sectional area taken normal to the direction of the fibers.[20] The latter parameter can be used to estimate the maximum contractile force they can generate along a line-of-action of the muscle in the manner of Ikai and Fukunaga.[21]

No conflict of interest reported by the author(s).
Linda Brubaker led the review process.
Work performed at the University of Michigan School of Nursing and University of Michigan Medical Center.
†Research Assistant Professor and Assistant Research Scientist.
‡Distinguished Research Professor.
¶Norman F. Miller Professor of Gynecology; also Director of Pelvic Floor Research Group and Fellowship in Female Pelvic Medicine and Reconstructive Surgery.
*Correspondence to: J.M. Miller, 400 N. Ingalls, Division 2, Ann Arbor, MI 48109-0482. E-mail: janismm@umich.edu
Received 26 September 2006; Accepted 3 January 2007
Published online 13 March 2007 in Wiley InterScience (www.interscience.wiley.com)
DOI 10.1002/nau.20407

Recent imaging research has demonstrated levator damage at vaginal birth[17,22,23] and demonstrated the association with damage and stress incontinence and pelvic organ prolapse.[24] These morphological studies indicate the importance of alterations in levator ani in pelvic floor dysfunction but do not yield information about levator muscle function in vivo.

Over the past years we have developed an instrumented vaginal speculum that measures vaginal closure force in the sagittal plane during maximal volitional contraction of the levator ani.[25] We reported data obtained with this instrumentation on longitudinal assessments of pelvic floor strengths in pregnant women.[26] A similar device was developed independently by Dumoulin et al.[27] and recently tested for its repeatability[28] and validity.[29,30] The correlation between vaginal closure force and digital palpation measurements has been found to be 0.73 in continent women,[29] though the issue of poor reliability with the subjective digital measure again poses problems. Likewise, a similar device has been used to measure how vaginal closure force in the frontal plane changes with vaginal diameter by Verelst and Leivseth.[31] Taken together, it is important to address the reliability issues of these new devices.

The purpose of this article is to report test-retest reliability for our instrumented speculum in measuring maximum vaginal closure force in healthy women and compare these results data with those reported on the instrument developed by Dumoulin et al.[28]

MATERIALS AND METHODS

Instrument

A technical description of the two-billed instrumented speculum is given in Ashton-Miller et al.[25] Two parallel aluminum bills, similar in size and shape to a pediatric speculum, are cantilevered and held in a fixed, substantially parallel relationship to one another. The bills had an insertion length of up to 7 cm and a width of 2.5 cm (Fig. 1). Each bill has three moisture-proofed strain gages located in a narrowed region near its root. The strain gages are connected to Wheatstone bridge circuits, amplifier, 12-bit analog-to-digital converter, and laptop. Two of the gages are wired as in a differential shear beam[32] such that, once calibrated using static weights from 0 N to 20 N, will measure the magnitude of an equivalent force acting normal to the plane of the speculum so as to approximate the bills toward one another.

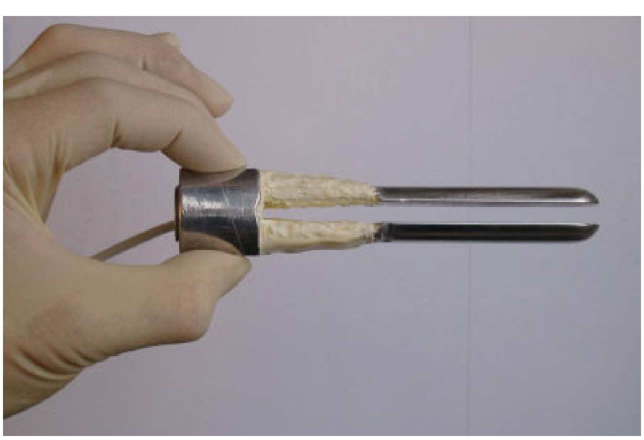

Fig. 1. Photograph showing the device and how it was held between the tips of the thumb and first two fingers for the measurements. The strain gages are covered with a white silicone sealant to protect them from moisture. The vertical line in the background illustrates the 5 cm depth to which the bills were inserted. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

