



Published in final edited form as:

*Female Pelvic Med Reconstr Surg.* 2018 ; 24(4): 295–300. doi:10.1097/SPV.0000000000000485.

## Comparing 3D Ultrasound to 3D MRI in the Detection of Levator Ani Defects

Camille S Calderwood, M.D.<sup>1</sup>, Amy Thurmond, M.D.<sup>2</sup>, Amanda Holland, M.P.H.<sup>1</sup>, Blake Osmundsen, M.D.<sup>3</sup>, and W. Thomas Gregory, M.D.<sup>1</sup>

<sup>1</sup>Division of Urogynecology and Reconstructive Pelvic Surgery, Department of Obstetrics and Gynecology, Oregon Health & Science University

<sup>2</sup>Womens Imaging & Intervention, Lake Oswego, OR

<sup>3</sup>Legacy Good Samaritan Medical Center, Portland, OR

### Abstract

**Objective**—To compare the detection of levator ani defects (LAD) between 3D ultrasound (US) and 3D MRI.

**Methods**—This is a secondary analysis of the Pelvic Floor Nerve Injury following Childbirth Study (NIPP). Nulliparous women underwent a standardized protocol of pelvic floor evaluations between January 2008 and December 2013, prior to *pregnancy* (V1) and at 2 points postpartum: 6 weeks (V2) and 6 months (V3). Those women who underwent a high-resolution 3D MRI pelvic floor sequence were selected. Comparisons were made to concomitantly acquired 3D perineal US. Eight tomographic slices were examined in the axial plane, each side independently scored with a 0 (no defect) or 1 (defect). A similar tomographic approach was applied to the MRI. For both MRI & US, the right (R) and left (L) sides were each scored. A total score of 0–8 was given to each side. A dichotomous variable “Complete LAD” was defined. Cohen’s kappa was used as a measurement of agreement of “complete LAD” between MRI and US. Kendall’s tau-b was used to correlate total scores.

**Results**—On the right side, 80/89 (90%) pairs were in agreement (concordant in the diagnosis or not of a “defect”). On the left side, 72/89 (81%) pairs were in agreement. Correlation (Cohen’s kappa) of complete LAD was 0.65 ( $p < 0.001$ ) on the right and 0.37 ( $p < 0.001$ ) on the left. Correlation of total scores was 0.47 ( $p < 0.001$ ) on the right, and 0.41 ( $p < 0.001$ ) on the left.

**Conclusion**—Moderate agreement was found between 3D US and 3 D MRI LAD detection. More LADs and discordance were seen on the left.

### Introduction

Pelvic injury after vaginal childbirth is as ubiquitous as sports injuries, with MRI studies reporting that up to 26% of women sustain pelvic muscle injuries after a 1<sup>st</sup> vaginal delivery.

<sup>1-4</sup> Injuries to connective tissues, levator ani muscles, and nerves, lead to weakening of the pelvic floor and a predisposition for pelvic floor disorders (PFD) such as urinary incontinence, fecal incontinence and pelvic organ prolapse. <sup>5, 6</sup> A current area of research interest is the relationship of levator ani “defects” and their relationship to PFD with US and MRI evaluating cohorts of women separately for the presence of these defects. However, US and MRI have not been compared to each other in their ability to detect levator ani defects in the same women.

Magnetic resonance imaging (MRI) and 3-dimensional ultrasound (3D US) of the pelvic floor are the predominant radiological modalities used to demonstrate the morphological changes of the pelvic floor that lead to PFD<sup>7-9</sup>. Initially, 2D MRI was the imaging method of choice for pelvic floor imaging due to its superior spatial resolution and ability to differentiate muscle groups <sup>9-10</sup>. However, as ultrasound technology has evolved, particularly 3D (and 4D) ultrasound’s convenient, multiplanar imaging capacity and superior temporal resolution, it has contributed to a growing volume of work with this technology <sup>11-15</sup>.

To date, very few studies have compared the findings of ultrasound and MRI modalities in the same patient. Three dimensional US and 2D MRI have been compared for their ability to measure the levator hiatus<sup>11</sup>. However, 2D MRI has two drawbacks: it can require laborious 3D rendering by manual segmentation, and, on any given axial image, it fails to capture the full extent of the origin-insertion plane of the puborectalis muscle, a consequence of its axial-to-body slicing technique<sup>11</sup>. In contrast, the readily accessible use of post-processing imaging software in ultrasound has provided an intimate look at the distal constituents of the levator ani complex. The combination of the wider use of 3D MRI in the evaluation of the pelvic floor as well as high speed post-processing computer software, allows for high resolution (in all planes) multiplanar reconstruction, volumetric rendering, image manipulation, and 3D reconstruction. These features significantly enhance the efficiency in evaluating the complex morphology of the components of the pelvic floor.

The primary aim of this study was to correlate 3D MRI and 3D ultrasound in their ability to detect levator ani defects (LADs).

## Methods

This is a secondary analysis of the IRB approved Pelvic Floor Nerve Injury following Childbirth Study (NIPP). NIPP is a longitudinal prospective cohort study of nulliparous women who were planning pregnancy (i.e. the woman was not yet pregnant). Subjects enrolled in NIPP underwent a standardized protocol of evaluations during three different visits. The first visit (V1) was prior to pregnancy, and, for those who became pregnant (just over 50%) the second (V2) and third (V3) visits were at six weeks and at six months postpartum, respectively. As part of the evaluations, subjects underwent pelvic imaging with both 3D US and MRI, at each visit. We had previously evaluated and reported on the comparison of 2D MRI acquisition to 3D Ultrasound as it applied to levator hiatus biometric measurements, utilizing 2 axial angles of acquisition to assure appropriate comparisons. During the course of the study we began utilizing 3D MRI acquisition sequences rather than

2D MRI. This secondary analysis thus compares the detection of levator ani defect between 3D US and 3D MRI for those subjects in whom these two image modalities were both obtained at the same visit, and does not represent an overlap of our previous work.

