Intra-Rater Reliability and Diagnostic Accuracy of a New Vaginal Dynamometer to Measure Pelvic Floor Muscle Strength in Women With Urinary Incontinence

Georgia Romero-Cullerés,1,4 Esteban Peña-Pitarch,2 Celia Jané-Feixas,1 Anna Arnau,3 Jesus Montesinos,3 and Montserrat Abenoza-Guardiola4

1Physical Medicine and Rehabilitation, Althaia Xarxa Assistencial Universitària de Manresa, Manresa, Barcelona, Spain
2Escola Politècnica Superior d’Enginyeria de Manresa (EPSEM), Universitat Politècnica de Catalunya, Manresa, Barcelona, Spain
3Clinical Research Unit, Althaia Xarxa Assistencial Universitària de Manresa, Manresa, Barcelona, Spain

Aims: The first choice treatment in urinary incontinence (UI) is rehabilitation of the pelvic floor in order to improve muscle strength. However, no entirely reliable instruments for quantifying pelvic floor muscle (PFM) strength are currently available. Our aim was to test the intra-rater reliability and diagnostic accuracy of a new vaginal dynamometer for measuring PFM strength. Methods: Test-retest reliability study. One hundred and four women with stress urinary incontinence (SUI) were recruited. Patients were excluded if they had a history consistent with urge urinary incontinence or pelvic organ prolapse, pregnancy, previous urogynecological surgery, severe vaginal atrophy, or neurological conditions. The examination comprised digital palpation quantified by the modified Oxford scale and by two consecutive dynamometry measurements obtained using a new prototype dynamometer. This instrument comprises a speculum in which an inductive displacement sensor (LVDTSM210.10.2.KT model, Schreiber) is attached to a spring of known stiffness constant (k). The intraclass correlation coefficient (ICC) was calculated to assess intra-rater reliability. Diagnostic accuracy was assessed using Receiver Operating Characteristics (ROC) curves analysis. Results: Of the 104 subjects included, 59.6% presented scores between 0–2 on the Oxford scale. Intra-rater reliability was 0.98 (95%CI: 0.97–0.99). In the Bland & Altman plot, the distribution of disagreements was similar in the lowest and the highest strength values. The diagnostic accuracy of the dynamometer with regard to digital palpation showed an area under the curve of 0.85 (95%CI: 0.77–0.93). Conclusions: Our results suggest that this new vaginal dynamometer is a reliable and valid instrument for quantifying PFM strength. Neurourol. Urodynam. © 2015 Wiley Periodicals, Inc.

Key words: dynamometer; pelvic floor muscle strength; rehabilitation; urinary incontinence

INTRODUCTION

The International Continence Society (ICS) defines urinary incontinence (UI) as any involuntary urine leakage.1 Stress urinary incontinence (SUI) is the most frequent type of UI, defined as involuntary leakage on effort or physical exertion, or upon sneezing or coughing.2 The prevalence of UI ranges from 14% to 69% and increases with age. More than 50% of patients attending geriatric clinics are incontinent.3 UI is far more prevalent than other pathologies such as hypertension, depression or diabetes mellitus, and may have a highly negative impact on quality of life,4 reducing social, personal and sexual relations,5 and physical activity.6

Stress urinary incontinence (SUI) appears to result from multiple failures in the continence mechanism, where defects in sphincteric function,7 pelvic floor muscle (PFM),8 connective tissues,9 or neural structures10 may all play a role. Physical therapy, mainly involving strength training of the PFM and training motor control strategies to prevent urine loss,6–8 has been shown positive results in women with SUI. As such, physical therapy has been recommended by the International Continence Society as the first-line treatment for women whose primary complaint is SUI.8,9 This treatment helps support pelvic organs (urethra, uterus, and bladder), thus avoiding genital prolapse and other dysfunctions of the pelvic floor such as incontinence.10