In clinical use the speculum is first covered by a disposable condom which is then covered with lubricant (K-Y Gel^{TM}) and introduced gently into the vagina along the distal axis of the vagina. The base of the speculum is held between the fingertips and inserted to a depth of 5 cm (Fig. 1). The ventral bill rests against the anterior vaginal wall and ventral movement is resisted by compression of the urethra against the pubic symphysis. The posterior bill is pulled toward the anterior bill in the distal region of the vagina by the contraction of a U-shaped muscular "loop" formed bilaterally by the pubovisceral muscle, the puborectalis muscle, the ventral-most fibers of the iliococcygeus, and the bulbocavernosus muscle (the latter contributing little force). A three-dimensional illustration of these muscles reconstructed from MR scans may be found in Lien et al.[33] Volitional, or involuntary, contraction of these muscles applies a ventrally directed net force to the posterior bill in the mid-sagittal plane, which in turn pushes the ventral aspect of the ventral bill anteriorly against the dorsal aspect of the urethra compressing it against the pubic bone. A reaction force then acts in the mid-sagittal plane on the ventral bill from the pubic symphysis in a posterior-cephalad direction, normal to the longitudinal axis of the speculum bills. The instrumented speculum was calibrated using known five static weights at two locations on the bills separated by 25 mm and found to have a sensitivity of 0.401 V/N. The coefficient of determination of the linear regression line relating output voltage to input force was 0.99. The thermal drift rate in the output of the force transducer averaged +0.14 N/min when the distal 5 cm of both the upper surface of the superior bill and the lower surface of the lower bill initially at ambient room temperature, contacted tissue at body temperature for 5 min.

Subjects

Twelve continent nulliparous women, mean age 25.4, range 21–44, years volunteered. Exclusion criteria were history of urinary incontinence, neurological disease or pain on vaginal exam. Written, informed consent was obtained from each subject, and the University of Michigan Institutional Review Board approved the study.

Procedures

Prior to insertion, the speculum was held horizontally in air while the transducer output was balanced to read zero. Preparatory to the measures, each woman was positioned comfortably in the lithotomy posture and an examiner experienced in pelvic muscle assessment provided brief (less than 5 min) instruction in pelvic muscle contraction. Digital palpation was used to provide feedback on technique, and suggestions to improve technique were provided, with opportunity for repeat practice within the 5-min instructional period. Instrumented speculum measures were then taken during maximal voluntary contraction (MVC). All measures were obtained by the same investigator (DP), who had conducted these same measures in well over 50 women. The investigator was not blinded. However the measures obtained earlier were purposely not reviewed prior to subsequent data collection. We performed multiple examinations on the same day, so it would have been quite difficult for the examiner to be able to accurately remember numbers from prior visits.

Measures of maximum volitional vaginal closure force were obtained from each woman on three different visits, with approximately 1 week between each of the three sessions. On each visit women were coached to perform two measures of MVCs, which were obtained with at least a 30-sec intervening rest period. Aside from this brief instruction during the visit time, no additional verbal or written instruction was provided, and there was no instruction to practice at home.

Data Analyses

The vaginal closure force during an MVC was calculated as the difference in the peak force reading measured during MVC minus the force measured in air at the beginning of the study. We analyzed the test-retest data using an analysis strategy suggested by Bland and Altman.[34] Calculations included (1) descriptive statistics, (2) the mean difference between two measures ($\bar{d}$), (3) the standard deviation of the difference, and (4) the repeatability coefficient (± 2 times the standard deviation of the difference). Scatter diagrams were plotted of the differences against zero to illustrate the distribution of results and to search for potential bias. We chose this statistical analysis over alternative possibilities based on the arguments presented by Bland and Altman.[35] The repeatability coefficient offers the advantage of remaining in the original units of measurement, which subsequently allows clinical judgments about adequacy. It is interpreted as 95% of the sample demonstrating a difference (from one measure to another) that falls within these limits. Ultimately, the clinician decides if the outer limits are tolerable for the specific purposes of application. Ideally, the limits should hover close to zero.