### Masking of Examiners

The examiners (CSC and WTG) were masked to visit time, the parous state of the subject, and the mode of delivery. Masking was achieved by first recoding the study subject's image files into random non sequential numbers. This was done separately for the ultrasound and the MRI files. Second, the ultrasound and MRI studies were analyzed at different time periods with the results from each modality unknown to the examiner at the time of review and analysis. Each image file was separately reviewed by each examiner. If there were disagreements, the examiners reviewed the image together and reached a consensus.

### Image Acquisition

**Three D US was acquired** with the participant in the supine position with an empty bladder. The ultrasound transducer was covered with a sheath and applied to the perineum in the midsagittal plane. Proper positioning was verified by identification of the symphysis pubis, urethra, bladder and anal canal. Images were acquired with an acquisition angle of 85 degrees, maximum field of view of 70 degrees in the sagittal plane and depth of 6.8 cm. Three dimensional US images were captured at rest, and after verbal instruction, were captured at maximal pelvic contraction and at maximal valsalva. One optimal "datacube" was saved for each maneuver for later analysis. Ultrasound examinations were acquired by one investigator (W.T.G), and performed using GE Voluson 730 (BT05) or GE Voluson e (BT12) equipped with RAB 4–8 MHz curved array 3D transducer (GE Medical Systems, Zipf, Austria).

**Three D MRI was acquired** using a Siemens 3T Trio Tim and a body coil (Siemens Medical Solutions, Malvern, PA). They were acquired using either an isotropic proton density sequence or an isotropic T2 weighted sequence. The subject lay supine in the MRI and, for this set of high-resolution sequences, was at rest.

### Anatomic Review and Definitions

In both, MRI and US images, levator ani disruptions were defined as an interruption in the continuity of the muscle, a detachment of levator ani from the inferior pubic ramus, and/or abnormal insertion of levator ani.

**US measurements** were done using GE 4D View (version 10.5 BT12). Using 4D View software, the ultrasound volumes (rest images) were displayed in the midsagittal, coronal, axial views, and in a rendered volumetric view. The rendered view had a slab thickness equal to the areas selected in the three planes. We oriented the image so as to follow the puborectalis muscle plane. Once the rendered image was in this plane of minimal hiatal dimension, we utilized the tomographic ultrasound imaging feature of the software to independently assess multiple slices. These tomographic images were assessed at 2.5 mm slice intervals from 5 mm below to 12.5 mm above the plane of minimal hiatal dimension

(Figure 1), producing eight slices in the axial orientation as seen in Figure 2. The third slice at 0 mm was considered the reference slice.

After applying standard post-processing enhancements of hues and contrasts to improve conspicuity, we analyzed each image, right side separate from the left side. For each of the eight images, a score of 0 or 1 was assigned, with “1” denoting a levator ani disruption; these scores were added to provide a total score for each side, ranging from 0 to 8. Zero was equivalent to an intact levator ani in all eight frames and eight to an obvious disruption in all eight frames.

**MRI measurements** were done using Osirix Lite (Pixmeo SARL, Switzerland), an advanced open source PACS workstation DICOM viewer, which is an application for Macintosh computers. To maximize comparability, a tomographic approach, that closely resembled how the ultrasound images were analyzed, was chosen. Using the 3D Multiplanar Reconstruction (MPR) feature, the midsagittal image was selected, and the plane of minimal hiatal dimension determined. Then, 24 slices at 0.8 -1 mm (the slice thickness of the source MRI) intervals, from 12 mm above and 12 mm below the plane of minimal hiatal dimension, were selected for simultaneous viewing in the axial plane as depicted in Figure 3. As with US, we improved conspicuity by adjusting the Color Look Up Table (CLUT) to “Hot Green”. A 24 frame “cineloop”, of the axial views corresponding to each sagittal slice selected, was subsequently made. For each side, 3 contiguous frames were analyzed as an entity (matching the 2.5 mm thickness of the ultrasound slices).

The right and left sides were scored separately. A levator ani disruption seen in any of the three frames was considered a positive muscle disruption in that section for that side. A value of 0 was given to a section if there was no disruption and a value of 1 was given if there was an identifiable disruption. The range of possible total scores for each side was 0–8 with a minimum total score of 0, if no disruptions were seen in any of the sections, or a maximum total score of eight if a disruption was seen in all sections.

**Definition of “Defect” (LAD)**—In keeping with prior work in this area,<sup>15</sup> we required a minimum score of 3 (out of 8) for it to be designated a “defect”. Additionally, the abnormal slices needed to be contiguous. Two definitions were created to assign minimal requirements for the presence of a complete levator ani defect (LAD):

- a. The “standard” definition, requiring the three central slices to be abnormal (the reference slice and the two contiguous cephalad slices) Figure 2.
- b. An “alternate” definition, requiring any three contiguous slices to be abnormal.

### Statistical Considerations

In our data analysis we correlated these two variables between 3D MRI and 3D US, assigning a unilateral disruption total score of the tomographic sliced images (ranging from 0–8), and the presence or absence of a complete LAD (0=no avulsion, 1=avulsion). Statistical analysis was done with IBM SPSS statistics v22. The total scores, for each side, were correlated with Kendall Rank coefficient (a measure of association between two ordinal variables). Cohen’s kappa (intra-rater correlation measure) was used as a

measurement of agreement between MRI and US complete levator ani defect results. Lastly, because we planned at the outset to analyze all the image pairs that had been obtained, we calculated that 90 pairs of images would achieve 80 % power to detect a kappa of at least 0.32 (given a null hypotheses kappa of no correlation) with an alpha of 0.01 (multiple testing adjustment).