At present, the standard instruments for evaluating pelvic floor muscle strength are vaginal palpation and manometry. Vaginal palpation using the modified Oxford grading scale11 is currently the gold standard, although it is a totally subjective method with low test-retest and inter-rater reliability.12 Vaginal manometry13 is not an efficient method for measuring intravaginal pressure, as it must always be performed at the same anatomical level, and taken into account that measurements of abdominal pressures can alter the PFM response.14

The ideal location appears to be the urethra rather than the vagina, with the same baseline pressures, the same units of measurement and the same size and shape of instrument, but these requirements are difficult to achieve.15 In addition, vaginal manometry evaluates pressure, and is mistakenly identified as an instrument for measuring strength. Some authors suggest assessing PFM strength using prototype vaginal dynamometers, but none are commercially available. Other techniques used to evaluate PFM function include electromyography (EMG), ultrasound, and dynamic magnetic

Dr. Fred Milani led the peer-review process as the Associate Editor responsible for the paper.

Potential conflicts of interest: Dr. Peña and Dr. Romero report a European patent application number 12765883.9 licensed to OIARSO S. COOP.

Grant sponsor: Spanish Government; Grant number: DPI2013-40882-P

*Correspondence to: Dr. Georgia Romero-Cullerés, Physical Medicine and Rehabilitation, Althaia Xarxa Assistencial Universitària de Manresa, Avinguda Bases de Manresa, 6-8, 08242 Manresa, Barcelona, Spain.

E-mail: gromero@althaia.cat

Received 23 June 2015; Accepted 16 October 2015

Published online in Wiley Online Library (wileyonlinelibrary.com).

DOI 10.1002/nau.22924

© 2015 Wiley Periodicals, Inc.
resonance imaging. Many of these techniques have limited clinical utility as the result of poor validation or reliability, limited accessibility and high cost. 16–19

One of the possible causes of SUI is bladder neck hypermobility, which might result from PFM weakness. 20 PFM strength in women is directly related to UI and the appearance of genital prolapse. When the bladder neck is located above the pelvic floor, the pressure transmitted to the bladder is equally transmitted to the urethra, resulting in a simultaneous increase of urethral closure pressure. If the bladder neck position is lower than the pelvic floor, greater pressure is transmitted to the bladder than to the urethra, threatening urethral closure, and continence status. A voluntary counterbracing contraction technique, termed “counterbracing” or “Knack,” which involves contracting the PFM just before the increase in intra-abdominal pressure, can significantly reduce the descent of the bladder neck during a cough. 21 Pelvic-floor muscle strengthening has been widely used to treat SUI since Kegel introduced this kind of exercise. 22 Intensive PFM strengthening is thought to increase muscle size and stiffness and to stabilize the bladder neck during the increase of intra-abdominal pressure. 7 The effectiveness of the treatment has been demonstrated by numerous randomized controlled trials, according to international practice guidelines. 23,24 With the recent creation of new inpatient and outpatient multidisciplinary pelvic floor units, reliable data are needed to evaluate PFM function and strength in the initial assessment of a patient consulting for incontinence and to evaluate the results obtained after treatment and rehabilitation. Researchers from our pelvic floor rehabilitation unit and biomechanical engineers recently designed a prototype dynamometric speculum (patent pending). This new vaginal dynamometer measures PFM strength and provides an objective assessment of the results of rehabilitation treatment. The aim of this study was to test the reliability and diagnostic accuracy of this new instrument for measuring PFM strength in women with urinary incontinence.

MATERIALS AND METHODS

PFM strength was measured by a new prototype dynamometer (European patent application number 12765883.9) comprising a speculum in which an inductive displacement sensor (LVDT SM210.10.2.KT model, Schreiber) is attached to a spring of known stiffness constant (k) (Fig. 1). The measurement obtained is the displacement—that is, the extent to which the spring is compressed. The instrument comprises a speculum formed by two pivoting pieces, each with a handle and a frontal area, which is introduced in the vagina. A displacement sensor is attached to the frontal area. The spring has a wire diameter between 0.5 and 1 mm.

