Pelvic floor muscle therapy, pessary wearing or both?

M.K. Schulten

September 2023

Pelvic floor muscle therapy, pessary wearing or both?

Evaluating the effect of pelvic floor muscle therapy additional to pessary wearing, compared to pessary wearing alone, for the treatment of pelvic organ prolapse by means of 4D transperineal ultrasound

Author
M.K. Schulten

Date
September 22nd, 2023

Graduation committee

Chairman
Prof. dr. M.M.P.J. Reijnen
University of Twente

Medical supervisor
Drs. A. van der Steen
Ziekenhuisgroep Twente

Technical supervisor
Dr. A.T.M. Bellos – Grob
University of Twente

Process supervisor
Dr. M. Groenier
University of Twente

External committee member
Dr. M.M.A. Kip
University of Twente

UNIVERSITY OF TWENTE.

*Front page image by Kaitlin Walsh of Lyon Road Art, URL: https://www.lyonroadart.com/collections/male-reproductive/products/pelvic-floor-watercolor-print (accessed 10-09-2023)

Abstract

Objective Limited studies evaluated if pelvic floor muscle therapy (PFMT) combined with pessary wearing (combined therapy) provides greater efficacy compared to pessary wearing alone (pessary therapy) in the treatment of pelvic organ prolapse (POP).[1–3] Besides, they only focused on biometric parameters and did not quantitatively assess the behavior and function of the muscle. Therefore, in this study muscle strain is utilized to evaluate the change in muscle function of the musculus puborectalis (PRM) after combined therapy, compared to pessary therapy, by means of 4D transperineal ultrasound (TPUS). Additionally, interobserver variability in PRM segmentation was investigated, to evaluate its impact on strain analysis.

Method A retrospective study was conducted including 17 women divided into two groups, the Combined therapy group (n = 7) and the Pessary therapy group (n = 10). 4D TPUS performed before and after treatment was used to assess the accumulated displacement estimates (cm), accumulated principal strain (%), and the mean principal strain (MPS; %) of the PRM. Besides, the area of the levator hiatus (LH) in rest and the difference in anteroposterior (AP) diameter of the LH between rest and maximum contraction are evaluated. Interobserver variability was assessed by the Dice Similarity Coefficient (DSC), average Hausdorff distance (HD) and 95% HD.

Results For the MPS, no significant difference is observed between first and follow-up ultrasound in both groups (Combined therapy, $p = 0.263$; Pessary therapy, $p = 0.489$). For the LH area, there was a significant increase for the Combined therapy group ($p = 0.036$), although a decrease was expected. No significant difference was observed for the Pessary therapy group ($p = 0.729$). For the AP-diameter the statistical test did not reveal a significant difference between first and follow-up ultrasound for both groups (Combined therapy, $p = 0.110$; Pessary therapy, $p = 0.610$). Furthermore, for all three parameters no significant difference was observed between the two groups (MPS, $p = 0.422$; LH area, $p = 0.475$; AP-diameter, $p = 0.165$), and no correlation existed between the parameters. However, acccumulated displacement estimates and accumulated principal strain images reveal an unexpected, highly variating pattern in muscle deformation. Therefore, it cannot be presumed that quantitative outcomes are accurate.

Segmentations of different observers showed moderate overlap in comparison metrics, but a variety in MPS values.

Conclusion A step has been made in the evaluation of the effectivity of PFMT additional to pessary wearing, compared to pessary wearing alone. However, the unexpected variability of the accumulated displacement estimates and accumulated principal strain, prevents the formulation of a conclusion. Therefore, this study highlights the need for further research with a prospective dataset and larger study population. Besides, strain software and image tracking need to be optimized. Automatic image segmentation via deep learning needs to be implemented, to prevent observer bias in PRM segmentation.

Table of contents

List of abbreviations

AP	Anteroposterior
cm	Centimetre
DSC	Dice Similarity Coefficient
EMG	Electromyography
HD	Hausdorff Distance
ICIQ	International Consultation on Incontinence Questionnaire
kg	kilogram
LH	Levator hiatus
LSQSE	Least-Squares Strain Estimator
m^2	Square metre
mm	Millimetre
mm^2	Square millimetre
MPS	Mean principal strain
PFDI-20	Pelvic floor disability index
PFMT	Pelvic floor muscle therapy
PGI-I	Global impression of improvement
POP	Pelvic organ prolapse
POP-Q	Pelvic organ prolapse quantification
POP-SS	Pelvic organ prolapse symptom score
PRM	Musculus Puborectalis
QoL	Quality of life
RCT	Randomized controlled trial
ROI	Region of interest
SD	Standard deviation
TPUS	Transperineal ultrasound

1. Introduction

Pelvic organ prolapse (POP) is a common condition in which the pelvic floor organs, i.e., the bladder, uterus and/or rectum, descend into or out of the vagina. Up to 50% of middle-aged women suffer from POP, with a prevalence increasing with age, vaginal birth and increasing parity.[4–7] Symptoms caused by prolapse, such as the feeling of a vaginal bulge, urine incontinence, sexual dysfunction and back pain, cause inconvenience and negatively impact quality of life, which makes adequate treatment important.[8]

Treatment of POP consists of surgery or conservative therapy, including pessary wearing and pelvic floor muscle therapy (PFMT). Conservative therapy is generally preferred as first-line treatment, because of its non-invasive and easily accessible character.[5,9] Depending on different factors, such as patient preference, the severity of the prolapse, and the functionality of the pelvic floor muscles, women may start with pessary wearing or PFMT, or a combination of those two therapies.[10]

However, not all women benefit from PFMT, and it is challenging to predict for which women the therapy is effective and for whom it is not. To predict the effectiveness of PFMT, it is crucial to understand the additional value of PFMT to pessary wearing. Limited studies evaluated if PFMT combined with pessary wearing provides greater efficacy compared to pessary wearing alone. Richter et al. (2010)[3] performed an exploratory study with 16 participants to investigate this topic. The results demonstrated a high degree of variability, although there was a minor advantage in the Combined therapy group in the outcome of the POP symptom score and quality of life. They recommended the initiation of a randomized controlled trial (RCT). However, the RCT was never implemented due to feasibility issues. Giroux et al.[2] published a literature review in 2020 about the effectiveness of pessary wearing in combination with PFMT, and pessary wearing and PFMT alone for the management of stress and mixed urinary incontinence. They reviewed ten studies, of which only one compared the effect of pessary wearing to PFMT and a combination of these therapies.[3] Despite this study being a large multicentre RCT (n=446), the researchers only conducted qualitative research to assess symptom reduction and patient satisfaction, using questionnaires.[11–13] Considering that POP is associated with pelvic floor muscle dysfunction[14], it is valuable to quantify muscle function to evaluate the effectiveness of conservative therapy.[9]

In recent years, there has been an increasing trend towards quantifying pelvic floor muscle function. Several methods, i.e., digital palpation, electromyography (EMG), and measurement of levator hiatus (LH) biometry through medical imaging, are used in literature to assess muscle function.[1,5,15–18] However, all these methods have limitations. Digital palpation measures the force of overall muscle group contraction rather than individual muscles, the results are affected by abdominal pressure, and it has a low interobserver reliability.[19] EMG records the activity of the muscle. Although electrical stimuli are crucial to generate force, electrical activity does not always indicate the occurrence of muscle contraction. Therefore, EMG alone does not provide a quantitative assessment of muscle function.[17,20] Biometric measurement provides a comprehensive assessment of changes in the shape or form of the muscle.[1,19] However, alterations in muscle length, volume, or hiatal diameter do not necessarily imply changes in muscular function. For instance, a decrease in muscle length may indicate a stronger muscle, but it can also be associated with muscle avulsion. Therefore, it is important to consider the extent to which these parameters provide comprehensive understanding of muscle function.

To develop a method that takes all these disadvantages into account, Das et al.[9] developed a 3D transperineal ultrasound (TPUS) strain algorithm which tracks the deformation of the musculus puborectalis (PRM) during contractility and relaxation. When a muscle undergoes stretching or compression, its shape changes, and the tissue experiences a certain level of deformation. The strain algorithm provides a quantitative value that reflects the percentage change in the length of a muscle compared to its original length and is therefore a good alternative in the quantification of pelvic floor muscle function.[21]

Therefore, the aim of this study is to quantify the effect of pessary treatment in combination with PFMT (Combined therapy) compared to pessary wearing alone (Pessary therapy) by utilizing muscle strain to evaluate the function of the PRM. This leads to the following research question:

What is the effect of combined therapy for the treatment of POP, compared to pessary therapy, on the strain of the PRM measured with TPUS?

The hypothesis is that both, combined therapy and pessary therapy, improve the ability to contract the PRM, and thus lead to a more negative strain. However, it is expected that combined therapy leads to a more substantial difference between the strain of the muscle in rest and the strain of the muscle during maximum contraction compared to pessary therapy.

To answer the research question, several sub-questions will be examined in this thesis:
1. What is the difference in PRM strain before and after treatment in patients with POP who had combined therapy versus patients receiving pessary therapy?
2. What is the difference in biometric parameters (LH area and LH-diameter) of the pelvic floor before and after treatment in patients with POP who had combined therapy versus patients receiving pessary therapy?
3. Is there a correlation between muscle strain and biometric parameters derived from TPUS? If so, what is the additional value of muscle strain compared to other parameters?
4. What is the interobserver variation in segmentation of the PRM and how does this variation impact strain calculation?

This thesis is divided into five chapters to answer the research question and sub-questions as described in the introduction (Chapter 1). The next chapters, i.e., Method, Results, Discussion and Conclusion, follow a uniform structure, beginning with muscle displacement and deformation (strain), followed by the biometric parameters, and concluding with the interobserver variability.

2. Method

This retrospective study aimed to quantify the effect of combined therapy compared to pessary therapy by evaluating the function of the PRM based on anatomical and functional (strain) 4D TPUS measurements. In this chapter, study design and population, data collection and (statistical) analysis are discussed.

2.1 Study design and population

For this retrospective study, participants' data were collected from an existing database. Patients underwent 4D TPUS measurements during their first visit at the Gynaecology department of the Bergman Clinics, Women's Health Care, Hilversum, the Netherlands between May 2018 and February 2020 (GYNIUS dataset). Participants were divided into two groups: the Combined therapy group and the Pessary therapy group.