## Results

Over the course of the 6-year project, 135 women were enrolled, from which 281 ultrasound/MRI pairs were collected. Due to the later initiation of 3D MRI technology for this project, we obtained 90 ultrasound/MRI pairs in which 3D MRI was utilized. One of these pairs had an uninterpretable 3D US; this left 89 pairs of 3D MRIs and 3D US's available for comparison. The 89 pairs of images came from 57 individual subjects. The 89 pairs of images were a collection of concomitant matched 3D MRI and 3 D US images obtained from participants at the same visits. There were a total of three different visits in this study, thus several of the 57 participants provided more than one pair of images for analysis.

Participants' average age was 30.8 (SD 3.7) years, and mean BMI 24.7 (SD 5.2) kg/m<sup>2</sup>. Eighty-six percent were Caucasian. There were 37 vaginal deliveries, 4 operative vaginal deliveries, 11 cesarean deliveries (5 stage 1, 6 stage 2). Five of the 57 subjects did not achieve a pregnancy during the allotted time of the study. Since 3 D MRI was introduced once the study had already begun, unsurprisingly we had less subjects with 3 D MRI from V1. Of the 89 image pairs, 7 were from V1, 39 from V2 and 43 from V3. Of the 7 V1 patients, 2 went on to V2 and V3, and of the 39 V2 visit, 32 went on to have a V3.

### Comparison of Imaging Modalities for “Total Scores”

On MRI, the right and left sided mean total scores (of the 8 evaluated slices observed to be disrupted) were 1.5 (SD 2.7) and 1.3 (SD 2.5) respectively. On US, the right and left sided mean total scores were 1.3 (SD 2.7) and 2.1 (SD 3.0) respectively. The median score for both sides and both modalities was 0. The frequency distribution of these disruptions is on Table 1. The correlation of the total scores obtained was low to moderate with a Kendall's tau-b score of 0.46 ( $p < 0.001$ ) on the right and 0.41 ( $p < 0.001$ ) on the left.

### Description of Diagnosis of “Complete LAD”

On MRI assessment, from the pool of all 89 image sets, we observed 16 (18%) defects on the right side and 13 (15%) on the left side using the “standard” definition for complete LAD. Using the same “standard” definition for US, we observed 15 (17%) defects on the right side and 20 (22.5%) defects on the left. To provide a useable comparison for the “incidence” of defects after delivery, we isolated the V3 time point visit. From the 43 V3 pairs, 9 (21%) MRI, and 7 (16%) US LAD were seen on the right with 8 (19%) MRI and 9 (21%) US LAD seen on the left. The proportion of LAD for each visit group, and for each modality is described in Table 2.

### Comparison of Imaging Modalities for Diagnosis of “Complete LAD”

On the right side using the “standard” definition for complete LAD, we had 80/89 (90%) pairs that were in agreement (concordant). On the left side, using the standard definition for complete LAD, we had 72/89 (81%) pairs in agreement. Using the “alternate” definition, our concordances were 77/89 (83%) on the right and 65/89 (73%) on the left.

The correlation of complete LAD was additionally evaluated using Cohen’s Kappa, which accounts for chance agreement. The summary of agreement of LAD between 3D MRI and 3D US, as calculated by a Cohen’s kappa, is described in table 3.

### Discussion

In this study, we correlated the detection of LAD with 3D MRI and 3D US. We found that on the right side 90% of the observations matched, whereas on the left, only 81% matched. Using Kappa Correlation coefficients, we found moderate agreement between 3D US and 3D MRI LAD detection. Regardless of the complete LAD definition utilized, 3D MRI and 3D US were not fully equal in the detection of LAD. Using the previously described “standard” definition<sup>15</sup>, the correlation was highest on the right side, kappa 0.65 (vs. kappa 0.37 on the left). In general, using ultrasound, we detected more defects than when using MRI.

We created an “alternate” definition for the variable of “complete LAD”. Our purpose in doing this was to create flexibility from the “standard” definition. The “standard” definition of complete LAD is based on previously described mathematical modelling in which the 3 central slices are required to be abnormal.<sup>15, 17</sup> Tomographic slicing is dependent on post-imaging processing, in which the area of interest is manually selected. The area selected is then sliced in 2.5 mm increments. A small deviation of area selection, may place a defect on slices outside the three central slices, thus we focused the “alternate” definition on any three contiguous abnormal slices. Interestingly, a slightly lower correlation of complete LAD was found with the more flexible definition of any 3 contiguous slices than with the “standard” definition.

We were perplexed by the finding of more LADs (as well as discordance) on the left side. More left sided defects were detected by ultrasound. This is in contrast to Dietz’s (Ultrasound Obstet Gynecol 2011; 37: 723–726) finding of a greater number of right sided defects when unilateral avulsion was noted<sup>18</sup>. Interestingly, the finding of left sided avulsions (and discordance) was also observed in the 7 pairs of images evaluated from V1, women who were nulliparous. Ostensibly, the muscular attachment to the pubic bone, and arcus tendineus fascia pelvis more proximally, should be intact in a nulliparous woman. So the natural conclusion would be that the ultrasound diagnosis is incorrect. However, given the lack of a true gold standard, we cannot be sure, so it’s possible that the US over-diagnosed a defect, or the MRI underdiagnosed a defect.