Once the measurements are recorded, the data are processed to give the strength in newtons (N), applying a linear regression equation:

\[
\text{Strength}(N) = (23.245x + 156.457) \times 9.81
\]

Where x is the spring displacement recorded, the results are multiplied by 9.81 to convert into N. This measurement indicates to the physician or physiotherapist the real force exerted by the patient on the pelvic floor muscle.

Design

Test-retest reliability study.

Patients

Patients visiting a pelvic floor rehabilitation specialist in our Center were screened for possible inclusion in this study. One hundred and four consecutive women over 18 years old with stress urinary incontinence (SUI) were recruited between November and December 2011. Patients were excluded if they had a history consistent with urge urinary incontinence or pelvic organ prolapse, current pregnancy, previous urogynecological surgery, severe vaginal atrophy, neurological conditions, or cognitive impairment.

Procedures

Examinations were conducted by a specialist in pelvic floor rehabilitation with the participants in the lithotomy position. The first examination comprised bidigital palpation quantified by the modified Oxford scale, to estimate pelvic floor muscle strength during maximum contraction. The modified Oxford
grading scale records the patient’s capacity to contract the pelvic floor muscles, as perceived by the specialist performing the palpation (the patient was instructed to contract concentrically in a cranio-ventral direction. The use of palpation and visual inspection during the physical examination revealed that all the subjects correctly understood the instructions). The score ranges from 0 to 5.

Subjects were given guided instructions by the specialist on how to perform a PFM contraction. Prior to the assessment, speculum branches were properly disinfected and covered with a female condom. PFM strength was evaluated twice in all women. Two consecutive measurements of the strength of contraction of pelvic floor were recorded using our dynamometric speculum. The rest period between the two measurements was 30 sec. Two values were recorded for each measurement: one initial value of the passive force after opening the device for 5 sec, and the maximal voluntary strength registered by the device in a period of 10 sec. In each measurement the dynamometric speculum was inserted into the vagina and the strength exerted was recorded. The strength of the contraction was calculated as maximum contraction strength minus baseline strength. PFM strength was recorded in N. All the measurements were performed by a single rater, a rehabilitation physician. The study protocol was approved by an independent Clinical Research Ethics Committee. Before recruitment, all patients provided written informed consent to participate.

**Statistical Analysis**

Continuous variables were summarized using means, standard deviations (SD), minimum and maximum value. Categorical variables were summarized using absolute values and relative frequency.

The mean dynamometric value was computed for each category of the modified Oxford grading scale and compared using a one-way ANOVA. The Scheffé post-hoc procedure was used to compare differences between pairs of assessment categories.

The intraclass correlation coefficient for two-way mixed effect model and for absolute agreement ICC (A,2) was calculated to assess the intra-rater reliability of the dynamometer measures. The Bland & Altman plot and a survival-agreement plot were also used. The receiver operating characteristic (ROC) curve graph and the area under the curve (AUC) were calculated in order to obtain the diagnostic accuracy of PFM strength measured by dynamometric speculum with the gold standard. Vaginal palpation was measured with the Modified Oxford grading scale grouping the categories from 0 to 2 and from 3 to 5.

A two-sided 5% (P < 0.05) significance level was assumed. All statistical analyses were performed using IBM® SPSS® Statistics for Windows v.20 (IBM Corporation, Armonk, NY) and Stata® v.10 (Statacorp LP, College Station, Texas).

**RESULTS**

First baseline and maximum contraction measurements were performed using the dynamometric speculum in 104 patients. Second baseline and maximum contraction measurements were obtained from 102 patients. The second measurement could not be made in two patients due to technical problems. Mean age was 56 years (SD = 10.3). Mean PFM strength obtained was 0.58 N (SD = 0.40) in the first measurement and 0.60 N (SD = 0.42) in the second measurement (Table I).