Both groups included women who wear a pessary. Pessaries come in various types and sizes, with the most frequently used ones being the open ring, the ring with support, the Gellhorn, and the donut pessary.[22] The choice of ring type depends on the type and severity of prolapse, the size of the vagina, level of sexual activity, the patient's cognitive ability and manual ability.[10,22]

The women within the Combined therapy group received PFMT performed by a trained pelvic floor physiotherapist. PFMT starts with an explanation of the pelvic floor function and the importance of the pelvic floor muscles, followed by a physical examination to assess pelvic floor muscle function. Subsequently, the physiotherapist assists patients in awareness of their pelvic floor muscles, learning to contract the appropriate muscles, and learning to relax these muscles. This is achieved by guiding the patient through a program consisting of diverse exercises targeting proprioception, coordination, strength, endurance, and relaxation of the pelvic floor muscles.[10,15]

In addition, the pelvic floor physiotherapist provides explanation and instruction for regulating intra-abdominal pressure. These exercises, commonly referred to as "Knack exercises", involve the simultaneous contraction of the pelvic floor muscles while preventing the breath from being held during moments of increased intra-abdominal pressure, such as lifting, sneezing, and coughing.[23] Furthermore, physiotherapists offer advice on toilet habits and provide lifestyle recommendations regarding weight management and diet control.[12]

Despite existing PFMT protocols, treatment is unique for each patient. Effectivity of PFMT depends, among others, on therapy adherence, and therefore the length of the treatment trajectory and the frequency of the therapy sessions varies per patient.[10,24] For this study, the frequency is listed.

2.2 Data collection and analysis

Several studies have reported the possibility of muscle quantification by means of 4D TPUS.[1,9,19] This imaging modality is widely accessible and provides detailed morphological information, including changes in muscle volume, muscle length, and diameter of the LH during contraction. Therefore, in this study patients are included for whom 4D TPUS was performed during their first and follow-up visit at the Gynaecology department. The 4D TPUS images were retrieved with the Philips EPIQ 7G ultrasound machine in combination with a X6-1 matrix transducer (Philips Healthcare, Bothell, WA, USA) covered with a 2-cm thick gel pad. The 4D TPUS measurements consist of a series of 3D volumes (frames) acquired over time. During the measurement, the patient was asked to contract the pelvic floor muscles to its maximum.

For all ultrasound images, the PRM in rest position was manually segmented out of a 3D-volume, according to the article of van den Noort et al.[25], using 3D Slicer 5.2.1 (3D Slicer, Boston, Massachusetts, USA). Ultrasound recordings were available for all participants with the pessary *in situ*. However, recordings without a pessary were not available for all participants. Therefore, the decision was made to exclusively utilize ultrasound images with the pessary *in situ* for both segmentation and strain measurement purposes.

2.2.1 Accumulated displacement estimates and accumulated principal strain

The behaviour of the muscle is analysed using strain to quantify the function of the muscle. For this purpose, the 4D TPUS images and the segmentation of the PRM are employed. The steps involved in calculating strain included inter-volume displacement estimations, tracking, and strain calculations, which were performed using MATLAB R2019b (The MathWorks, Inc., Natick, MA, USA). The existing algorithm of Das et al. is utilized.[9]

First, the ultrasound volume and mask of the segmentation are used to estimate the inter-volume displacement, which means that the displacement of the segmentation between two consecutive frames is estimated. The two frames are divided into 3D blocks (kernels and templates), with the kernels matched with the templates to find cross-correlation in the next frame.

Subsequently, tracking is applied because the muscle moves and may exit the region of interest (ROI). Therefore, the ROI needs to be updated based on the estimated displacement from the previous step.

After finishing the inter-volume displacement estimates, the displacements are accumulated from rest position to maximum contraction. This outcome, the accumulated displacement estimates, are utilized to visually evaluate movement of the muscle during contraction.

Subsequently, strain calculations are performed. The 3D strain tensor is computed from the accumulated displacement estimates, using a 3D Least-Squares Strain Estimator (LSQSE).[9] This leads to individual strain values for the x-, y-, and z-directions, and thus three separate accumulated strain images.

To combine the strain of three directions into one single image, the principal component of the strain was calculated by the software. This represents the chosen direction in which the strain is the most significant and returns an accumulated principal strain image with its corresponding directions. An overview of the strain calculation steps is displayed in Figure 1.

2.2.2 Visual analysis

To explain the outcome of the accumulated displacement estimates and the accumulated principal strain, for each patient the original ultrasound recordings obtained during the first visit at the Gynaecology department, as well as during the follow-up visit were visually inspected. This evaluation encompassed the assessment of muscle contraction, classified in three categories: *good,* when the ultrasound video showed a clear contraction and a reduction of the minimal hiatal dimension; moderate, when there was a small contraction noticeable; or poor, when there was little to no movement of the pelvic floor muscles visible. Additionally, an evaluation was conducted concerning the precision of muscle tracking by the software. This was categorized as *accurate*, when the software seemed to track the PRM correctly; *moderately accurate*, when the software tracked the PRM partially correct; and *inaccurate* when the software tracked surrounding tissue instead of the PRM.

2.2.3 Mean principal strain

To retrieve a quantitative value that represents strain, the MPS value was calculated from the accumulated principal strain. The value is calculated by the software, taking the average of the strain of all muscle fibres. Because the magnitude of the MPS value differs for each person, the data is normalized by using the difference of the MPS value between resting state and maximum contraction. According to Das et al., a negative MPS value indicates contraction of the muscle region, while a positive value indicates elongation of the muscle.[9]

2.2.4 Biometric parameters

Besides muscle strain, biometric parameters are assessed to compare results with existing literature. Sainz-Bueno et al.[1] described in their article that PFMT attributes to shrinkage of the LH by hypertrophy of the pelvic floor muscles, which leads to a decreased LH area after PFMT. Volløyhaug et al.[19] measured the change in anteroposterior (AP) diameter of the LH between rest and maximum contraction to evaluate muscle function. They noticed a decrease in AP-diameter during contraction. Nevertheless, they did not evaluate the difference in AP-diameter before and after POP-treatment.

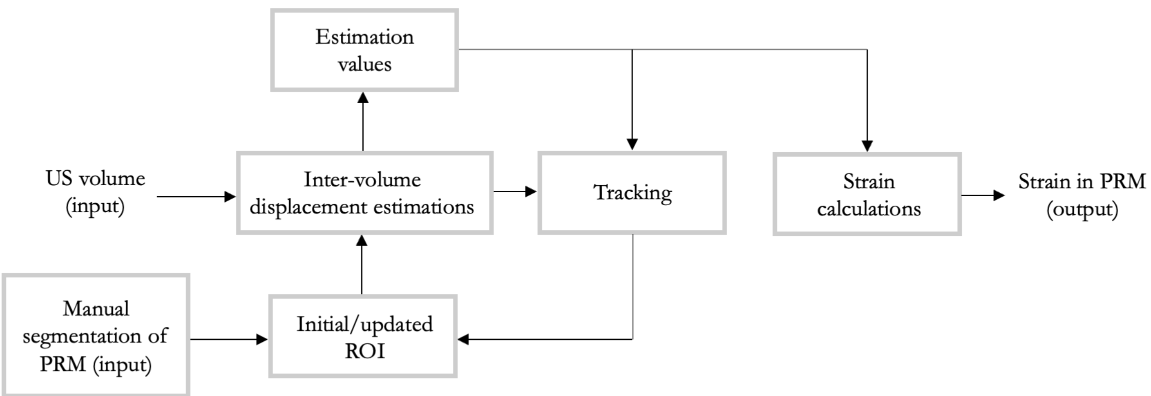


Figure 1: *Diagram of the steps to obtain the accumulated principal strain. US = ultrasound; PRM = puborectalis muscle; ROI = region of interest.*

Although both studies investigated a different patient population and had a different research aim, the articles showed relevance of calculating biometric parameters in the evaluation of muscle function. Therefore, in this study both LH area in rest and the difference in AP-diameter of the LH between rest and maximum contraction were assessed from TPUS. All parameters were compared between pre- and post-treatment. Both parameters were determined using the measuring tape tool facilitated by 3D Slicer.[26,27]

2.2.5 Correlation between parameters

When both biometric parameters and the strain measurement are effective predictors of muscle function, the additional value of muscle strain can be questioned. To explore the relationship between the biometric parameters and strain, an assessment of correlation is performed, using Pearson's correlation coefficient.

2.2.6 Statistical analysis

Despite the small dataset, quantitative analysis is performed to evaluate muscle function of both groups. The quantitative dataset consists of MPS values, LH area in rest position, and the difference in AP-diameter of the LH between rest and maximum contraction. To assess whether a statistically significant difference exists in the parameter outcomes before and after treatment, a paired t-test was applied in case of normal distribution for each group individually. When data did not exhibit a normal distribution, the Wilcoxon signed rank test was used. To determine any significant difference between the two groups, an independent t-test was conducted, or a Mann-Whitney U test when data was not normally distributed. For the statistical analysis of the data, IBM SPSS Statistics version 21 software (IBM SPSS, Armonk, NY, USA) was utilized.

2.2.7 Interobserver variability

Part of the accuracy of this study relies upon the accuracy of the PRM segmentation from the TPUS volume. The accuracy of the segmentation might vary between observers. Therefore, the interobserver variability was evaluated. For this purpose, five TPUS images were segmented by three different observers. Observer 1 had extensive experience with PRM segmentation. Their segmented images were considered as the ground truth. Observer 2 had experience in segmenting ultrasound images, although not specifically in PRM segmentation. Observer 3 was completely new to image segmentation. Both Observers 2 and 3 were trained by Observer 1 on how to perform the segmentation of TPUS images. The segmentations were performed using MeVisLab (MeVis Medical Solutions, Bremen, Germany) by Observer 1, and 3D Slicer by Observers 2 and 3.