Both imaging modalities are associated with potential artifacts. While generally thought to be high resolution, isotropic (3D) MRI requires several minutes of acquisition time, during which, patient movement can affect resolution. Unlike 2D MRI though, isotropic sequences

are subject to much less chemical shift artifact which has been blamed for asymmetric measurements in the frequency encoding direction.<sup>19</sup>

With respect to ultrasound, the senior author, obtained all the ultrasound images for this evaluation with his left hand (working the ultrasound keyboard with his right hand). He made every effort to reduce any artifact, including avoiding any air in the probe cover, assuring sufficient gel for acoustic window throughout the sector sweep. However, there may have been undetected shadowing artifact, for example, that limited the ability to interpret the images optimally. In a repeated study, we would ensure a greater proportion of masked nulliparous women to further investigate this finding, including determining whether this were also to be found on the right.

With respect to the level of agreement, only 2 individuals evaluated the images. Both examiners each independently examined each fully masked data set, and provided his/her own score for each tomographic slice. The data points were compared, and where there were disagreements, we reached consensus. Although we did not perform a separate, formal, interrater reliability for each imaging modality separately because that has been done previously<sup>7, 20, 21</sup>, the proportion of disagreements in assessment was less than 10% of the tomographic slices for both MRI and ultrasound. It is feasible, that had more individuals reviewed the images from each modality, a different consensus, and therefore different Cohen's Kappa would have been achieved.

We were curious to see what happened to the degree of agreement of complete LAD as the total scores increased. We wanted to know if images were more likely to agree with each other as the defect spanned over more tomographic slices. We chose to perform a logistic regression to determine if odds of agreement for complete LAD between 3D MRI and 3D US indeed increased as total scores increased. In order to proceed with this regression, we had the option to choose the total scores of either MRI or US. Arbitrarily we chose the total scores of the MRI. We expected to find a greater degree of agreement as total scores increased, however, on both right and left sides, odds of agreement decreased as total scores increased: OR 0.83 (0.69, 1.0  $p < .04$ ) left, OR 0.76 (0.61, 0.94  $p < .01$ ) right. We then performed a logistic regression of only those subjects who had a complete LAD. Because a minimal total score of 3 was needed to have a complete LAD, those subjects with total scores of 0, 1 and 2 were removed from the calculations. However, when we removed all the patients with total scores of 0, 1 and 2, we reduced the numbers such that we could not perform a robust analysis.

The principal strength of this study is that this is one of the few correlations of 3D US to 3D MR in detecting LADs. Previous studies comparing ultrasound to MRI, have chosen MRI to be a gold standard<sup>22</sup> or used a proxy as evidence of a defect<sup>23</sup>. Additionally, our study focuses on the peripartum time period, when mounting evidence suggests that this is when the injury occurs. Furthermore, this is the first study that applies the easy-to-use 3D US technique (of measuring and grading the puborectalis muscle, along the plane of minimal hiatal dimension) to 3D MRIs. Achieving this was made significantly faster and easier by using the Dicom Viewer Osirix Lite. Historically, comparing MRI and US to each other, for the detection of levator ani defects, has been difficult. This difficulty is because different

classification systems (DeLancey's system for MRI and Dietz's system for 3D US), have traditionally been used to grade the observed defects. In previous study by Notten et al., 3D US was compared to MRI in the detection of levator ani defects, using MRI as the reference standard and applying DeLancey's scoring system. Further strengths of this study include, the prospective design, the masking of the examiners, and the reduction of technique variability by having one examiner perform all the ultrasounds. Lastly, the comparison of imaging modalities "Total Scores" and "Complete LAD" were sufficiently powered, and the probability of finding a result as extreme as we did, if there was no correlation between our imaging modalities, was 0.001. Based on these results, we rejected the hypothesis that there was no correlation among the imaging modalities.

This study also had some limitations. We had to compare the resting 3D volumes to the resting 3D MRI, where ideally we would be able to compare contracted volumes of both modalities. Contracted muscles have an increased ratio of muscle cells per extracellular matrix (fat and collagen), which in turn can create a more conspicuous ultrasound image. Authors have suggested that levator ani defect might be improved by analysis of contracted muscle volumes.<sup>24</sup> Comparison of contracted muscle volumes might have provided a different level of correlation. Another limitation is that the majority of our study were in individuals without levator injuries. The prevalence of levator ani defects in this group of women was consistent with previous investigations.<sup>24-25</sup> We also did not have enough subjects in the operative vaginal deliveries or low transverse cesarean groups to do any sub-analysis. We also point out that the distance of axial slices analyzed on US was 20 mm, and 24 mm for MRI, but one might expect this discrepancy to result in more observed defects on MRI, and this was not the case. Further limitations, are the absence of a gold standard in the detection of levator ani injuries, if a gold standard existed, we could calculate conditional probabilities for both MRI and US. Notten et al. used MRI as a reference standard, allowing investigators to calculate sensitivity, specificity, predictive values and likelihood ratios. Some authors support that the superior resolution of MRI, is enough to declare it the gold standard, however to our knowledge this has not been accepted yet as a standard. The validity of declaring MRI as the gold standard, due to its excellent resolution and beautiful images, could be argued. Very few imaging modalities findings have been validated with cadaveric anatomic studies (which is considered the gold standard). A recent multicenter study by Da Silva et al., through a cadaveric study, aimed to correlate ultrasound findings of levator ani avulsions with anatomic findings. A difference between anatomical and US findings was found, the authors found that avulsions seen on ultrasound were not considered to be true avulsions, and recommended that the term "levator ani defect or damage" be used instead<sup>26</sup>. In the absence of a robust body of literature correlating imaging findings with anatomic dissection findings, our study did not set to find out which was a better imaging modality, but rather how did the modalities correlate to each other in the findings of levator ani disruptions.