RESULTS

Within-Visit Repeatability of Maximal Voluntary Contraction Force (Tables I and II)

Raw data are shown in Table I, with results portrayed for each individual at clinic visits 1, 2, and 3, with MVC repeated at each visit day. The mean group difference of measures obtained on 1st and 2nd efforts was close to zero on all three different visit days (top half of Table II). The best repeatability of within-visit measures during MVC was demonstrated on the third visit. On the third visit, the group average (range) of the two measures obtained during MVC was 6.2 (2.1–12.7) Newtons (N). The coefficient of within-visit repeatability on this best of the three days was ± 3.8 N. In this small sample size, a statistically, and likely clinically, insignificant bias in the direction of the first measure being higher than the second measure is indicated as a positive group mean difference of 0.6 N.

To further explore the data, plots of the differences in first and second measures by individual were plotted against zero for each visit. The distribution was examined and bias in the order of the measurements can be seen on all three visits, as the points are not distributed evenly above and below the zero line (Fig. 2). On Visit 1, the bias was in the direction of the second effort producing the highest intravaginal closure force (a negative difference score seen when plotted against zero). On Visits 2 and 3, the bias was in the direction of the first effort producing the highest intravaginal closure force (a positive difference score seen when plotted against zero). However, in all three visits, the group mean bias was small (about 0.5 N).

Across-Visit Repeatability of Maximal Voluntary Contraction (Tables I and II)

Across-visit repeatability was analyzed in two ways; first by using averaged scores of the two measures obtained at each discrete visit and second by using the better of the two measures obtained at each visit. We did so to investigate whether using averaged scores improves repeatability, but may underestimate actual intravaginal closure force; whereas using best effort scores may improve precision in measuring a woman's true volitional muscle contraction ability, but with lesser repeatability. As expected, the data demonstrated better repeatability when using averaged scores rather than best effort; and the data demonstrated higher group mean scores when calculated using best effort rather than averaged scores (please see lower half of Table II).

In the analysis of averaged scores, repeatability was better for comparisons made between Visits 2 and 3, rather than between Visit 1 and Visit 2 or 3. In this sample of women who received rather brief training, the group average (range) of the measures of MVC computed for characterizing the data from

TABLE I. Instrumented Speculum Measurements of Vaginal Closure Force (in N) for the First (MVC1) and Second (MVC2) Maximal Voluntary Contractions at Each Visit

	Visit 1		Visit 2		Visit 3	
Subject	MVC 1 (N)	MVC 2 (N)	MVC 1 (N)	MVC 2 (N)	MVC 1 (N)	MVC 2 (N)
1	2.5	2.6			2.5	1.7
2	2.6	2.9			3.6	5.2
3			3.7	2.2	1.1	4.1
4			4.8	3.3	8.3	6.1
5	0.8	3.2	4.1	7.3	5.3	7.3
6	3.3	3.6	5.1	3.7	7.5	7.4
7	6.1	9.9	3.5	2.4	4.0	4.0
8	6.5	7.8	8.6	10.1	9.2	7.2
9	2.2	2.1	10.2	8.5	5.8	2.1
10	5.0	0.8	8.5	4.7	13.0	12.5
11	8.4	9.0	8.8	10.6	8.2	8.6
12	8.8	10.3	10.1	10.3	7.8	7.0

*Missing items are due to participant inability to attend Visit (ID = 1) and computer failure during 1 day of data collection (ID = 2, 3, 4).

