

Interrater Reliability and Physical Examination of the Pubovisceral Portion of the Levator Ani Muscle, Validity Comparisons Using MR Imaging

Rohna Kearney,* Janis M. Miller, and John O. L. DeLancey

Department of Obstetrics and Gynecology, University of Michigan Medical School, Ann Arbor, Michigan

Aims: Defects in the pubovisceral portion of the levator ani muscle are seen with MR imaging. This study aims to determine interrater reliability of physical examination in detecting these defects, and to validate findings from physical examination using comparisons with MR images. **Methods:** Two examiners palpated the pubovisceral muscles of 29 women to assess for defects in this muscle. Each examiner was blinded to the others findings. MR scans were acquired on a further 24 women after structured clinical examination by one examiner. These images were read to determine pubovisceral muscle defects, blinded to patient identifiers. Agreement between raters and between MR imaging and clinical examination were calculated. **Results:** The two examiners had positive agreement (presence of a defect) of 72.7% and negative agreement (absence of a defect) of 83.3%. The positive agreement between physical examination and MR imaging was 27.3% and the negative agreement 86.5%. **Conclusion:** The structured physical examination to detect defects in the pubovisceral portion of the levator ani muscle can be learned as shown by good interrater agreement. However, examination alone underestimates these defects compared with MR imaging. *NeuroUrol. Urodynam.* 25:50–54, 2006. © 2005 Wiley-Liss, Inc.

Key words: levator ani muscle; MRI; pubovisceral; pelvic floor

INTRODUCTION

Defects in the levator ani muscle have long been recognized as a contributing cause of pelvic floor dysfunction [Halban and Tandler, 1907; Berglas and Rubin, 1953; Koelbl et al., 1989; Hanzal et al., 1993; DeLancey et al., 2003]. MR imaging has objectively demonstrated these defects in women with stress urinary incontinence and pelvic organ prolapse particularly in the region of the pubovisceral portion of the levator ani [Tunn et al., 1998; Hoyte et al., 2001]. This part of the muscle, also known as the pubococcygeal muscle attaches the pelvic organs, including the tissues that support the urethra, to the pubic bones. Imaging studies, however, are expensive, and at present, not obtained during routine clinical care.

The levator muscle is palpable through the vaginal wall on physical examination. Damage occurring after vaginal birth, as assessed by digital palpation, has previously been recorded [Gainey, 1943, 1955]. The accuracy of physical examination, however, against objective imaging techniques needs to be clarified. We therefore sought to determine the accuracy of physical examination compared with MRI to detect these abnormalities and whether acceptable agreement between examiners could be achieved on physical examination without MR imaging. Without knowledge of the reliability and validity of inexpensive physical examination as a determinant of defects, we are limited in our ability to study questions such as: “Is levator ani injury associated with operative failure?” or “Do certain obstetrical factors cause levator ani injury?”

The purpose of this study was to determine interrater reliability in detecting abnormalities in the pubovisceral portion of the levator ani muscle, and validate findings from physical examination using comparisons with MR images.

MATERIALS AND METHODS

Test-Retest in Physical Examination

In 29 women presenting for routine care in a urogynecology clinic, two physician examiners assessed the pubovisceral portion of the levator ani muscle by using a structured examination while remaining blinded to one another's results. The first examiner (JD), was a recognized expert in the field of urogynecology with greater than 15 years of experience in anatomical dissection, pelvic floor imaging, and assessing the levator ani muscle on physical examination. The second

Abbreviations: BS, bulbospongiosus; HM, hymenal ring; ICM, iliococcygeal muscle; LM, labia majora; OI, obturator internus; PE, physical examination; PB, pubic bones; PVM, pubovisceral muscle; R, rectum; U, urethra; V, vagina.

*Correspondence to: Rohna Kearney, 1 Upper Gwydir Street, Cambridge, CB1 2LR, UK. E-mail: rkearney@doctors.net.uk

Received 13 November 2004; Accepted 10 June 2005

Published online 22 November 2005 in Wiley InterScience (www.interscience.wiley.com)

DOI 10.1002/nau.20181

examiner (RK), we will refer to as a “trained examiner,” having trained with the expert for 6 months time, was felt to represent the skill level that an individual wishing to carry out research on the levator ani muscle would possess.