Different metrics are used in literature to calculate the overlap of two segmentations.[28,29] The Dice Similarity Coefficient (DSC) is the most common metric used in medical image segmentation. It quantifies the degree of overlap between two segmentations by computing twice the intersection volume of the two segmentations divided by the sum of their individual volumes:

$$DSC = \frac{2(A \cap B)}{(A + B)}$$

The DSC ranges between 0 and 1, with 0 indicating no overlap and 1 indicating maximum segmentation overlap. A DSC > 0.7 is regarded as good overlap.[30]

To address potential limitations of the DSC in cases where segmentations do not perfectly overlap, for instance, when one segmentation extends only a few millimetres beyond the other, a situation may arise where the DSC is low, despite good agreement between the segmentations. Therefore, the average Hausdorff Distance (HD) and the 95% HD have been assessed. The average HD is computed by determining the shortest distance from each point in segmentation A to segmentation B. The largest distance obtained from these calculations represents the HD, while the average HD corresponds to the mean of all these distances. Nevertheless, the HD is sensitive to outliers, which can lead to elevated values for the average HD. To account for this, the 95% HD was also considered, which calculates the mean distance between 95% of the boundary between the two segmentations. The DSC, average HD, and 95% HD were obtained using the Slicer software.

2.3 Thesis writing

For the writing of this thesis, ChatGPT was occasionally used. The AI tool was employed as a translation tool, with its input being both Dutch and English versions of text segments, originally written by the researcher. The output generated by ChatGPT was not directly copied into this thesis. Instead, it was thoroughly reviewed and modified by the researcher. No content was derived from ChatGPT.

3. Results

This chapter presents the results of patient recruitment, the accumulated displacement estimates, accumulated principal strain, and visual analysis. Furthermore, the mean principal strain, biometric parameters, and the correlation between them are demonstrated. Additionally, findings of the interobserver variability assessment are presented.

It is worth noting upfront that the results of this research do not exhibit a consistent or uniform trend. Instead, a variety in displacement and deformation of the PRM is seen for both groups. This diversity makes it challenging to provide a summary of the outcomes. To address this issue, a few examples from both groups are represented in this chapter, illustrating the outcome and its variety.

3.1 Patient demographics

Out of the GYNIUS dataset, consisting of 412 women, this study included a total number of 17 women who met the inclusion criteria, see Figure 2. The baseline characteristics of the study population are shown in Table 1. In this study all women wore an open ring or ring with support. For one participant the type of pessary was unknown. In the Combined therapy group, the specific number of PFMT sessions was known for three participants and ranged between two to eight sessions. There is no significant difference in baseline characteristics between the two groups, except from the menopausal status of the participants. In the Combined therapy group, four out of seven women were postmenopausal, while in the Pessary therapy group, all women were in the postmenopausal phase.

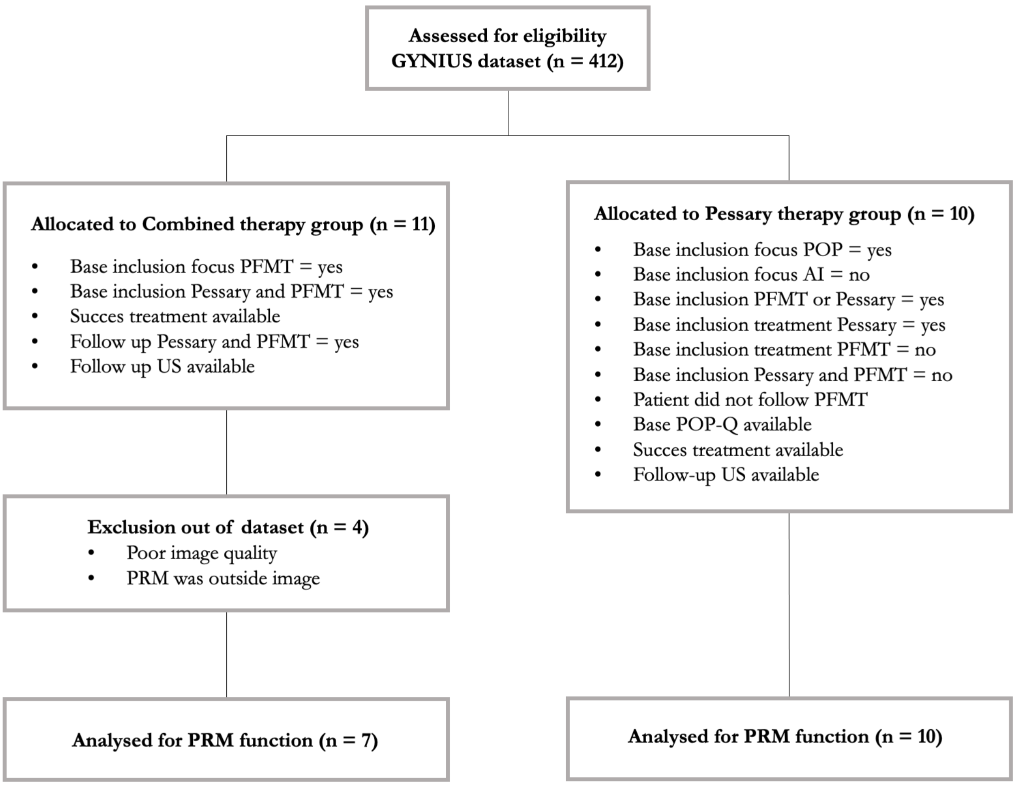


Figure 2: Flowchart of patient recruitment. AI = faecal Incontinence; PFMT = pelvic floor muscle therapy; POP(-Q) = pelvic organ prolapse (-quantification); PRM = musculus puborectalis; US = ultrasound.

Table 1: Baseline characteristics[a]

	Combined therapy (n = 7)	Pessary therapy (n = 10)	p-value
Age (SD)	58.29 (± 11.15)	63.30 (± 6.48)	0.154
Body mass index (SD), kg/m²	24.31 (± 4.31)	26.24 (± 3.62)	0.936
Parity (SD)	2.29 (± 0.76)	2.00 (± 0.667)	0.366
Postmenopausal			<0.001
Yes	4 (57.1)	10 (100)	
No	3 (42.9)	-	
Type of prolapse			0.784
Anterior vaginal wall prolapse	7 (100)	9 (90)	
Uterine prolapse	3 (42.9)	5 (50)	
Posterior vaginal wall prolapse	4 (57.1)	7 (70)	
Stage of prolapse[b]			0.674
1	0 (0)	0 (0)	
2	4 (57.1)	5 (50)	
3	3 (42.9)	5 (50)	
Avulsion			0.617
No avulsion	3 (42.9)	6 (60)	
Partial unilateral avulsion	-	1 (10)	
Partial bilateral avulsion	-	-	
Complete unilateral avulsion	2 (28.6)	1 (10)	
Complete bilateral avulsion	1 (14.3)	1 (10)	
Partial and complete avulsion	1 (14.3)	1 (10)	
Type of Pessary			0.058
Ring without support	3 (42.9)	5 (50)	
Ring with support	3 (42.9)	5 (50)	
Unknown	1 (14.3)	-	
Pessary size			0.200
Size 2	1 (14.3)	1 (10)	
Size 3	-	1 (10)	
Size 4	3 (42.9)	4 (40)	
Size 5	2 (28.6)	4 (40)	
Unknown	1 (14.3)	-	
Frequency PFMT			
2 sessions	1 (14.3)		
5 sessions	1 (14.3)		
8 sessions	1 (14.3)		
Unknown	4 (57.1)		

[a]Data is reported as number (percentage) of participants unless otherwise indicated.
[b]Stage of most prolapsed compartment.

3.2 Accumulated displacement estimates and accumulated principal strain

To be able to interpretate the results of the accumulated displacement estimates and the accumulated principal strain, an example of a healthy PRM is represented in Figure 3 and 4, respectively. The accumulated displacement estimates include images of the displacement into the z-, x-, and y-directions:

- Within the **z-direction,** a positive value indicates movement away from the os pubis, whereas a negative value indicates movement towards the os pubis;
- The **x-direction** captures lateral or side-to-side movement. In this direction minimal displacement is seen for a healthy PRM;
- Displacement along the **y-direction** indicates movement towards the ultrasound transducer (negative values) or away from the transducer (positive values). During maximum contraction, a healthy PRM moves away from the transducer.

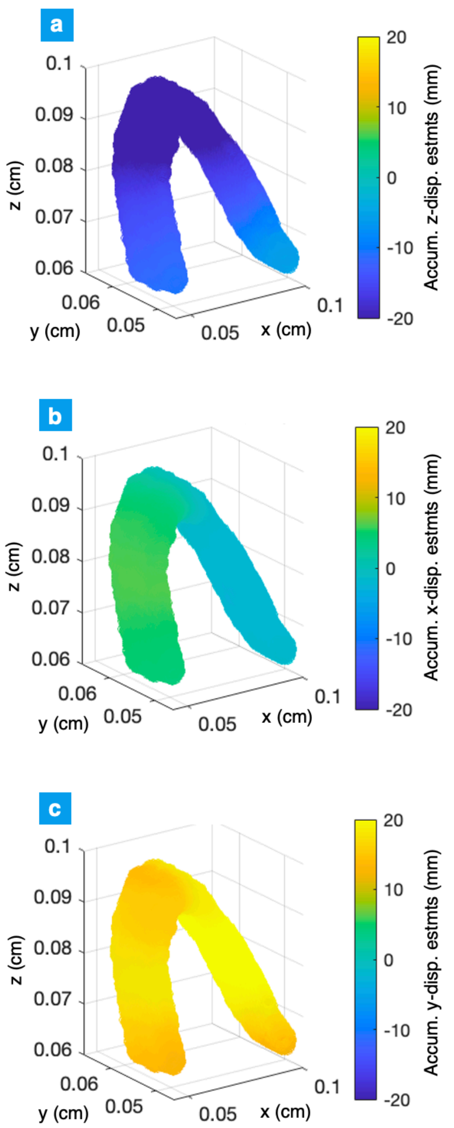


Figure 3: *Example of the accumulated displacement estimates of a healthy PRM. a) Accumulated z-direction displacement estimates. b) Accumulated x-direction displacement estimates. c) Accumulated y-direction displacement estimates*