In summary, for the evaluation of the distal pelvic floor muscles, specifically the puborectalis, we believe that current 3D US technology and current 3D MRI sequences with multiplanar reconstruction have acceptable correlation to one another. Certainly, more comparative work needs to be done using these two modalities particularly in women who



have pelvic floor disorders, but at this time, it appears that the modalities can be reliably used interchangeably.

## Acknowledgments

### Source of Support

N/A

### Source of Funding

National Institute of Child Health & Human Development Grant R01 HD049541

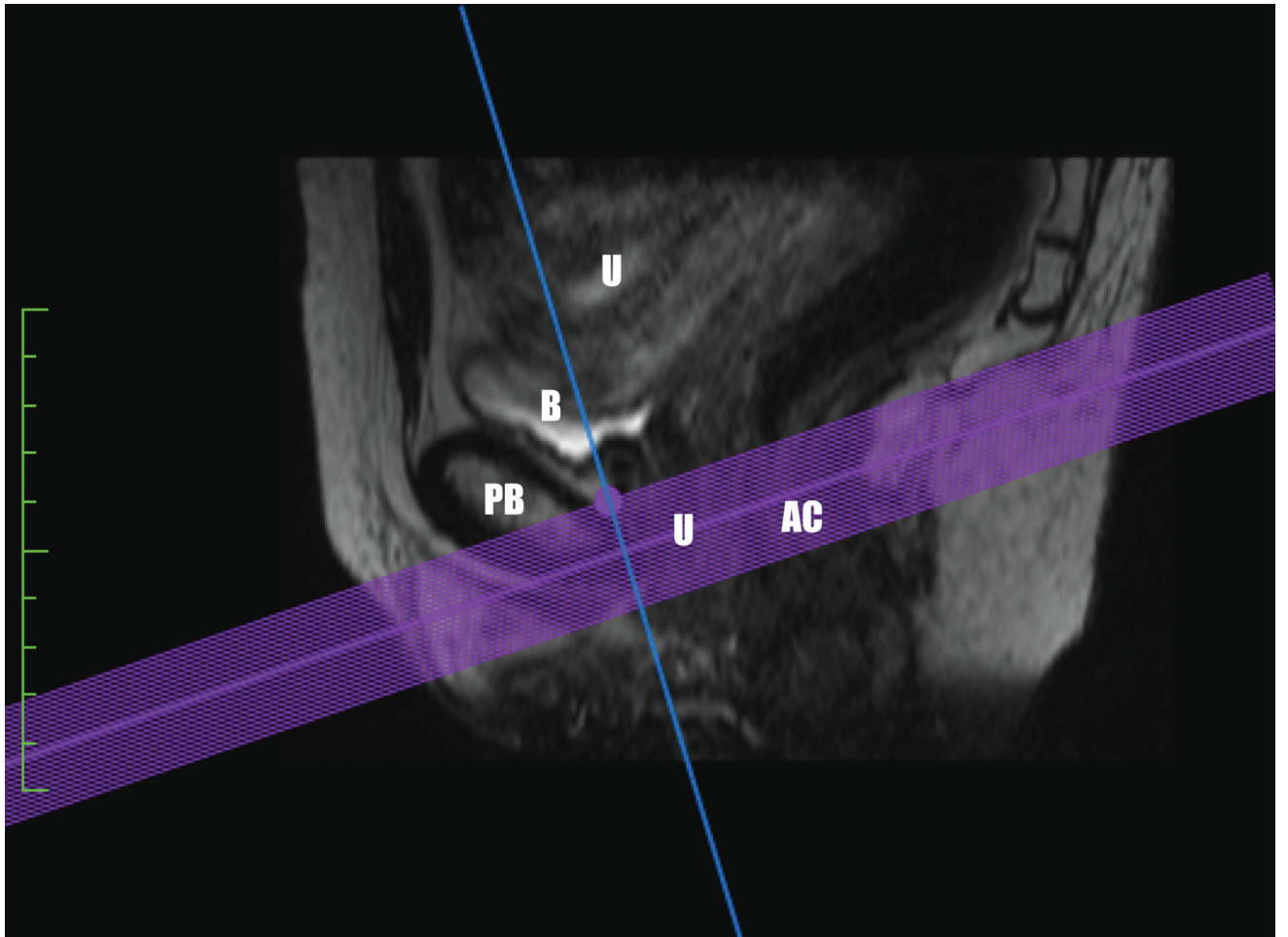
## Abbreviations

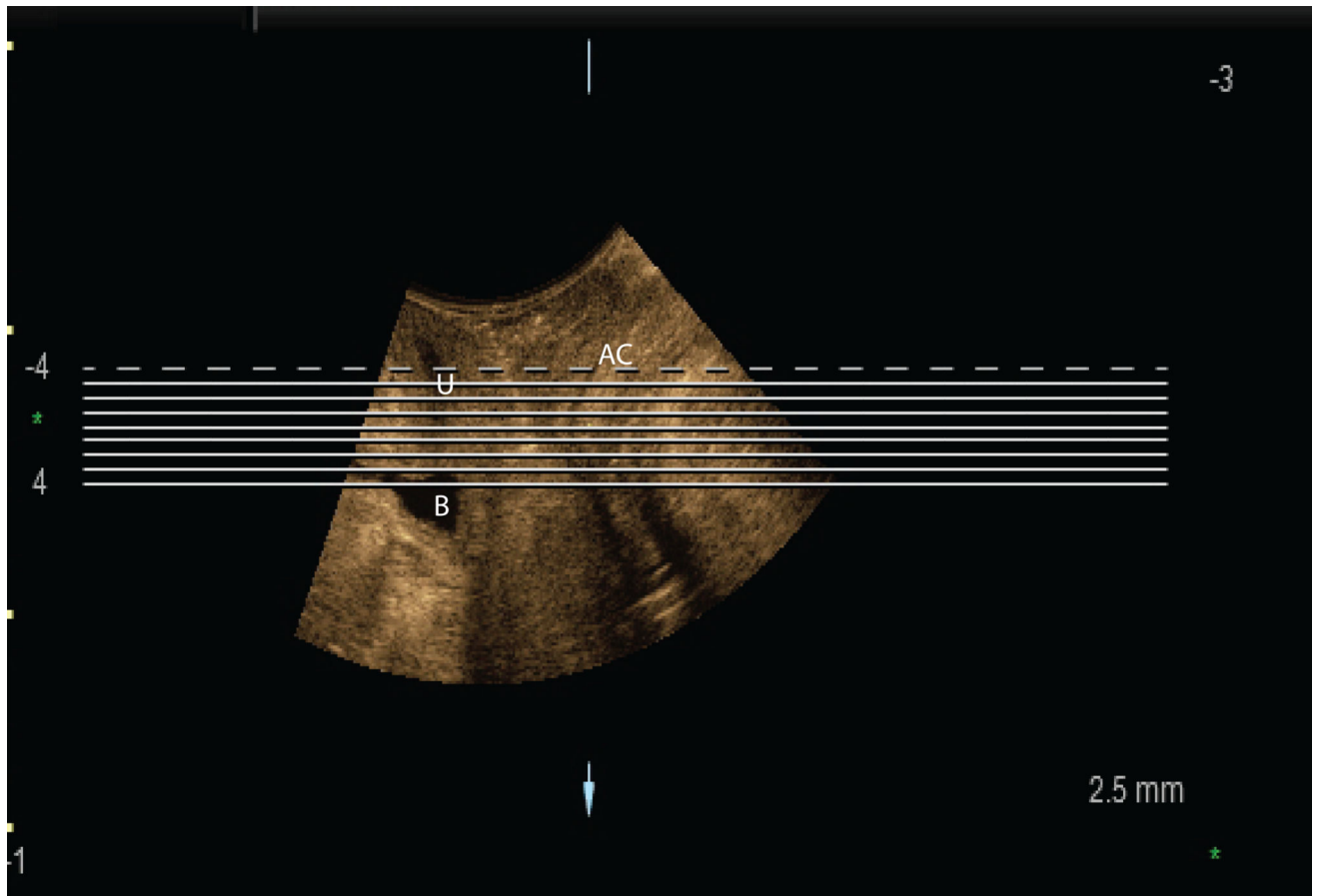
<b>PFD</b>	Pelvic floor disorders
<b>LAD</b>	Levator Ani Defect
<b>LAD</b>	Levator Ani Avulsion
<b>3D</b>	Three Dimensional
<b>US</b>	Ultrasound
<b>MRI</b>	Magnetic Resonance Imaging
<b>V1</b>	Visit one
<b>V2</b>	Visit two
<b>V3</b>	Visit three