Each evaluated the presence of muscle defects using the following maneuvers: (1) the pubovisceral muscle was palpated by placing the index finger laterally in the lower one third of the vagina (within approximately 2–3 cm of the hymen) so that the middle of the distal phalanx lies on the normal location of the pubovisceral muscle. Palpation assesses presence or absence of muscle bulk at rest and during contraction (contraction assists in identifying the presence of small amounts of muscle). Contraction of the bulbospongiosus muscle (outside the hymen) and iliococcygeal muscle (greater than 3 cm above the hymen) are differentiated from the targeted pubovisceral muscle and assist in confirming that a woman understands what to do.

For each physical examination, both the left and right sides of each muscle were evaluated and a defect judged to occur when the bulk of the pubovisceral portion of the levator ani muscle was found to be palpably different than that found by prior experience in assessing nulliparous women. In addition, to the levator ani assessment, the pelvic support of each woman was measured and recorded as the pelvic organ quantification score (POP-Q) [Bump et al., 1996].

Comparisons Between Physical Examination and MRI

An additional 24 women, recruited for an Institutional Review Board approved parent research study on pelvic floor disorders, also received a structured physical examination (by only the trained examiner) along with a multiplanar proton density MR image of the levator ani. The images were acquired using 2-dimensional fast spin (echo time, 15 msec; repetition time, 4,000 msec) at 5 mm intervals using a 1.5 T superconducting magnet (Signa; General Electric Medical Systems, Milwaukee, WI) with version 5.4 software, as previously described [Chou and DeLancey, 2001]. The resulting MR images were read to determine defects in the pubovisceral portion of the levator ani while remaining blinded to patient identifiers linking the MR to the physical examination.

An example of the appearance of normal and abnormal pubovisceral muscle as seen on MR images are provided in Figure 1. Note the loss of pubovisceral muscle bulk between the vagina and the internal obturator muscle in the axial and coronal scans. Determinations of muscle integrity were based on our previous experience with assessing normal muscle morphology [Strohbehn et al., 1996; Chou and DeLancey, 2001; DeLancey et al., 2003].

Statistical Analysis

For statistical analysis, positive or negative findings of defects were documented for each subject on physical examination and MR image. Agreement between raters, and