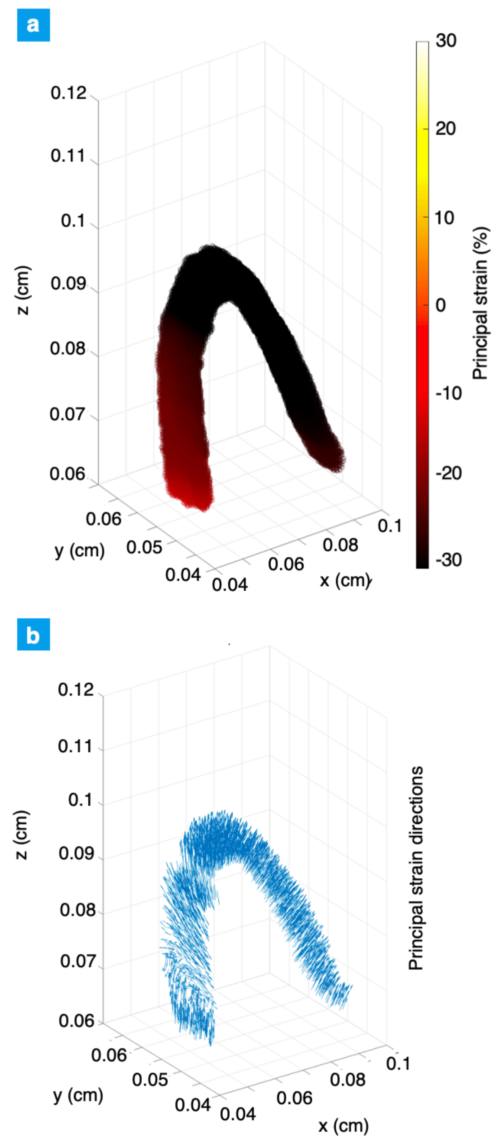


Figure 4: *Example of the accumulated principal strain of a healthy PRM. a) Accumulated principal strain image. b) Direction corresponding to the accumulated principal strain image above.*

In Figure 4, an example of the accumulated principal strain is displayed along with the corresponding directions. This figure illustrates the deformation of the muscle, with a more negative value indicating contraction and a positive value indicating elongation. For a healthy PRM, negative values for the accumulated principal strain are seen during contraction. In Figures 5 to 12, the examples of respectively two patients from each group are represented.

Patient 1 I Combined therapy

Looking at Figure 5, minimal displacement is visible along the z-axis of the first ultrasound. Analysing the follow-up image, modest displacement appears towards the os pubis in the central region of the PRM, accompanied by movement away from the os pubis on both sides of the muscle.

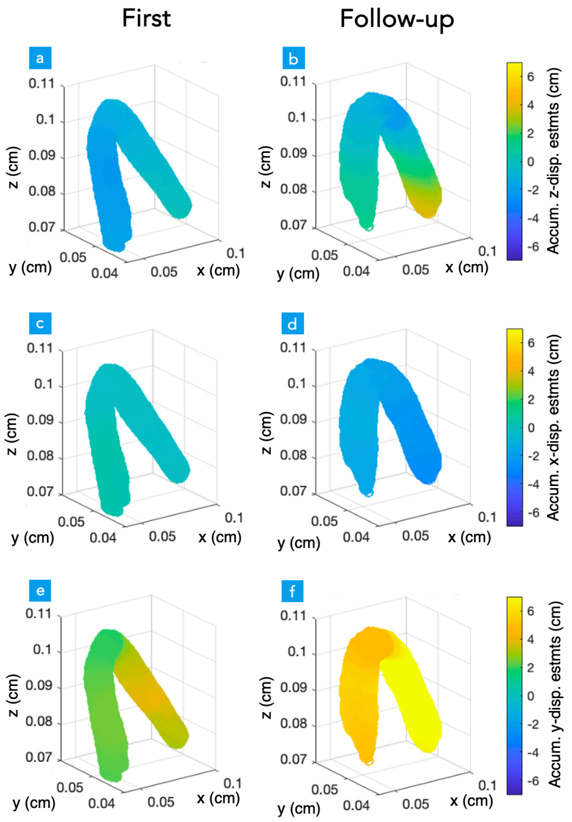


Figure 5: The accumulated displacement estimates of ***patient 1*** *from the* ***Combined therapy*** *group during the first and follow-up ultrasound. a-b) Accumulated z-direction displacement estimates. c-d) Accumulated x-direction displacement estimates. e-f) Accumulated y-direction displacement estimates*

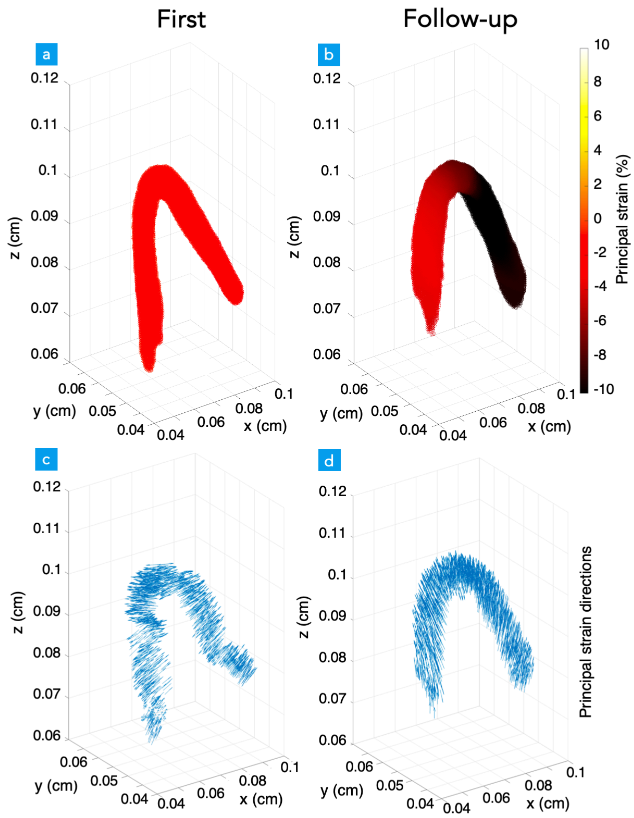


Figure 6: The accumulated principal strain of ***patient 1*** *from the* ***Combined therapy*** *group. a-b) Accumulated principal strain images of the first and follow-up ultrasound c-d) Directions corresponding to the accumulated principal strain images above.*

For the x-direction, the first ultrasound image reveals no evident displacement, while in the follow-up image, the PRM shows a small shift to the left. The most substantial difference occurs within the y-direction, where the first ultrasound image displays minor movement away from the transducer, while the follow-up image reveals a more prominent level of displacement away from the transducer.

Figure 6 displays the accumulated principal strain corresponding to the accumulated displacement estimates of the participant from Figure 5. From this figure, it appears that minimal deformation is existing in the first ultrasound image, while a significant shift in negative direction can be observed in the follow-up image, especially concentrated in the right segment of the PRM.

Patient 2 I Combined therapy

The accumulated displacement estimates of patient 2 from the Combined therapy group, displayed in Figure 7, shows minimal positive displacement in the z-direction during the first ultrasound and no displacement during the follow-up ultrasound. Similarly for the x-direction, no movement is observed in both ultrasound images. For the y-direction of the first ultrasound, positive displacement is visible in both ends of the PRM, while the centre of the muscle reveals no displacement. During follow-up the muscle shows no movement except for the right segment of the PRM.

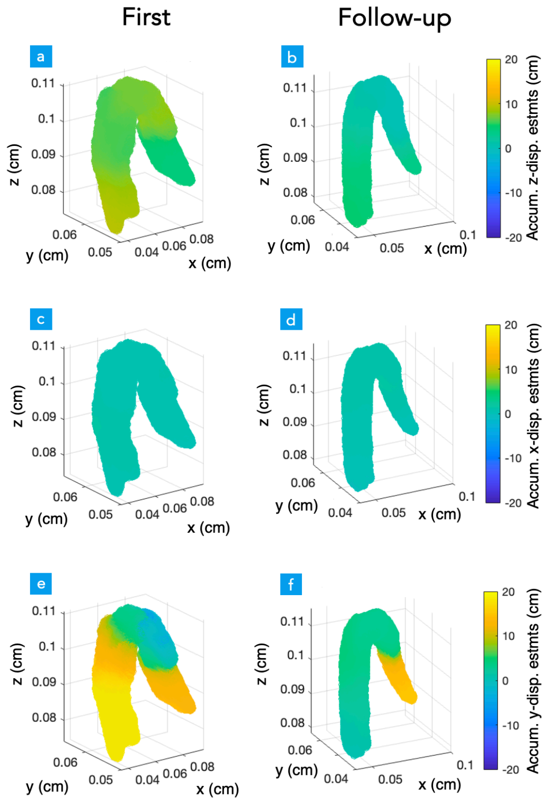


Figure 7: *The accumulated displacement estimates of* ***patient 2*** *from the* ***Combined therapy*** *group during the first and follow-up ultrasound with varying displacement. a-b) Accumulated z-direction displacement estimates. c-d) Accumulated x-direction displacement estimates. e-f) Accumulated y-direction displacement estimates*

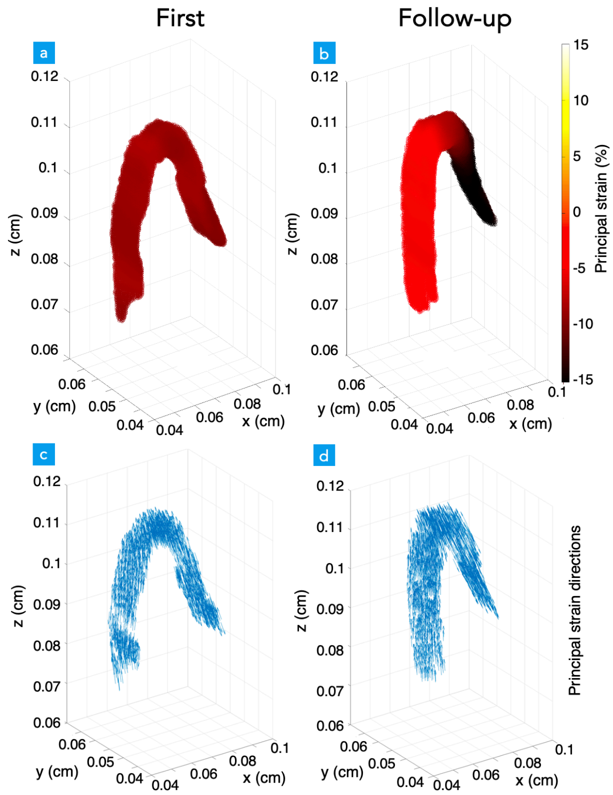


Figure 8: *The accumulated principal strain of* ***patient 2*** *from the* ***Combined therapy*** *group. a-b) Accumulated principal strain images of the first and follow-up ultrasound c-d) Directions corresponding to the accumulated principal strain images above.*

Looking at the accumulated principal strain of patient 2 in Figure 8, moderate strain is visible in the entire PRM during the first ultrasound. During the follow-up ultrasound no strain is visible in the PRM, except for the right segment of the muscle.