TABLE II. Results of Within- and Between-Visit Repeatability of Intravaginal Closure Force at Maximal Voluntary Contraction of the Levator Ani Muscles

MVC repeatability	Mean of MVC values (N)	Range (N)	Mean difference (N)	SD_{diff} (N)	Coefficient of repeatability
Within-visit comparisons of two trials					
Visit 1					
MVC 1 versus MVC 2	4.9	2.0–9.5	−0.6	2.1	±4.2
Visit 2					
MVC 1 versus MVC 2	6.5	3.0–10.2	0.2	2.1	±4.1
Visit 3					
MVC 1 versus MVC 2	6.2	2.1–12.7	0.6	1.9	±**3.8**
Between-visit comparisons based on average of two trials within visit					
Visit 1 versus Visit 2 using average	5.3	2.5–9.8	−1.8	3.5	7.0
Visit 1 versus Visit 3 using average	5.6	2.3–8.5	−2.1	3.9	7.7
Visit 2 versus Visit 3 using average	6.1	2.1–9.7	−0.3	3.3	**6.6**
Between-visit comparisons based on best of two trials within visit					
Visit 1 versus Visit 2 using best	6.0	2.6–10.3	−1.8	4.1	8.2
Visit 1 versus Visit 3 using best	6.3	2.5–9.0	−1.4	3.9	7.8
Visit 2 versus Visit 3 using best	6.9	2.5–10.8	−0.1	2.8	**5.5**

Coefficient of repeatability indicates that 95% of this sample produced repeated measures that fell within the upper and lower limits of the coefficient. Results in bold show the best repeatability.

Visit 2 and 3 was 6.9 (2.5–10.8) N. The coefficient of repeatability was ±5.5 N. In this small sample size, a statistically insignificant and likely clinically insignificant bias in the direction of the third visit's measures being higher than the second visit's measures is indicated as a group mean difference of −0.14 N. The differences between measures obtained across the various visits were distributed evenly above and below zero when comparing Visits 2 and 3; Visit 1, however, shows a bias of lower measures when compared to either Visit 2 or Visit 3.

DISCUSSION

Measurements of vaginal closure forces during a MVC appear to have the best repeatability between the 2nd and 3rd visits, and lesser repeatability between the 1st and 2nd visits' measures. This suggests a between-visit learning effect. Thus, the most likely explanation for the finding of a larger mean value obtained on the 2nd visit, which has also been observed when testing other striated muscle,[36] is a practice effect due to improved awareness and skill resulting from the individualized instruction. Because the visits were spaced close together (about a week apart), we do not believe that pelvic floor muscle hypertrophy could have occurred. It is possible that women practiced at home, despite the fact that we did not specifically ask them to do so. Regardless of explanation, we shall see that these findings provide independent corroboration of the results from Dumoulin et al.[28] using a similar dynamometer also designed to measure vaginal closure force.

In terms of within-day repeatability Dumoulin et al.[28] found a coefficient of dependability of 0.88 (a statistic similar to the intra-class correlation coefficient) and a standard error of the mean (SEM) of 1.49 N, or 21% of the approximately 6 N mean force, for measurements taken three times on any single day. Our data, recalculated using the intra-class correlation coefficient and SEM, are comparable. For measures repeated twice on any single day, the best intra-class correlation coefficient was 0.83 and a SEM of 0.86 N, or 13.9% of the 6.22 N mean force. There was a trend in our data toward the second MVC measure of the day being greater than the first MVC measure (Table II, first three rows). The most likely explanation for this is the thermal drift in the strain gage transducer output (noted in Materials and Methods) as the distal 5 cm of its speculum bills warmed from ambient air temperature to body temperature during the 5 minutes of the measurements. This effect could have been obviated by storing the transducer in a warming drawer at 37°C. It also could have been obviated by conducting all analyses with the difference between peak force during MVC and the force at rest measured immediately before and after the MVC attempt. However, such an analysis would not have given the absolute vaginal closure force, rather the volitional augmentation in closure force. Finally, it is unlikely that the effect of temperature drift had any effect on between-visit results because a standard test protocol was used for all visits.