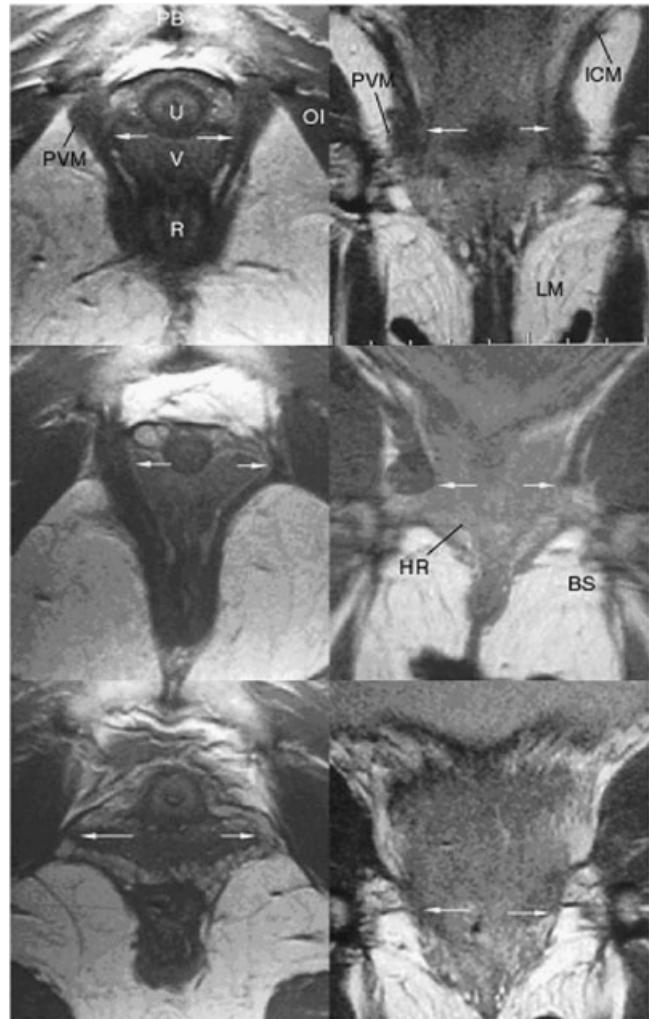


Fig. 1. MR proton density axial images (left) and coronal images (right) of three women showing both normal and abnormal pubovisceral muscles. The top images are from a nulliparous women. The pubovisceral muscle (PVM) is shown between the urethra (U), vagina (V), rectum (R) and obturator internus muscle (OI) as it attaches to the pubic bones (PB). The iliococcygeal (ICM) portion of the levator ani is shown in the coronal images. The arrows indicate where the bulk of the pubovisceral muscle is palpated on physical examination with the index finger in the vagina. The hymenal ring (HM), labia majora (LM) and the bulbospongiosus (BS) are indicated for reference. The middle images are of a woman with a unilateral abnormality on the right side of the images. The difference between the bulk of the pubovisceral muscle is appreciated between the two sides. The lower images are of a woman who has a bilateral abnormality of the pubovisceral muscles. No muscle is palpated on examination.

between MR image and physical examination findings, are reported using overall proportion of agreement, positive agreement, negative agreement, agreement expected by chance, and agreement corrected for chance (Cohen's kappa). These statistical computations, each with advantages and disadvantages in describing the data, are detailed in Kundel and Polansky [2003].

TABLE I. Demographics of the Two Cohorts PE (Physical Examination)

	PE/MR	PE/PE
Age (years) mean \pm SD	58.2 \pm 11.8	61.6 \pm 12.1
Parity (median, range)	2, 0–6	3, 1–5
Vaginal birth (median, range)	2, 0–6	2, 1–3
Forceps (median, range)	0, 0–4	0, 0–1
Cesarean section (median, range)	0, 0–1	0, 0

RESULTS

The age, parity, and obstetric history of the women in the study are shown in Table I. The stages of prolapse are shown in Table II.

Results of Test-Retest in Physical Examination

The overall proportion of agreement was 79.3%, however, there was an imbalance in the proportion of positive and negative findings within the sample. The trained examiner detected muscle defects in 9 of 29 women, the expert examiner detected muscle defects in 13 of 29 women, and both were in agreement that defects existed in 8 women. We therefore calculated positive agreement (presence of defect) 72.7%, negative agreement (absence of defect) 83.3%, and expected agreement by chance 51.9%. Cohen's kappa and standard error for agreement between the two raters in physical examination of the pubovisceral muscle was 0.569 ± 0.18 (Table V). Details of the agreement between the two examiners on physical examination are reported in Table III.

Results of MRI and Physical Examination Comparisons

Cohen's kappa and standard error for agreement between the MR images and the physical examination of the pubovisceral muscle was 0.444 ± 0.18 . The overall proportion of agreement was 79.2%, positive agreement (presence of defect) was 27.3%, negative agreement (absence of defect) was 86.5%, and expected agreement by chance was 62.6%. The MR images demonstrated muscle defects in eight women and the trained examiner detected muscle defects in three of these same women. Details of the agreement between the MR images and the examiners results by physical examination are reported in Table IV.

TABLE II. Stages of Prolapse

Stage	PE/MR	PE/PE
0	0	0
I	3	0
II	18	16
III	2	13
IV	1	0

TABLE III. Different Investigators Judgments of Significant Levator Ani Defect Presence or Absence by Physical Exam of 29 Women

JD physical exam	RK physical exam		Total
	Presence of defect	Absence of defect	
Presence of defect	8	5	13
Absence of defect	1	15	16
Total	9	20	29

DISCUSSION

This is the first study to focus attention on detecting defects in the levator ani muscle by physical examination with comparisons between raters and to an objective measure obtained by MR imaging. The study confirmed by good correspondence between blinded examiners that the physical examination can be readily learned. Defects in the pubovisceral portion of the levator ani muscle discovered on physical examination were confirmed by MR imaging. No false positive findings on examination were observed in this small sample, suggesting minimal risk of overestimating defects on physical examination. However, there was only moderate strength of agreement beyond chance in comparing physical examination findings to MR image findings, and physical examination estimated the number of women with defect at less than half the rate of those detected by MR imaging. Accuracy in findings based only on physical examination by an examiner of trained examiner capability was found to be insufficient for answering questions such as prevalence rates of the abnormality in various populations, or for confidently addressing potential research questions on associations of levator ani defects with comorbidities such as prolapse or incontinence. It may be possible to improve on this with extensive training.