Patient 3 I Pessary therapy

Looking at the results from a patient of the Pessary therapy group, Figure 9, a displacement in the z-direction towards the os pubis can be observed for both first and follow-up ultrasound. In the x-direction there is very little movement in both images. For the y-direction, a difference can be observed between the first and the follow-up images. Before treatment, the PRM shows no displacement towards or away from the transducer, while after treatment, a clear movement away from the ultrasound transducer can be observed.

Looking at the accumulated principal strain for this patient in Figure 10, minimal deformation is observed in the first ultrasound, while contraction can be observed during follow-up.

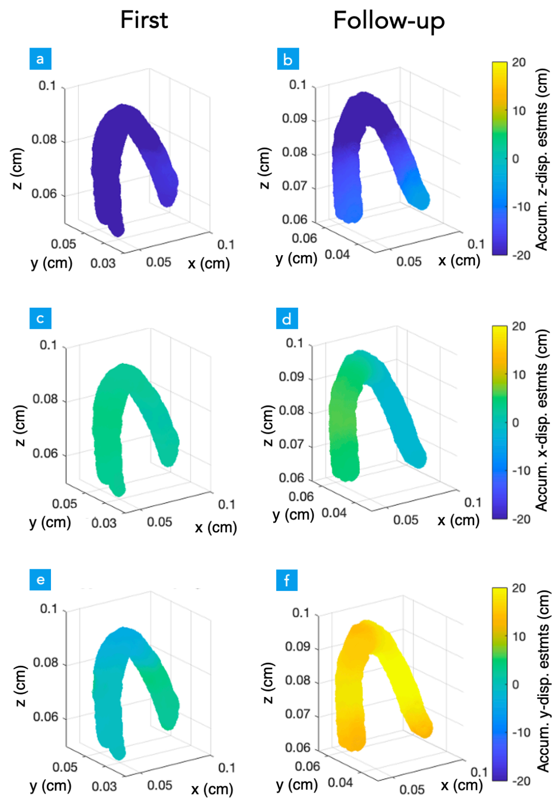


Figure 9: The accumulated displacement estimates of ***patient 3*** *from the* ***Pessary therapy*** *group during the first and follow-up ultrasound. a-b) Accumulated z-direction displacement estimates. c-d) Accumulated x-direction displacement estimates. e-f) Accumulated y-direction displacement estimates*

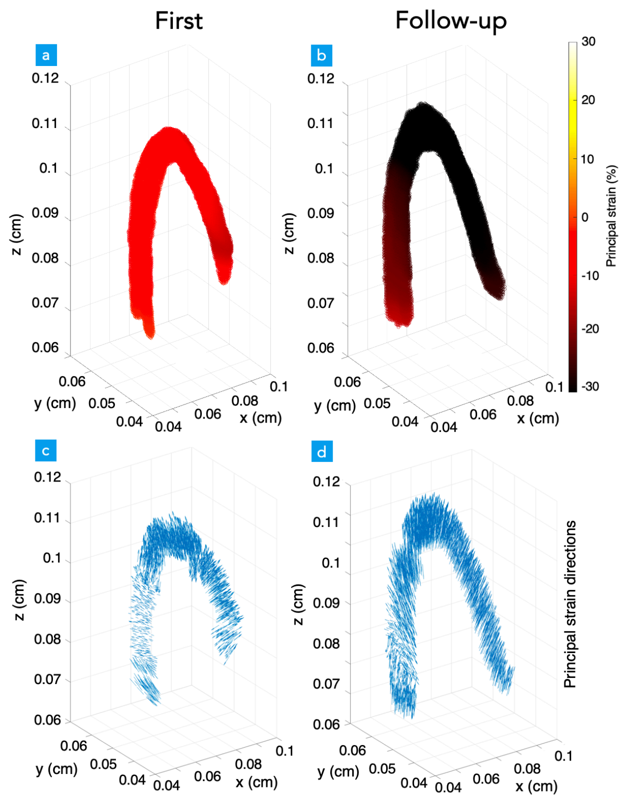


Figure 10: The accumulated principal strain of ***patient 3*** *from the* ***Pessary therapy*** *group. a-b) Accumulated principal strain images of the first and follow-up ultrasound c-d) Directions corresponding to the accumulated principal strain images above.*

Patient 4 I Pessary therapy

Figure 11 displays the accumulated displacement estimates of another patient from the Pessary therapy group. In both first and follow-up ultrasound image, the PRM moves towards the os pubis in the z-direction, with most displacement occurring in the centre of the muscle. In the x-direction, limited movement is observable in both the first and follow-up image. In the y-direction during the first ultrasound, multiple displacements occur, both away from the transducer (in the left and right segments) and towards the transducer (at the centre of the muscle). During the follow-up ultrasound displacement away from the transducer is observed.

For the accumulated principal strain of the first ultrasound, displayed in Figure 12, there is a lack of deformation in the left segment, while the right segment shows substantial contraction. For the follow-up ultrasound, contraction of the entire PRM is visible.

For visualization of the diversity in the complete dataset, accumulated principal strain images of all participants are added to Appendix I and II.

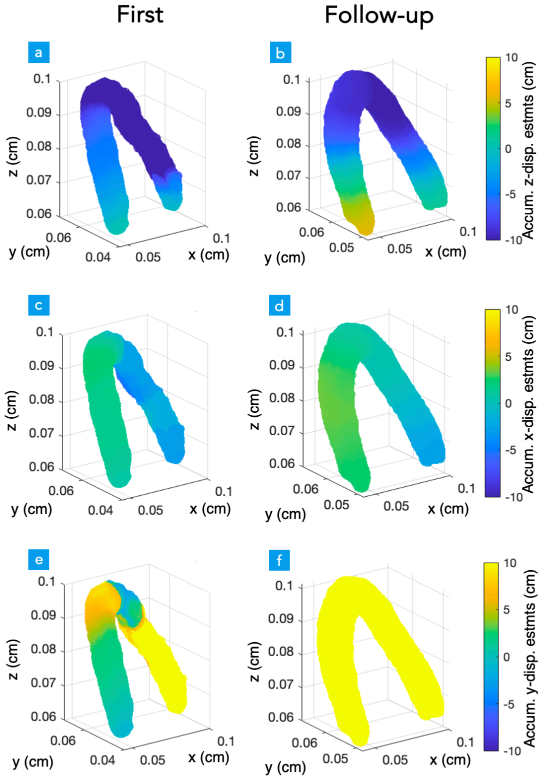


*Figure 11: The accumulated displacement estimates of **patient 4** from the **Pessary therapy** group during the first and follow-up ultrasound with varying displacement. a-b) Accumulated z-direction displacement estimates. c-d) Accumulated x-direction displacement estimates. e-f) Accumulated y-direction displacement estimates*

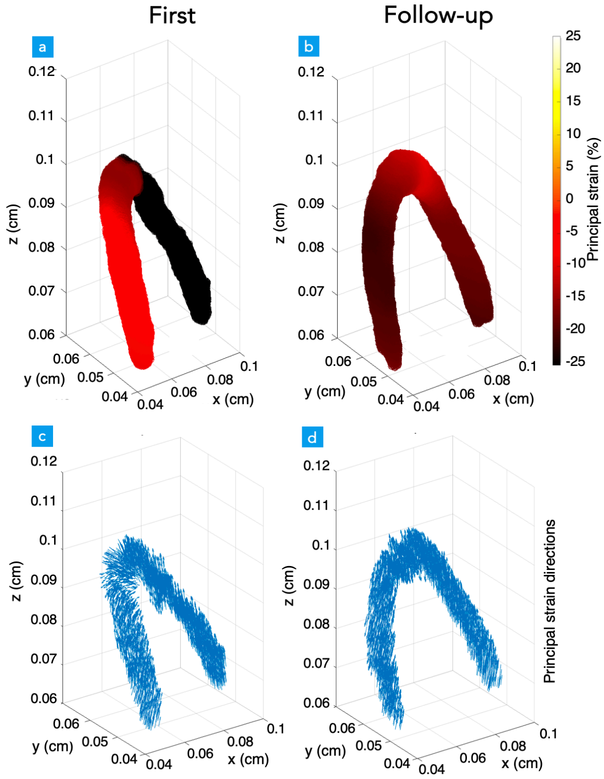


*Figure 12: The accumulated principal strain of **patient 4** from the **Pessary therapy** group. a-b) Accumulated principal strain images of the first and follow-up ultrasound c-d) Directions corresponding to the accumulated principal strain images above.*

3.3 Visual analysis

While evaluating the ultrasound recordings and segmentation tracking videos, a diverse range in the degree of contraction was observed, along with variations in the accuracy of segmentation tracking. The outcomes for all participants during both the initial and follow-up ultrasound sessions have been added to Appendix III.

3.4 Mean principal strain

The average MPS value for both groups is presented in Table 2. For the Combined therapy group, the MPS value increases after therapy. For the Pessary therapy group, the MPS value decreases, while showing a broad variation as can be observed from the standard deviation. However, the difference between the MPS value before and after treatment is not significant for both groups individually and compared.

Table 2: The absolute difference between the MPS value in rest and during maximum contraction. Statistical p-values are computed to compare first and follow-up ultrasound and to compare both groups. SD = standard deviation.

	Combined therapy (n = 7) Mean (SD)			Pessary therapy (n = 10) Mean (SD)			Difference between groups
	First	**FU**	**p-value**	**First**	**FU**	**p-value**	**p-value**
Difference MPS value between rest and max. contraction	-2.5 (± 1.8)	-1.4 (± 2.4)	0.263	-5.1 (± 10.44)	-7.9 (± 9.3)	0.241	0.133

3.5 Biometric parameters

The biometric parameters that were assessed are the hiatal area in rest and the difference in AP-diameter of the LH between resting state and maximum contraction. Given the limited dataset, it could not be definitively concluded whether the data was normally distributed. As these parameters exhibit a normal distribution in the overall population, it has been assumed that the data in this study also adheres to a normal distribution.[31] The outcomes of the biometric parameters are displayed within Table 3.