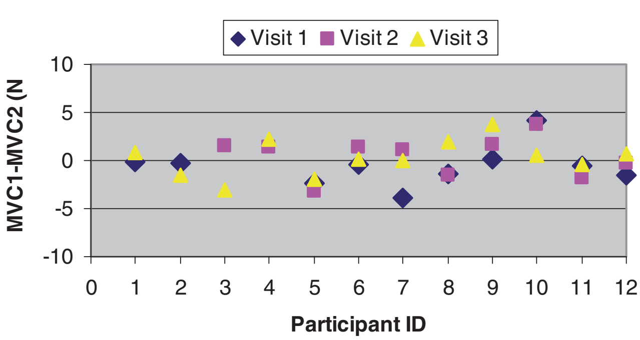

Fig. 2. Plots of the differences between repeated intravaginal closure force measures obtained in newtons (N) during maximal voluntary contraction (MVC) maneuvers on Visit 1, Visit 2, and Visit 3. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

We were unable to conduct a full generalization analysis[37] as reported by Dumoulin et al.[28] since our sample size of N = 12 lacked sufficient degrees of freedom to calculate the variance components for the three-way interaction: "subject" × "day" × "trial." However, in considering Dumoulin's findings that trial contributed negligible variance, we instead ran a comparable generalization analysis using just "subject" × "day." Results indicated that "subject" contributed

30.6%, "day" contributed 2.6%, and "subject $\times$ day" contributed the remaining 66.8% of the variance. These results are comparable to those found by Dumoulin on endurance measurements after 10 sec, respectively: 38.9%, 0%, and 61.2% for "subject," "day," and "subject" $\times$ "day."

There are inherent problems in assessing the repeatability of an instrument designed to measure pelvic muscle strength. Chief among these is the fact that 30–50% of women generally have poor motor control in voluntarily contracting this muscle, both in terms of the magnitude of the maximum contraction and the ability to refrain from a bearing down effort.[38,39] We believe our data reflect this poor motor control and that the majority of the observed variation in the readings of maximum vaginal closure pressure comes from a woman's imprecise activation of this muscle group by the central nervous system (known as central activation), rather than from variability in the input–output characteristics of the measuring device itself, which proved better than 2% in bench tests on the same or different days. This would explain why the present variability is considerably greater than test-retest values for maximum voluntary isometric muscle strength in other striated muscle. For example, in measurements of isometric neck muscle strength the mean difference from one day to the next has been shown to be about 10%, and the coefficient of repeatability can range from 11% to 33% of the mean of the first day depending upon the test direction.[36] A second problem is that both the perineometer and the instrumented speculum readings are subject to systematic bias due to the intra-abdominal pressure rise that inevitably accompany a maximum vaginal closure force attempt.[40] The higher the intra-abdominal pressure, the greater will be the vaginal closure force that is registered. The insertion length of the speculum bills is 7 cm, thus approximately 4 cm length could extend into the abdominal cavity beyond the "high pressure zone" due to the levator ani located approximately 2–3 cm inside the hymenal ring. To counter this effect, we now routinely measure intra-abdominal pressure separately and do not accept vaginal closure force readings obtained when the accompanying rise in intra-abdominal pressure exceeds 5 cm H_2O.[4] However, that policy is not feasible to implement in certain vulnerable populations, such as pregnant women. Considering this, the policy was not implemented in this study, or that of Dumoulin et al.,[28] so the present results are conservative estimates of repeatability. Finally, some variability can be attributed to the operator holding the device (Fig. 1) with slightly different bias force in the sagittal plane while resisting the tendency of intra-abdominal pressure to push the device out of the abdominal cavity. As long as this bias force is directed axially along the device it will have no effect; but if it contains a force component that acts ventral or dorsal to the long axis of the device, then it can affect the sensed force slightly.

Intravaginal closure force provides a global assessment of muscle and connective tissue function responsible for closure of the urogenital hiatus, and indirectly an assessment of the support available to the pelvic organs. Measurement of intravaginal closure force is therefore useful for assessing and monitoring changes that may occur by age, from childbirth or surgical injury mechanisms, as well as from treatment outcome. The instrumented speculum, in contrast to digital measures, intravaginal balloon type devices, or EMG studies, offers the advantage of an objective measure of the isometric contractile force, in a known direction within the mid-sagittal plane, generated by the pelvic floor muscles to close the urogenital hiatus.