## References

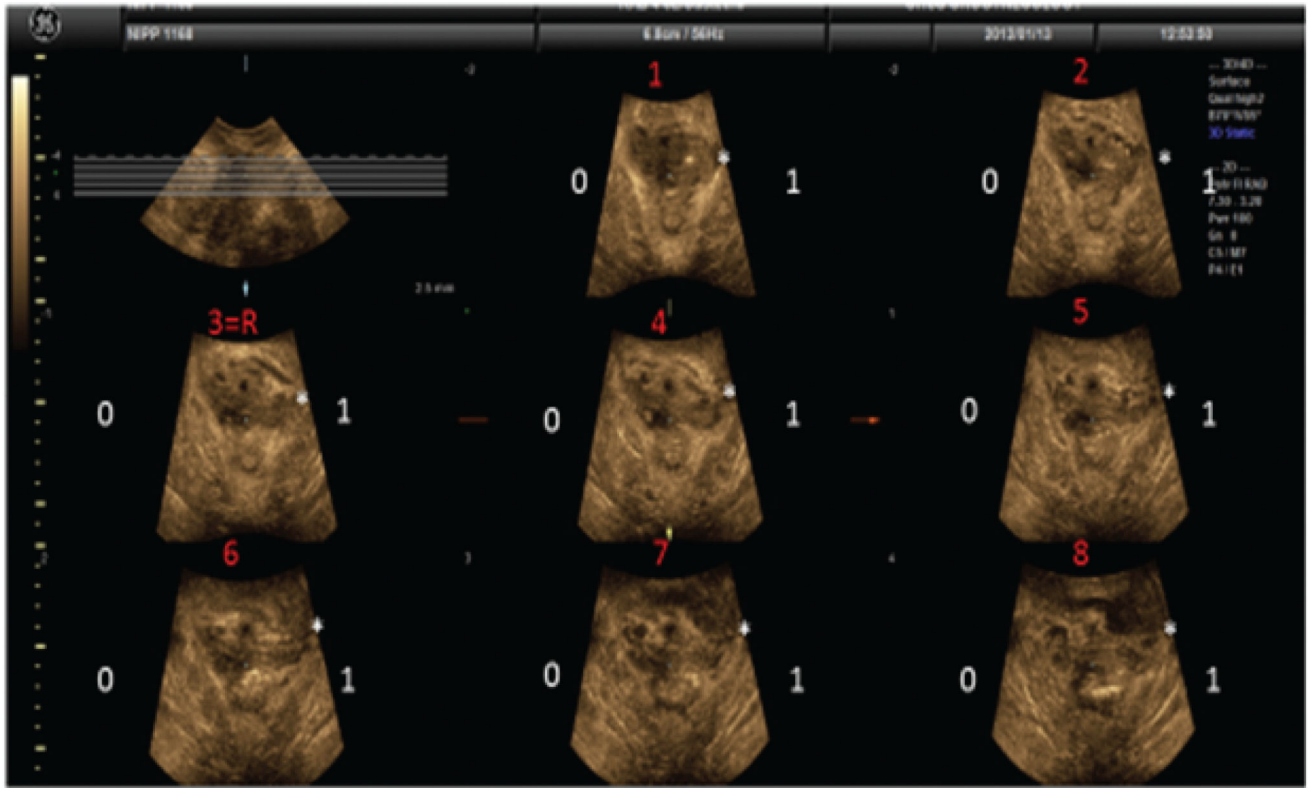
1. Viktrup L, Rortveit G, Lose G. Risk of stress urinary incontinence twelve years after the first pregnancy and delivery. *Obstet Gynecol.* 2006; 108:248–254. [PubMed: 16880292]
2. DeLancey JO, Kane Low L, Miller JM, et al. Graphic integration of causal factors of pelvic floor disorders: an integrated life span model. *Am J Obstet Gynecol.* 2008; 199:610.e1–610.e5. [PubMed: 18533115]
3. DeLancey JO, 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. [PubMed: 12517644]
4. Dietz HP, Lanzarone V. Levator trauma after vaginal delivery. *Obstet Gynecol.* 2005; 106:707–712. [PubMed: 16199625]
5. Bump RC, Norton PA. Epidemiology and natural history of pelvic floor dysfunction. *Obstet Gynecol Clin Nort Am.* 1998; 25:723–46.
6. Fielding JR, Dumanli H, Schreyer AG, et al. MR-based three dimensional modeling of the normal pelvic floor in women: Quantification of muscle mass. *Am J Roentgenol.* 2000; 174:657–60. [PubMed: 10701604]
7. Dietz HP. Quantification of major morphological abnormalities of the levator ani. *Ultrasound Obstet Gynecol.* 2007; 29(3):329–334.
8. Dietz HP, Abbu A, Shek KL. The lavator-urethra gap measurement: a more objective means of determining levator avulsion? *Ultrasound Obstet Gynecol.* 2008; 32(7):941–945. [PubMed: 19035543]

9. Strohbehn K, Ellis JH, Strohbehn JA, et al. Magnetic resonance imaging of the levator ani with anatomic correlation. *Obstet Gynecol.* 1996; 87(2):277–285. [PubMed: 8559539]
10. Macura KJ. Magnetic resonance imaging of the pelvic floor defects in women. *Top Magn Reson Imaging.* 2006; 17(6):417–426. [PubMed: 17417089]
11. Nardos R, Thurmond A, Gregory WT, et al. Pelvic Floor Levator Hiatus Measurements: MRI versus Ultrasound. *Female Pelvic Med Reconst Surg.* 2014; 20(4):216–221.
12. Dietz HP. Ultrasound imaging of the pelvic floor. Part I: two dimensional aspects. *Ultrasound Obstet Gynecol.* 2004; 23:80–92.
13. Dietz HP. Ultrasound imaging of the pelvic floor. Part II: three-dimensional or volume imaging. *Ultrasound Obstet Gynecol.* 2004; 23:615–625. [PubMed: 15170808]
14. Dietz HP, Shek C, Clarke B. Biometry of the pubovisceral muscle and levator hiatus by three-dimensional pelvic floor ultrasound. *Ultrasound Obstet Gynecol.* 2005; 25:580–585. [PubMed: 15883982]