Hiatal area

The Combined therapy group shows a significant increase of the hiatal area after treatment, with small standard deviation. In the Pessary therapy group, a decrease is observed for the hiatal area, although not statistically significant. An independent t-test shows no significant difference in hiatal area between both groups.

AP-diameter

For the difference in AP-diameter of the LH between rest and maximum contraction, both groups show an increase after therapy. However, for both groups the increase is not significant. Moreover, the difference in the AP-diameter between the two groups is not statistically significant either.

Table 3: *Measurements of the hiatal area in rest and the difference in AP-diameter of the LH between rest and maximum contraction. Statistical p-values are computed to compare first and follow-up ultrasound and to compare both groups. SD = standard deviation.*

	Combined therapy (n = 7) Mean (SD)			**Pessary therapy (n = 10)** Mean (SD)			**Difference between groups**
	First	**FU**	**p-value**	**First**	**FU**	**p-value**	**p-value**
Hiatal area (mm^2) in rest	167.4 (± 3.8)	185.6 (± 4.6)	0.036	201.5 (± 9.7)	189.4 (± 3.7)	0.729	0.475
Difference in AP-diameter (mm) between rest and max. contraction	3.0 (± 3.1)	5.1 (± 2.9)	0.110	8.8 (± 3.6)	9.1 (± 2.9)	0.610	0.165

3.6 Correlation between parameters

The outcome of the correlation test between the biometric parameters and the MPS value is presented in Table 4. No significant correlations exist between the parameters.

Table 4: *Pearson's Correlation among biometric and strain parameters*

		1	**2**	**3**
1.	**Hiatal area (mm^2)**	-		
2.	**Difference in AP-diameter (mm) between rest and max. contraction**	r =-.037 p =.886	-	
3.	**Difference MPS value between rest and max. contraction**	r = 0.099 p = .707	r =-.129 p = .621	-

3.7 Interobserver variability

The interobserver variability is based upon six segmentations. The mean values of the DSC, average HD, and 95% HD are presented in Table 5. The DSC reveals moderate overlap between different observers, ranging between 0.43 and 0.51. The average HD varies between 2.09 and 2.41 mm with a small standard deviation. The 95% HD shows less overlap with a value ranging between 6.59 and 8.81 mm and a larger standard deviation.

Table 5: *Mean values of comparison metrics between the three observers.*

	DSC	Average HD (mm)	95% HD (mm)
Observer 1 – 2	0.48 (± 0.08)	2.21 (± 1.72)	7.66 (± 7.80)
Observer 1 – 3	0.43 (± 0.06)	2.52 (± 1.15)	9.23 (± 5.80)
Observer 2 – 3	0.54 (± 0.08)	1.88 (± 0.99)	6.62 (± 5.38)

To assess the influence of various observers on the strain, their MPS values are presented in Table 6. It can be observed that the difference between the MPS values for the three observers differ and do not show a clear correlation with the amount of overlap.

Table 6: *The amount of overlap between the observers and the corresponding MPS values during maximum contraction for the observers individually. The mean difference is the average difference between the MPS values of the different observers.*

		MPS value during max. contraction (%)			
Patient	**Overlap**	**Observer 1**	**Observer 2**	**Observer 3**	**Mean difference**
1	Good	-6.84	-8.79	-8.33	1.31
2	Good	-37.77	-44.65	-34.76	6.59
3	Good	-15.31	-16.64	-22.36	4.70
4	Moderate	-13.87	-13.53	-16.48	1.97
5	Moderate	-30.82	-22.0	-18.18	8.43
6	Poor	-12.24	-11.38	-11.05	0.79

In Figure 13, the 3D examples of one segmentation with sufficient overlap and another segmentation with lack of consensus among the three observers are represented. For the segmentation with sufficient overlap, the accumulated displacement estimates are displayed in Figure 14.

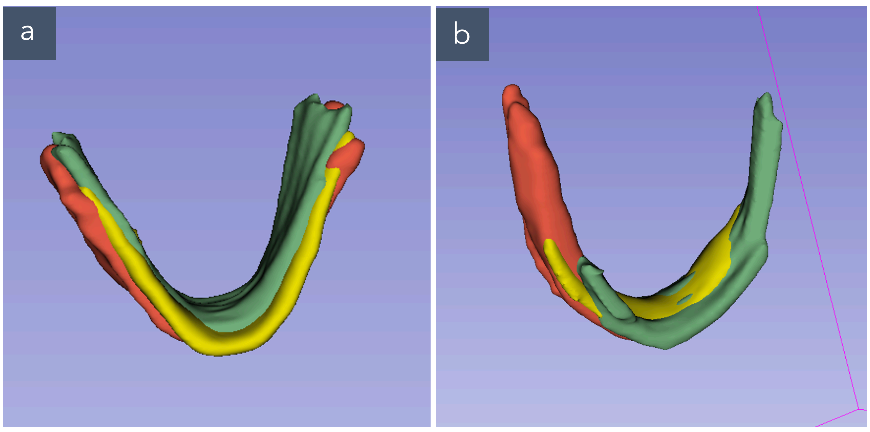


Figure 13*: Examples of the PRM segmented by Observer 1 (green), Observer 2 (red) and Observer 3 (yellow). a) Segmentation with sufficient overlap. b) Segmentation with lack of consensus among the three observers.*

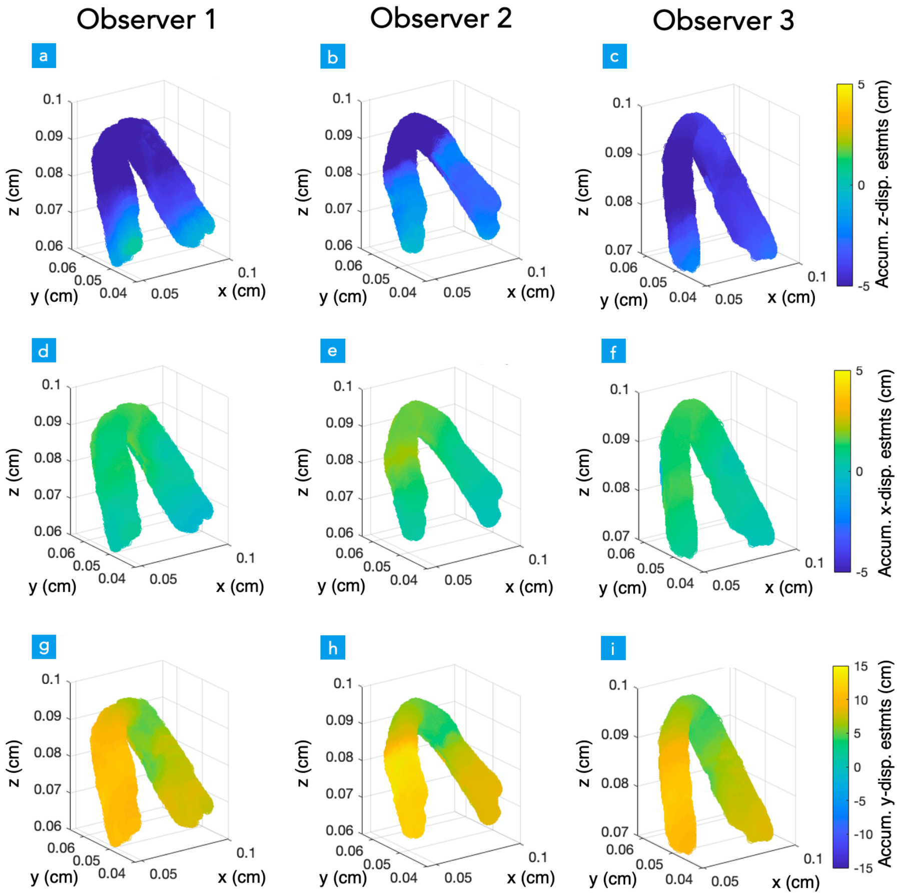
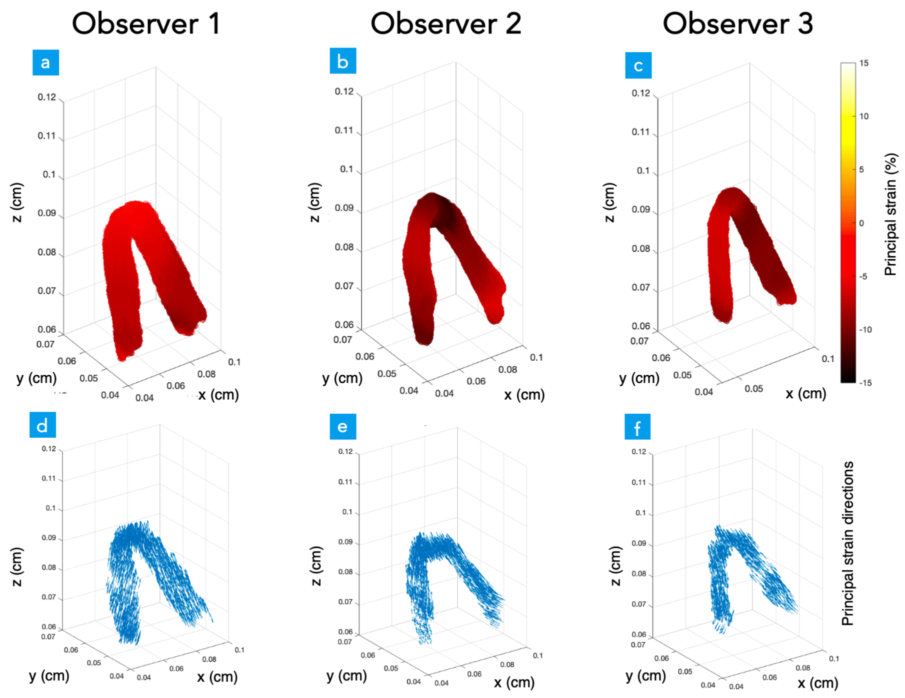


Figure 14: *Segmentation with sufficient overlap between different observers. a-c) Accumulated z-direction displacement estimates. d-f) Accumulated x-direction displacement estimates. g-i) Accumulated y-direction displacement estimates.*



Average MPS-value	-6,84%	-8,79%	-8,33%

Figure 15: *Segmentation with sufficient overlap between different observers. a-c) The accumulated principal strain magnitude of the three different observers. d-f) Corresponding directions of the accumulated principal strain images above. Below: average MPS-value of the three observers.*

The accumulated displacements along the x-, y-, and z-directions exhibit a consistent pattern among the three observers, despite some variation in the extent of displacement. The same applies to the corresponding accumulated principal strain, as depicted in Figure 15, although a smaller deformation of the PRM is apparent for Observer 1. This is confirmed by the MPS values, showing a more positive value for Observer 1 (-6.84%). MPS values of Observers 2 (-8.79%) and 3 (-8.33%) are almost similar.