CONCLUSIONS

The repeatability of the instrumented speculum for measuring maximal volitional isometric vaginal closure force is consistent with that reported by Dumoulin et al.[28] However, variability in the central activation of the pelvic floor muscles likely explains why both studies report the repeatability of the measurement of maximal vaginal closure force is less than one-third that reported for the maximum isometric strength of other striated muscle [e.g., neck muscle, as measured by Ylinin et al.,[36]].

REFERENCES

1. DeLancey JOL. Structural support of the urethra as it relates to stress urinary incontinence: The hammock hypothesis. Am J Obstet Gynecol 1994;170: 1713–23.
2. Miller JM, Perucchini D, Carchidi LT, et al. Pelvic muscle contraction during a cough and decreased vesical neck mobility. Obstet Gynecol 2001;97: 255–60.
3. Shafik A, Doss S, Asaad S. Etiology of the resting myoelectric activity of the levator ani muscle: Physioanatomic study with a new theory. World J Surg 2003;27:309–14.
4. Morgan DM, Kaur G, Hsu Y, et al. Does vaginal closure force differ in the supine and standing positions? Am J Obstet Gynecol 2005;192:1722–8.
5. Chen L, Ashton-Miller JA, Hsu Y, et al. Interaction among apical support, levator ani impairment, and anterior vaginal wall prolapse. Obstet Gynecol 2006a;108:324–32.
6. Peschers UM, Gingelmaie A, Jundt K, et al. Evaluation of pelvic floor muscle strength using 4 different techniques. Int Urogynecol J Pelvic Floor Dysfunct 2001;12:27–30.
7. Brink CA, Wells TJ, Sampselle CM, et al. A digital test for pelvic muscle strength in women with urinary incontinence. Nurs Res 1994;43:352–6.
8. Kalk CS. Nerve Injuries Committee. Medical Research Council, Great Britain. Aids to the investigation of peripheral nerve injuries War Memorandum #7, revised 2nd edition; 1943.
9. Anonymous. Kegel—perineometer. J Am Med Assoc 1954;154:679.
10. Shepherd AM, Montgomery E, Anderson RS. Treatment of genuine stress incontinence with a new perineometer. Physiotherapy 1983;69:113.
11. Dougherty MC, Abrams R, McKey PL. An instrument to assess the dynamic characteristics of the circumvaginal musculature. Nurs Res 1986;35:202–206.
12. Dougherty MC, Bishop KR, Mooney RA, et al. Variation in intravaginal pressure measurements. Nurs Res 1991;40:282–85.
13. Frawley HC, Galea MP, Phillips BA, et al. Reliability of pelvic floor muscle strength assessment using different test positions and tools. Neurourol Urodyn 2006;25:236–42.
14. Guaderrama NM, Nager CW, Liu J, et al. The vaginal pressure profile. Neurourol Urodyn 2005;24:243–7.
15. Bö K, Finckenhagen HB. Vaginal palpation of pelvic floor muscle strength: Inter-test reproducibility and comparison between palpation and vaginal squeeze pressure. Acta Obstet Gynecol Scand 2001;80:883–87.
16. Bö K, Stein R. Needle EMG registration of striated urethral wall and pelvic floor muscle activity patterns during cough, Valsalva, abdominal, hip adductor, and gluteal muscle contractions in nulliparous healthy females. Neurourol Urodyn 1994;13:35–41.
17. DeLancey JOL, Kearney R, Chou Q, et al. The appearance of levator ani muscle abnormalities in magnetic resonance images after vaginal delivery. Obstet Gynecol 2003;101:46–53.
18. Tunn R, DeLancey JO, Quint E. Visibility of pelvic organ support system structures in magnetic resonance images without an endovaginal coil. Am J Obstet Gynecol 2001;184:1156–63.
19. Margulies RU, Hsu Y, Kearney R, Stein T, Umek WH, DeLancey JO. Appearance of the levator ani muscle subdivisions in magnetic resonance images. Obstet Gynecol 2006;107:1064–9.
20. Chen L, Hsu Y, Ashton-Miller JA, et al. Measurement of the pubic portion of the levator ani muscle in women with unilateral defects in 3-D models from MR images. Int J Gynaecol Obstet 2006b;92:234–41.
21. Ikai M, Fukunaga T. A study on training effect on strength per unit cross-sectional area of muscle by means of ultrasonic measurement. Int Z Angew Physiol 1970;28:173–80.
22. Dietz HP, Lanzarone V. Levator trauma after vaginal delivery. Obstet Gynecol 2005;106:707–712.
23. Kearney R, Miller JM, Ashton-Miller JA, et al. Obstetrical factors associated with levator ani muscle injury after vaginal birth. Obstet Gynecol 2006;107:144–9.
24. Hoyte L, Jakab M, Warfield SK, et al. Levator ani thickness variations in symptomatic and asymptomatic women using magnetic resonance-based 3-dimensional color mapping. Am J Obstet Gynecol 2004;191:856–61.