Scores between 0 and 2 in the modified Oxford Grading Scale were recorded in 59.6% of patients. Table II shows the pelvic floor muscle strength measured with the dynamometric speculum for each category of the modified Oxford grading scale. The ANOVA analysis was significant (F = 27.862; P < 0.001), indicating differences in PFM strength across digital assessment categories. In the post-hoc analysis, no significant differences were found between adjacent assessment categories such as 0–1 and 3–4. Statistical differences were found between adjacent categories 1–2 and 2–3, and between all other pairs of categories.

Intra-rater reliability according to ICC (A,2) was 0.978 (95%CI: 0.968–0.985). According to the Bland & Altman plot, the mean difference between the two measurements of PFM strength was −0.02 N (SD = 0.12). Intra-rater disagreements were distributed similarly in the lowest and the highest strength values. According to the survival-agreement plot, the difference between the two measurements was 0.05 N or lower in 50% of patients, and 0.22 N or lower in 90% (Fig. 2).

The results for the diagnostic accuracy of the dynamometer compared with digital palpation calculated using the ROC analysis showed an area under the curve of 0.85 (95%CI = 0.77–0.93) (Fig. 3).

**DISCUSSION**

The purpose of this study was to evaluate the diagnostic accuracy and intra-rater reliability of a new dynamometric speculum measuring pelvic floor muscle strength, based on a mechanical inductive displacement sensor attached to a standard vaginal speculum. Our results show good diagnostic accuracy in SUI patients using the modified Oxford scale as the reference method, and very good intra-rater reliability.

The standard method used to evaluate the contraction capacity of the pelvic floor (digital examination assessed with the modified Oxford grading scale) has only limited capacity to reveal the true difference in strength. It is poorly suited to detect changes over time or to evaluate the response to rehabilitation, because of the high level of rater subjectivity and the low test-retest reliability and inter-rater reproducibility.

The dynamometric speculum described in this study presents several advantages over traditional measurement instruments. In addition to its accurate results as a measurement instrument, our device stands out for its simplicity of use, the fact that it can be used by different healthcare professionals, and that it

**TABLE I. Pelvic Floor Muscle Strength Measured With Dynamometric Speculum (N)**

<table>
<thead>
<tr>
<th>Measurement</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline strength (N) 1st</td>
<td>104</td>
<td>1.72</td>
<td>0.21</td>
<td>1.57</td>
<td>2.72</td>
</tr>
<tr>
<td>Baseline strength (N) 2nd</td>
<td>102</td>
<td>1.70</td>
<td>0.19</td>
<td>1.57</td>
<td>2.37</td>
</tr>
<tr>
<td>Contraction strength (N) 1st</td>
<td>104</td>
<td>2.31</td>
<td>0.49</td>
<td>1.57</td>
<td>3.33</td>
</tr>
<tr>
<td>Contraction strength (N) 2nd</td>
<td>104</td>
<td>2.30</td>
<td>0.49</td>
<td>1.57</td>
<td>3.27</td>
</tr>
<tr>
<td>Strength (N) 1st</td>
<td>104</td>
<td>0.58</td>
<td>0.40</td>
<td>0.00</td>
<td>1.46</td>
</tr>
<tr>
<td>Strength (N) 2nd</td>
<td>102</td>
<td>0.60</td>
<td>0.42</td>
<td>0.00</td>
<td>1.49</td>
</tr>
</tbody>
</table>

SD, standard deviation; N, newtons; Strength measurement = Contraction strength—Baseline strength.
can be made in different sizes to accommodate different morphologies and sizes of the vaginal canal.