The accumulated displacement estimates of a segmentation with no consensus between the three observers is displayed in Figure 16. What stands out is the shape of the segmentation and consequently, the displacement. Evaluating the z-direction, all three segmentations show displacement towards the os pubis in the middle segment of the muscle. Notably, where Observer 1 segmented a right muscle segment, there is visible movement away from the os pubis. A similar trend is observable for the left muscle segment, segmented by Observers 2 and 3. Figure 16 d-f reveal a consistent pattern for all three observers, with a slight positive displacement along the x-direction. However, a higher amount of displacement appears in the left segment, segmented by Observers 2 and 3, compared to the right segment, segmented by Observer 1. In the y-direction, predominantly positive displacement is observed, although Observer 1 also indicates a minor negative displacement in the right segment.

For the corresponding accumulated principal strain, displayed in Figure 17, the deformation is nearly similar for the segmentation of Observers 2 and 3, despite small difference in segmentation. The segmentation of Observer 1 shows significantly more deformation, especially in the right segment. However, the MPS value indicates nearly comparable measurements for Observers 2 and 3, with a subtle variation observed for Observer 1.

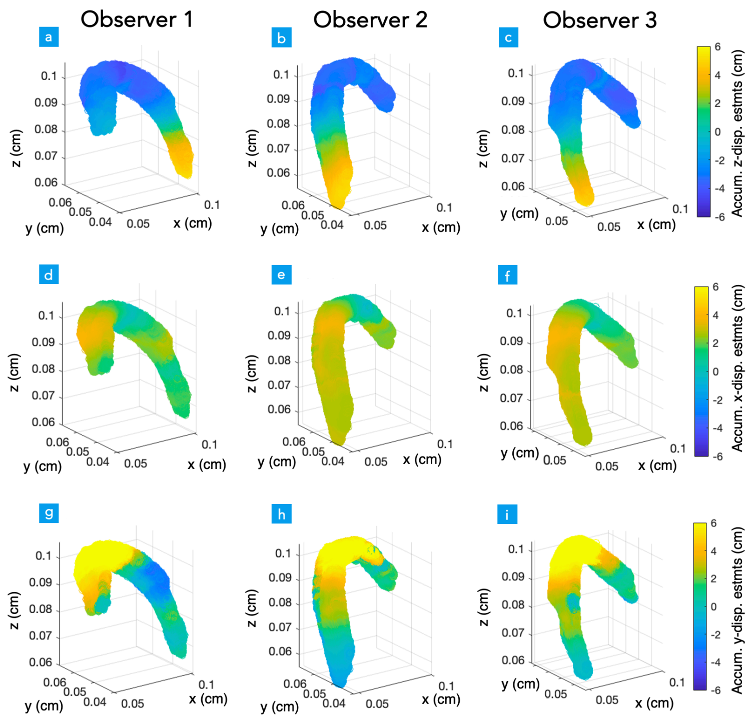
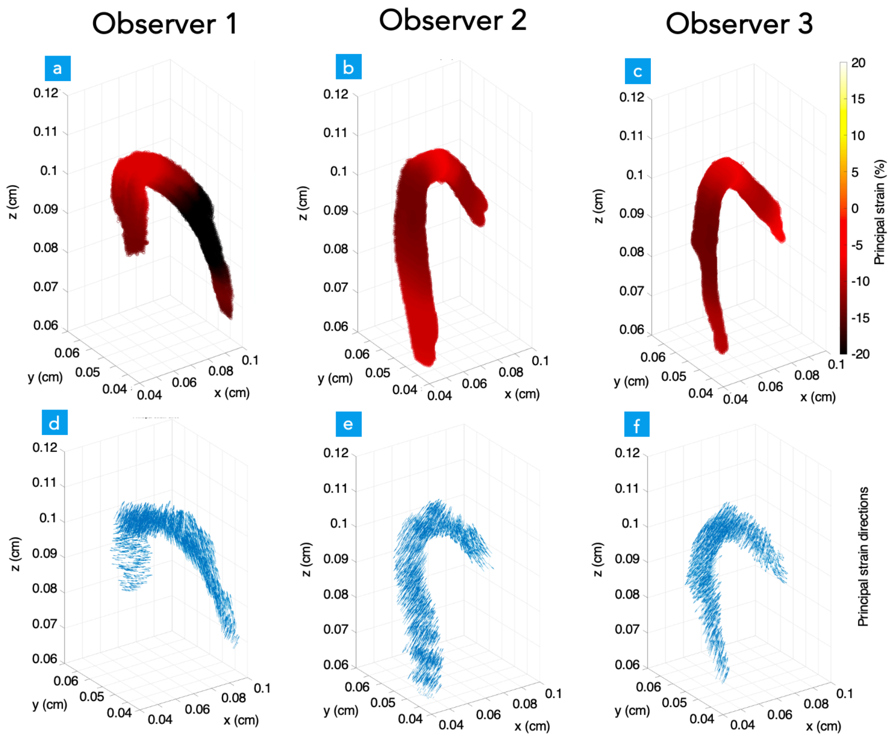


Figure 16: Segmentation with lack of consensus among the observers. a-c) Accumulated z-direction displacement estimates. d-f) Accumulated x-direction displacement estimates. g-i) Accumulated y-direction displacement estimates.



Average MPS-value	-12,24%	-11,38%	-11,05%

Figure 17: Segmentation with lack of consensus among the observers. a-c) The accumulated principal strain magnitude of the three different observers. d-f) Corresponding directions of the accumulated principal strain images above. Below: average MPS-values for the three observers.

4. Discussion

This study aimed to investigate the additional value of PFMT to pessary wearing in comparison to pessary wearing alone. The research involved the comparison of mean principal strain and biometric parameters before and after treatment by means of 4D TPUS. Furthermore, the study explored the influence of various observers and their segmentations on the strain measurements. This research is a first step towards predicting which women will benefit from PFMT and whom will not. It contributed to the understanding of PRM strain and exposed several software challenges that need to be considered in the future. Additionally, this thesis can be used as the foundation of a larger prospective RCT to gain further insight into this topic.

The results of this study provide the following answers to the sub-questions:

1. There is no significant difference in MPS-value between the first and follow-up ultrasound for both groups, individually and compared.
2. Hiatal area: a significant increase in hiatal area in rest was observed within the Combined therapy group after treatment, although a decrease was expected. The Pessary therapy group did reveal a decrease in hiatal area, although not significant. AP-diameter: no significant difference was observed in AP-diameter between the first and follow-up ultrasound for both groups, individually and compared.
3. There exists no correlation among the three parameters.
4. Interobserver variability revealed a moderate overlap among segmentations performed by three observers according to comparison metrics, contrasted by a similarity in MPS-values.

Key observations of this study encompass highly variating results which deviate from the initial expectations, and it cannot be presumed that quantitative outcomes are accurate. Based on these findings, a definitive conclusion on the main research question: "*What is the effect of combined therapy for the treatment of POP, compared to pessary therapy, on the strain of the PRM measured with TPUS?"*, cannot be drawn. Therefore, this chapter gives a detailed discussion of the findings, focusing on possible explanations for the deviating results and offering possible resolutions or recommendations, instead of focusing on the differences between the two groups.

4.1 Accumulated displacement estimates, accumulated principal strain & visual analysis

According to the research conducted by Das et al., a healthy PRM exhibits a negative displacement in the z-direction during contraction, accompanied by minimal to negligible displacement in the x-direction, and a positive displacement in the y-direction.[9] This indicates that the muscle moves towards the os pubis while moving away from the transducer. In simpler terms, the muscle becomes shorter and thicker.

Figures 5 to 12 give the examples of two cases extracted from each group: one with displacement and deformation that seem to align with the article of Das et al., and the other with deviating results. To explain these results, for all participants, the original ultrasound video was evaluated to compare the contraction as seen in the recording with the accumulated displacement estimates and accumulated principal strain images. This section discusses the results of the four patients and gives possible explanations and limitations for the deviating results.

Patient 1 | Combined therapy

The original ultrasound video of patient 1 exhibits minimal contraction of the pelvic floor muscles during the first ultrasound, while a slightly improved contraction is observed during follow-up. This aligns with the pattern observed in the accumulated displacement estimates and accumulated principal strain, displayed in Figure 5 and 6, and the hypothesis that the strain improves after treatment. Nevertheless, this case remains somewhat exceptional rather than representative for the rest of the dataset.

Patient 2 | Combined therapy

The second patient from the Combined therapy group gives an example of an unexpected outcome. In the accumulated displacement estimates of the first ultrasound, Figure 7, a small positive displacement in the z-direction is observed. Instead of contraction, this suggests elongation of the muscle. In the y-direction, there is displacement of the muscle in the expected direction, but the center of the muscle seems to stay in place. Looking at the initial ultrasound video for this patient, an unusual movement is observed, resembling a backward pull of the muscle rather than movement towards the os pubis. This observation could account for the displacement in the z-direction. However, when examining the accumulated principal strain, as shown in Figure 8, an overall muscle contraction appears to be present.

For the follow-up ultrasound, minimal displacement is observed in all three directions, except in the y-direction. In this direction, only the right segment of the muscle exhibits positive displacement. This could potentially be attributed to the complete unilateral avulsion of the PRM in this patient. However, the line between orange and green (moderate vs no displacement) in the right segment is so pronounced that it suggests the muscle segmentation may be inaccurate. From the accumulated principal strain image, Figure 8, a moderate contraction is seen in the right segment of the muscle. The rest of the muscle remains inactive, although the original ultrasound video shows a moderate contraction of the muscle. Consequently, it may be implied that the muscle is not accurately segmented.