It has long been recognized that the levator ani muscle plays a critical role in pelvic organ support [Halban and Tandler, 1907; Berglas and Rubin, 1953]. Despite this general appreciation of the muscle's importance, scientific study of the muscle has begun only recently with the advent of MR imaging that can objectively visualize normal and abnormal muscle anatomy. Evidence of muscle damage have been reported in women with pelvic organ prolapse and stress urinary incontinence and progress is being made to assess what specific role this muscle injury plays in the pathogenesis of these common problems [Berglas and Rubin, 1953; Koelbl et al., 1989; Kirschner-Hermanns et al., 1993; Hoyte et al., 2001]. Because the medial margin of the levator ani muscle is connected to the urethral supports [DeLancey, 1988] and the target of treatment with pelvic muscle exercise, the status of the levator ani muscle has particular implications for stress urinary incontinence.

The importance of establishing a technique that can reliably and accurately detect injury to the levator ani muscle that can be used by average examiners is important. Levator ani

TABLE IV. Same Investigator Judgment of Significant Levator Ani Defect Presence or Absence by MRI Readings Versus Physical Exam in 24 Women

MRI reading	RK physical exam		Total
	Presence of defect	Absence of defect	
Presence of defect	3	5	8
Absence of defect	0	16	16
Total	3	21	24

muscle damage is twice as common in women with new-onset stress incontinence after first birth [DeLancey et al., 2003] and there is suggestion, for example, that damage to the levator ani muscle is associated with operative failure for prolapse [Koelbl et al., 1989; Hanzal et al., 1993], significant in light of the fact that 29% of women experience operative failure [Olsen et al., 1997]. If we knew definitely the obstetrical factors that place women at increased risk of significant injury we may be able to prevent levator ani injuries in the delivery room, but accurate determination of levator ani injury is a necessary precursor to these studies. Availability of MR imaging as an accurate technique rather than reliance on physical examination alone may prove instrumental in completing these needed studies.

The structured examination used in this study focuses on detecting an absence of muscle substance in the pubovisceral portion of the levator ani muscle; the region we have found most often damaged by vaginal birth [DeLancey et al., 2003]. It is different than previously published techniques of examination [Worth et al., 1986; Sampselle et al., 1989; Brink et al., 1994]. Previous studies have generally employed a scale assessing three characteristics: pressure, duration, and displacement in plane [Brink et al., 1994; Sampselle et al., 1989]. These focus on the function of the muscle rather than its bulk or structural integrity. An earlier study included muscle ribbing as a 4th characteristic [Worth et al., 1986] but our technique focuses attention on detecting structural defects in the levator ani muscle that involve loss of muscle substance. Further research will elucidate how these techniques perform in assessing pelvic muscle function compared with other techniques such as EMG, force measurement, and intravaginal pressure quantification.

TABLE V. Agreement Between Raters, Examination and MRI

Agreement index	Type of agreement	JD versus RK examiner	MRI versus RK exam
P_o	Overall	0.793	0.792
P_{pos}	Positive for defect	0.727	0.273
P_{neg}	Negative for defect	0.833	0.865
P_e	Chance	0.519	0.626
$K \pm SE$	Chance corrected	0.569 ± 0.15	0.444 ± 0.18

It is possible that additional training could improve the results of physical examination, but we felt it best to evaluate the accuracy of the type of individual that would normally be expected to carry out examinations in a research setting. It must be emphasized that these results are specific to the two individuals chosen at a point in time. The trained examiner, whose readings were compared to MR findings, could improve to reduce false negatives by additional practice with comparisons to MR image results. However, the interrater reliability results, which were good, suggest that trained examiner's findings largely paralleled the expert's findings. Both of the examiners had excellent knowledge of the anatomy and previous experience with reviewing many MR scans of nulliparous women and comparing them to multiparous women who had defects seen in the levator ani. This anatomical insight provided by MR was the driving force behind developing a specific clinical examination to detect these defects.

Alternatively, it is possible that the defect findings on MR images could have been an overestimate. However, additional studies have confirmed the anatomical correctness of the MR portrayal of the pubovisceral muscle, and validated the same defects by comparisons with expected clinical findings such as higher rates of defects in women with prolapse [Hoyte et al., 2001, Singh et al., 2003]

Further work is suggested to determine if physical examination techniques in others' hands can achieve results that minimize the rate of false negatives that we obtained. Obstetricians and gynecologists examine many nulliparous women that provide a healthy sample of women for all interested clinicians to become familiar with the normal configuration of this muscle and become experienced in its assessment. But until proven otherwise, caution is indicated in making a definitive conclusion of "normal muscle" in a parous woman when assessing by physical examination alone.