In 2003, Dumoulin et al.\(^\text{27}\) described an instrument based on a non-standard speculum, comprising two aluminum branches. One branch was fixed while the other was equipped with a strain gauge. In spite of this and other attempts to modify standard or newly designed speculums as instruments to measure PFM strength, none of these instruments have been commercialized to date and they appear to be used only in research.\(^\text{28,29}\) Similarly, vaginal manometers or perineometers,\(^\text{13,15}\) instruments for measuring pressure, have been used to measure PFM strength, even though it is known that pressure and strength are different physical parameters. Vaginal manometers have known limitations such as the fact that intraurethral pressure should be measured rather than intravaginal pressure, and the huge variability in terms of type and size among the devices in the market. Furthermore, no standard reference parameters have been established to date, and the baseline pressures used by different researchers vary widely.

Most of the equipment currently available can only be used to measure pressure. The exception is the dynamometer, which can be used directly to measure strength and can provide reliable PFM strength measurements.\(^\text{30,31}\) In the conclusions of a systematic review of PFM training and adjunctive therapies for SUI, Neumann et al.\(^\text{32}\) called for a more standardized approach to outcome measurement in research with appropriate outcome measures reflecting clinical practice requirements.

The limitations of this study include aspects related to the prototype and aspects related to the study design. Measurements may have been contaminated by speculum deformities or the use of female condoms. Thus, the baseline measurement was obtained when the opening of the speculum blades adapted to the muscle tone at rest. Potential improvements to be considered would include an independent opening spring having a constant stiffness to adapt speculum blades to the PFM tone at rest, as well as the use of disposable materials. Regarding the study design, most of our patients were postmenopausal and all presented with SUI. Sixty per cent scored between 0 and 2 on the modified Oxford grading scale. In future studies, the patient group should be expanded to include women with high scores on the Oxford scale (4–5), different sizes of speculum and a control group of continent women in order to validate the results presented in our study.

### CONCLUSIONS

The dynamometric speculum used in this pilot study presented good test-retest reliability, and diagnostic accuracy. This new instrument offers considerable advantages over those currently available and appears to obtain objective clinical evaluations and an accurate measurement of pelvic floor muscle strength. It avoids the problem of rater subjectivity and allows close monitoring of the results of pelvic floor

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No contraction (0)</td>
<td>6</td>
<td>0.06</td>
<td>0.10</td>
<td>0.00</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>1st measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very weak contraction (1)</td>
<td>24</td>
<td>0.24</td>
<td>0.25</td>
<td>0.00</td>
<td>0.84</td>
<td>0.689</td>
</tr>
<tr>
<td>1st measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak contraction (2)</td>
<td>32</td>
<td>0.55</td>
<td>0.28</td>
<td>0.05</td>
<td>1.08</td>
<td>0.005</td>
</tr>
<tr>
<td>1st measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate contraction: maintenance of pressure (3)</td>
<td>36</td>
<td>0.83</td>
<td>0.32</td>
<td>0.02</td>
<td>1.32</td>
<td>0.003</td>
</tr>
<tr>
<td>1st measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good contraction: maintenance of tension with resistance (4)</td>
<td>6</td>
<td>1.17</td>
<td>0.26</td>
<td>0.83</td>
<td>1.46</td>
<td>0.172</td>
</tr>
<tr>
<td>1st measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SD, standard deviation; N, newtons; Measurement = Contraction strength / Baseline strength.

Post-hoc P-values for the Scheffe test between adjacent Oxford scale categories (0 vs. 1; 1 vs. 2; 2 vs. 3; and 3 vs. 4).

![Bland-Altman plot](image1)

**Fig. 2.** Intra-rater reliability according to Bland & Altman plot and to survival-agreement plot.

*Neurourology and Urodynamics* DOI 10.1002/nau
rehabilitation. More studies are now needed to confirm our results, evaluate inter-rater reliability, and compare the instrument with the alternatives currently available.

ACKNOWLEDGMENT

This work was partially supported by the Spanish government with the project DPI2013-40882-P.

REFERENCES


Dynamometer and Pelvic Floor Muscle Strength

Fig. 3. Diagnostic accuracy of the dynamometer with regard to the gold standard (digital manual examination evaluated using the modified Oxford Grading Scale). Receiver operating characteristic curve and area under the curve.