25. Ashton-Miller JA, DeLancey JOL, Warwick DN. An apparatus for measuring properties of the pelvic floor muscles. US Patent # 6,468,232 B1; 2002.
26. Sampselle CM, Miller JM, Mims BL, et al. Effect of pelvic muscle exercise on transient incontinence during pregnancy and after birth. Obstet Gynecol 1998;91:406–412.
27. Dumoulin C, Bourbonnais D, Lemieux MC. Development of a dynamomenter for measuring the isometric force of the pelvic floor musculature. Neurourol Urodyn 2003;22:648–53.
28. Dumoulin C, Gravel D, Bourbonnais D, et al. Reliability of dynamometric measurements of the pelvic floor musculature. Neurourol Urodyn 2004;23: 134–42.
29. Morin M, Dumoulin C, Bourbonnais D, et al. Pelvic floor maximal strength using vaginal digital assessment compared to dynamometric measurements. Neurourol Urodynam 2004a;23:336–41.
30. Morin M, Bourbonnais D, Gravel D, et al. Pelvic floor muscle function in continent and stress urinary incontinent women using dynamometric measurements. Neurourol Urodynam 2004b;23:668–74.
31. Verelst M, Leivseth G. Force-length relationship in the pelvic floor muscles under transverse vaginal distension: A method study in healthy women. Neurourol Urodyn 2004;23:662–7.
32. Dechow PC, Carlson DS. A method of bite force measurements in primates. J Biomech 1983;16:797–802.
33. Lien K-C, Mooney B, DeLancey JOL, et al. Levator ani muscle stretch induced by simulated vaginal birth. Obstet Gynecol 2004;103:31–40.
34. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
35. Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Comput Biol Med 1990;20:337–40.
36. Ylinin J, Salo P, Nykänen M, et al. Decreased isometric neck strength in women with chronic neck pain and repeatability of neck strength measurements. Arch Phys Med Rehab 2004;85:1303–8.
37. Shavelson RJ, Webb NM. Generalizability theory: A primer, Vol. 1. Newbury Park, California: Sage Publications Inc., 1991.
38. Bump R, Hurt G, Fantl A, et al. Assessment of Kegel pelvic muscle exercise performance after brief verbal instruction. Amer J Obstet Gynecol 1991;167: 322–329.
39. Sampselle C, Messer K, Seng J, et al. Learning outcomes of a group behavioral modification program to prevent urinary incontinence. Int Urogynecol J 2005;16:441–446.
40. Bö K, Kvarstein B, Hagen R, et al. Pelvic muscle exercise for the treatment of female stress urinary incontinence: II validity of vaginal pressure measurements of pelvic floor strength. The necessity of supplementary methods for control of correct contraction. Neurourol Urodyn 1990;9:479–87.