15. Dietz HP, Shek KL. Tomographic ultrasound of the pelvic floor: Which levels matter most? *Ultrasound Obstet Gynecol.* 2009; 33:698–703. [PubMed: 19434620]
16. Gregory WT, Nardos R, Worstell T, et al. Measuring the levator hiatus with axial MRI sequences: adjusting the angle of acquisition. *Neurourol Urodyn.* 2011; 30(1):113–6. [PubMed: 21046654]
17. Dietz HP. Pelvic floor trauma in childbirth. *Aust N Z J Obstet Gynaecol.* 2013; 53:220–230. [PubMed: 23452259]
18. Diez HP, Bhalla R, Chantarasorn V, et al. Avulsion of the puborectalis muscle is associated with asymmetry of the levator hiatus. *Ultrasound Obstet Gynecol.* 2011; 37(6):723–6. [PubMed: 21351178]
19. Zand KR, Reinhold C, Haider MA, et al. Artifacts and pitfalls in MR imaging of the pelvis. *J Magn Reson Imaging.* 2007; 26(3):480–97.
20. Morgan DM, Umek W, De Lancey JO, et al. Interrater reliability of assessing levator ani muscle defects with magnetic resonance images. *Int Urogynecol J Pelvic Floor Dysfunct.* 2007; 18(77): 773–8. [PubMed: 17043740]
21. Tan L, Shek KL, Atan IK, et al. The repeatability of sonographic measures of functional pelvic floor anatomy. *Int Urogynecol J.* 2015; 26:1667–1672. [PubMed: 26072127]
22. Notten KJ, Kluivers KB, Futterer JJ, et al. Translabial three-dimensional ultrasonography compared with magnetic resonance imaging in detecting levator ani defects. *Obstet Gynecol.* 2014; 124(6):1190–1197. [PubMed: 25415171]
23. Zhuang RR, Son YF, Chen ZQ, et al. Levator avulsion using a tomographic ultrasound and magnetic resonance-based model. *Am J Obstet Gynecol.* 2011; 205:232.e1–8. [PubMed: 21620359]
24. Dietz HP, Steensma AB. The prevalence of major abnormalities of the levator ani in urogynaecological patients. *BJOG.* 2006; 113(2):225–230. [PubMed: 16412002]
25. Kearney R, Miller J, Aston-Miller J. Obstetric factors associated with levator ani muscle injury after vaginal wall prolapse. *Obstet Gynecol.* 2006; 107:144–149. [PubMed: 16394052]
26. Da Silva AS, Digesu GA, Dell’Utri Chiara, et al. Do Ultrasound findings of levator ani “avulsion” correlate with anatomical findings: A multicenter cadaveric study. *Neurourol Urodyn.* 2016; 35:683–688. [PubMed: 25982354]