Findings from this study demonstrate acceptable interrater reliability between two examiners in assessing defects in the pubovisceral portion of the levator ani muscle by physical examination. Positive findings by physical examination were confirmed by MR images. However, negative findings were inaccurate in half of the women who by subsequent MR imaging were found to have a greater than twofold higher prevalence of pubovisceral defects compared with those detected by physical examination alone. Poor positive agreement on defects between physical examination and MR images suggests that using physical examination only may grossly underestimate the prevalence of levator ani injury compared with proven cases by MR imaging.

CONCLUSION

A structured physical examination to detect defects in the pubovisceral muscle can be learned as shown by good interrater reliability, however examination alone underestimates these defects and MR imaging remains the method of choice for assessing levator ani defects.

REFERENCES

- Berglas B, Rubin IC. 1953. Study of the supportive structures of the uterus by levator myography. *Surg Gynecol Obstet* 97:677–92.
- Brink CA, Wells TJ, Sampelle CM, et al. 1994. A digital test for pelvic muscle strength in women with urinary incontinence. *Nurs Res* 43:352–6.
- Bump RC, Mattiasson A, Bo K, et al. 1996. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol* 175:10–7.
- Chou Q, DeLancey JOL. 2001. A structured system to evaluate urethral support anatomy in magnetic resonance images. *Am J Obstet Gynecol* 185:44–50.
- DeLancey JOL. 1988. Structural aspects of the extrinsic continence mechanism. *Obstet Gynecol* 72:296–301.
- DeLancey JOL, Kearney R, Chou Q, et al. 2003. The appearance of levator ani muscle abnormalities seen in MR images after vaginal delivery. *Obstet Gynecol* 101:46–53.
- Gainey HL. 1943. Postpartum observations of pelvic tissue damage. *Am J Obstet Gynecol* 45:457–66.
- Gainey HL. 1955. Post postpartum observations of pelvic tissue damage: Further studies. *Am J Obstet Gynecol* 70:800.
- Halban J, Tandler J. 1907. *Anatomie und aetiologie der genitalprolapse beim weibe*. Vienna and Leipzig: Wilhelm Braumuller.
- Hanzal E, Berger E, Koelbl H. 1993. Levator ani muscle morphology and recurrent genuine stress incontinence. *Obstet Gynecol* 81:426–9.
- Hoyte L, Schierlitz L, Zou K, et al. 2001. Two- and 3-dimensional MRI comparison of levator ani structure, volume, and integrity in women with stress incontinence and prolapse. *Am J Obstet Gynecol* 185:11–9.
- Kirschner-Hermanns R, Wein B, Niehaus S, et al. 1993. The contribution of magnetic resonance imaging of the pelvic floor to the understanding of urinary incontinence. *Br J Urol* 72:715–8.
- Koelbl H, Strassegger H, Riss PA, et al. 1989. Morphologic and functional aspects of pelvic floor muscles in patients with pelvic relaxation and genuine stress incontinence. *Obstet Gynecol* 74:789–95.
- Kundel HL, Polansky M. 2003. Measurement of observer agreement. *Radiology* 228:303–8.
- Olsen AL, Smith VJ, Bergstrom JO, et al. 1997. Epidemiology of surgically managed pelvic organ prolapse and urinary incontinence. *Obstet Gynecol* 89:501–6.
- Sampelle CM, Brink CA, Wells TJ. 1989. Digital measurement of pelvic muscle strength in childbearing women. *Nurs Res* 38:134–8.
- Singh K, Jakab M, Reid WM, et al. 2003. Three-dimensional magnetic resonance imaging assessment of levator ani morphologic features in different grades of prolapse. *Am J Obstet Gynecol* 188:910–5.
- Strohbehn K, Ellis JH, Storhbehn JA, et al. 1996. Magnetic resonance imaging of the levator ani with anatomic correlation. *Obstet Gynecol* 87:277–85.
- Tunn R, Paris S, Fischer W, et al. 1998. Static magnetic resonance imaging of the pelvic floor muscle morphology in women with stress urinary incontinence and pelvic prolapse. *Neurourol Urodynam* 6:579–89.
- Worth AM, Dougherty MC, McKey PL. 1986. Development and testing of the circumvaginal muscles rating scale. *Nurs Res* 35:166–8.