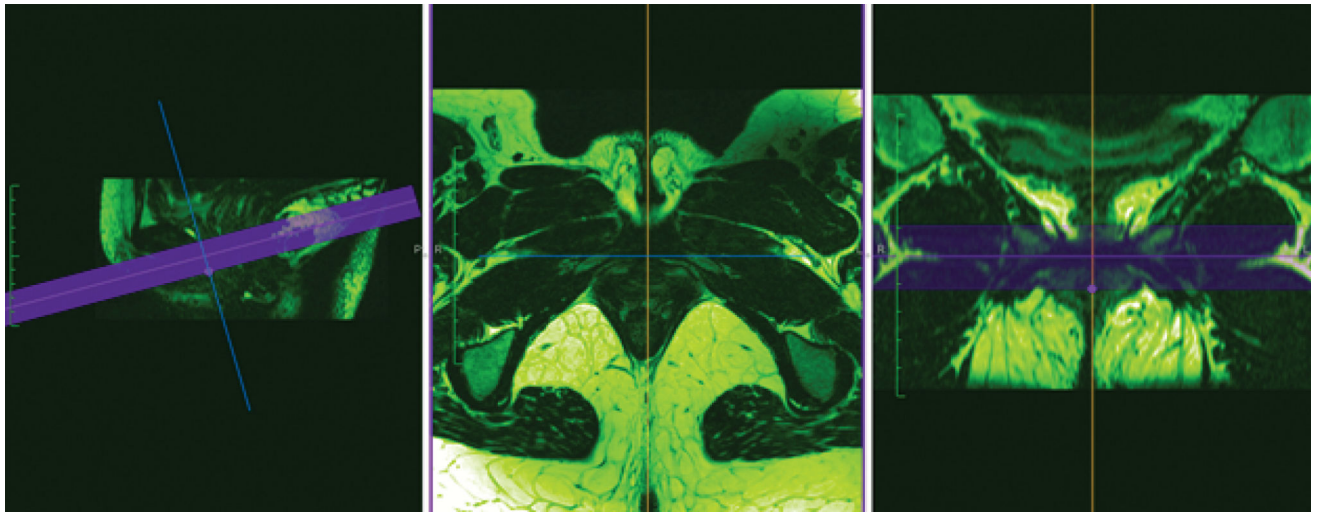




**Figure 1.**  
Plane of minimal hiatal dimension in the midsagittal plane as seen on a. MRI and b. US.  
Abbreviations: B bladder, PB-pubic bone, U-urethra, AC-anal canal.



**Figure 2.** Tomographic assessment of the puborectalis muscle. Tomographic slices set 2.5 mm apart, slice 3 is the reference slide. In this example a complete left sided puborectalis avulsion is observed.



**Figure 3.**  
MRI Tomographic selection in the midsagittal view, with an axial and coronal representation.

**Table 1**

Frequency of distribution of disruptions seen per side and per imaging modality.

Frequencies and Percentages				
	Total R MRI n (%)	Total L MRI n (%)	Total R US n (%)	Total L US n (%)
0	65(73)	65(73)	69(78)	55(62)
1	1(1)	1(1)	0	0
2	3(3)	4( 5)	2(2)	4(4)
3	2(2)	4(5)	3(3)	4(4)
4	3(3)	2(2)	1(1)	4(4)
5	2(2)	4(5)	2(2)	6(7)
6	5(6)	2(2)	2(2)	1(1)
7	0	0	3(3)	5(6)
8	8(9)	7 (8)	7(8)	10(11)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Complete levator ani defect for each visit group, and for each modality.

	3D MRI		3D US	
	R avulsions	L avulsions	R avulsions	L avulsions
Visit 1 (n=7)	0 (0%)	0 (0%)	0(0%)	2 (28 %)
Visit 2 (n=39)	7 (18%)	5 (12%)	8 (20%)	9 (23%)
Visit 3 (n=43)	9 (21%)	8 (20%)	7(16%)	9 (21%)
Totals (n=89)	16(17%)	13(15%)	15(17%)	20(22%)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



The summary of agreement of complete levator ani defect between 3D MRI and 3D US (Cohen's kappa).

**Table 3**

Summary of Agreement of LAD Detection between 3D MRI and 3D US					
	Concordant Intact	Concordant Disrupted	Discordant	Cohen's kappa	p-value
Standard Right	69	11	9	0.65	<0.001
Standard Left	64	8	17	0.37	<0.001
Alternate Right	63	14	12	0.53	<0.001
Alternate Left	52	13	24	0.35	<0.001