Patient 3 | Pessary therapy

Also, for the Pessary therapy group only few patients contracted their pelvic floor muscles according to expectations. For patient 3 from the Pessary therapy group, as depicted in Figures 9 and 10, a similar pattern is observed as for patient 1 from the Combined therapy group. Even a more pronounced displacement along the y-direction during the follow-up ultrasound is reflected in the accumulated displacement estimates, resulting in a substantial deformation in the accumulated principal strain. This is confirmed by the ultrasound videos of this participant, which reveal a sufficient contraction during the first assessment, and a better contraction during the follow-up assessment.

Patient 4 I Pessary therapy

Patient 4, displayed in Figures 11 and 12, gives an example of the Pessary therapy group in which the displacement and strain pattern appear to be different from the expected outcome. Although the first ultrasound video shows a good contraction, the first accumulated displacement estimates reveal a deviating pattern. Notably, in the y-direction, the displacement alternates towards and away from the transducer. Additionally, minimal deformation is observable in the left muscle segment, looking at the accumulated principal strain in Figure 12. Explanations for these and other abnormalities will be hypothesized:

Limitation 1 I Segmentation

In the example of patient 4, the issue might derive from an inaccurate segmentation of the muscle. When examining the accumulated principal strain in Figure 12, there seems to be no deformation in the left segment of the muscle in the first ultrasound. However, observing the strain from the follow-up ultrasound, deformation is noticeable in the right part of the muscle. Hence, there are two possible explanations: either the patient learned to contract the entire muscle instead of just one side between the first and follow-up ultrasound, or there is an inadequate segmentation on the left side of the muscle within the first ultrasound. Considering that the ultrasound video of this participant demonstrates a strong contraction during the first ultrasound, an incorrect segmentation seems to be a more reasonable explanation for the observed issue.

To investigate whether this is the issue, the PRM in the follow-up ultrasound was re-segmented for this patient, with a slight upward adjustment made to the left segment of the muscle. The results are depicted in Figures 18 and 19.

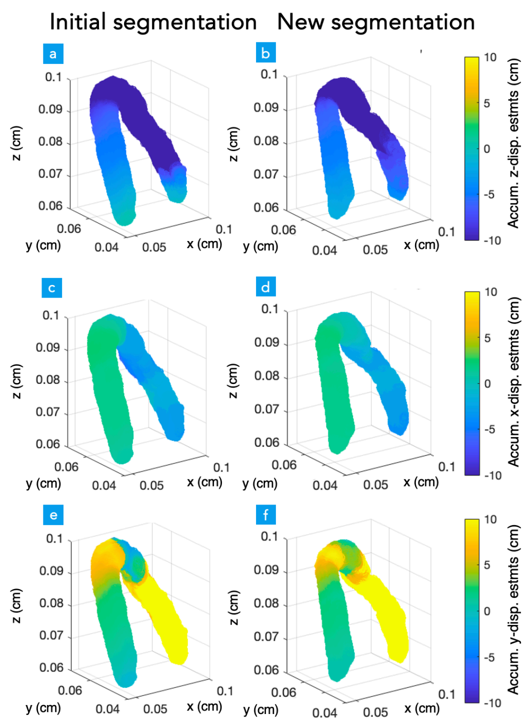


Figure 18: The accumulated displacement estimates from an initial PRM segmentation and a new segmentation. a-b) Accumulated z-direction displacement estimates. c-d) Accumulated x-direction displacement estimates. e-f) Accumulated y-direction displacement estimates

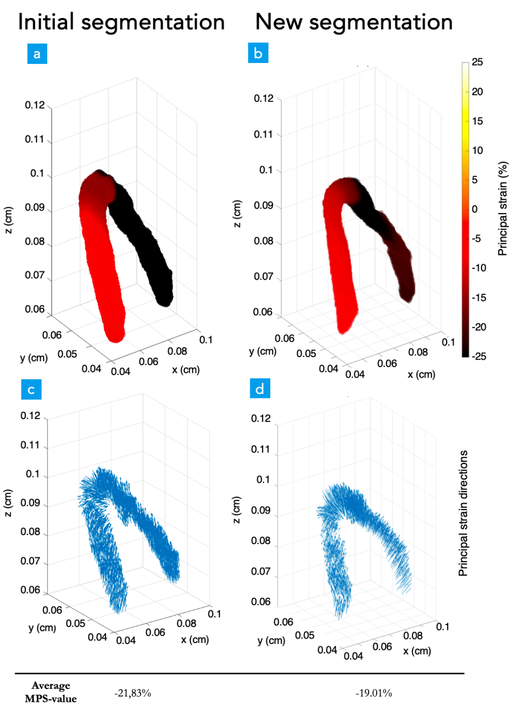


Figure 19: The accumulated principal strain from an initial PRM segmentation and a new segmentation. a-b) Accumulated principal strain images of the initial and new segmentation c-d) Directions corresponding to the accumulated principal strain images above.

Two notable observations can be made from these images. 1) The accumulated displacement estimates, and the accumulated principal strain appear to be nearly identical for both segmentations. 2) The new segmentation exhibits irregular bumps instead of a smooth boundary of the muscle. Since the segmentation was re-segmented for only one patient, it cannot be concluded from these figures that the deviating results stem from incorrect muscle segmentation. However, the unusual shape in the re-segmented muscle highlights a different issue: incorrect muscle tracking.

Limitation 2 | Tracking

One of the issues encountered in the process of the strain measurement seems to be the incorrect tracking of the segmentation mask from rest to contraction. An example is depicted in Figure 20, which shows that the software incorrectly estimates the location of the muscle to be in the rectum.

A possible hypothesis for this inaccurate tracking is an incorrect ROI update: When the pelvic floor muscles contract, the muscles undergo displacement. Accurate tracking of this displacement is crucial for strain calculation. Given that 4D ultrasound consists of 3D frames acquired over time, the position of the PRM varies across the frames due to the contraction. This is computed during intervolumetric displacement estimations using a ROI. The ROI must move along with the PRM to prevent it from moving out of the field of view. Therefore, the ROI is updated after each displacement estimation. However, the ROI is dependent of kernel and template sizes. Too small kernel or template sizes can result in the kernel and template not being matched, and to the absence of cross-correlation peaks. The developers of the strain software are currently investigating whether this may explain the incorrect tracking.

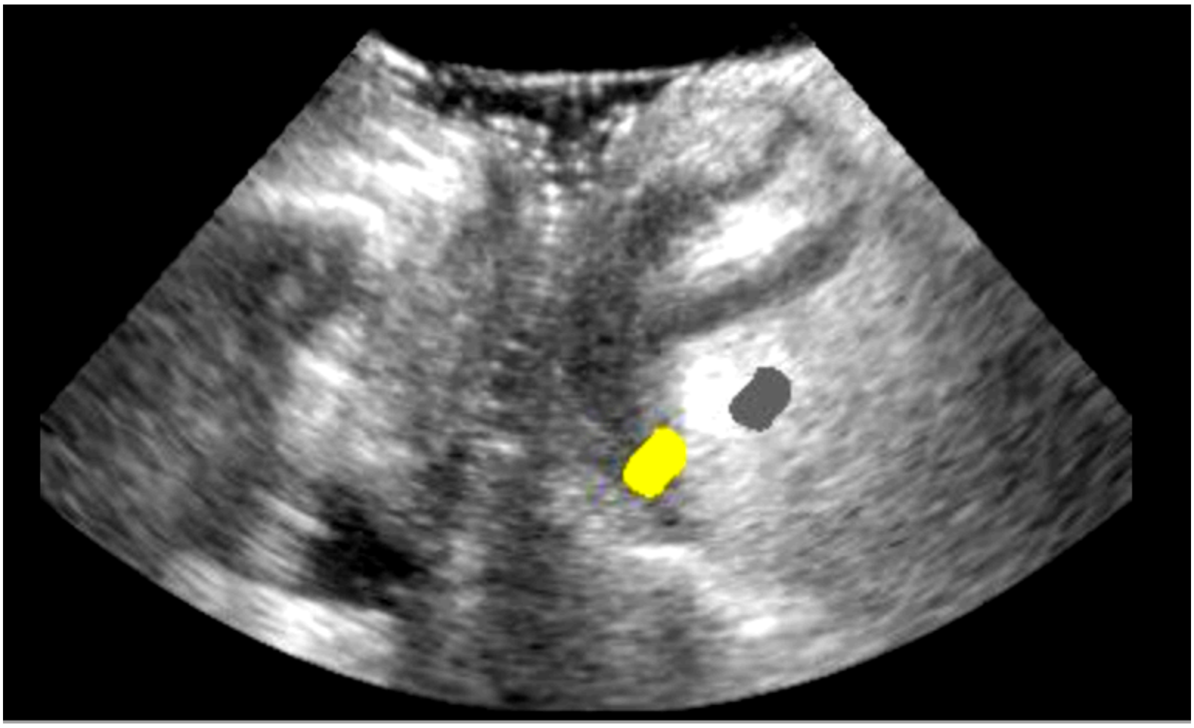


Figure 20: Sagittal ultrasound slice during maximum contraction of the pelvic floor muscles. Yellow area: position of the muscle during maximum contraction as estimated by the strain software. Dark grey area: position of the muscle in rest position, which can be ignored here.

Limitation 3 I Pessary

Another explanation for the incorrect tracking could be the presence of the pessary in the ultrasound image. For practical reasons, the method of this study involved using ultrasound images with the pessary *in situ* for all participants. However, the determination of strain in women wearing a pessary has not been explored before, thus the influence of a pessary on the image quality, muscle tracking and strain calculation is still unknown. An initial attempt to explore this is presented in Figure 14, where the accumulated displacement estimates and accumulated principal strain for a single participant are calculated both with and without the pessary i*n situ*.

As can be observed, minor differences are noticeable between the accumulated displacement estimates and accumulated principal strain with and without pessary. Displacement in the z-direction remains relatively consistent, with a slightly greater negative displacement in the right muscle segment for the muscle without the pessary. This suggests that the patient can slightly better contract the muscle without the pessary *in situ*. Displacement in the y-direction shows nearly identical patterns. Notably, the most significant difference is observed in the x-direction, wherein minimal displacement appears with the pessary *in situ*, while without the pessary *in situ*, the left muscle segment moves in a positive direction.

Commonly, displacement and deformation occur primary along the z- and y-axes during contraction, therefore movement along the x-axis is not expected. The discrepancy along the x-axis could be assigned to the sequence of ultrasound acquisition. The ultrasound with the pessary in situ was performed initially, followed by removal of the pessary and subsequent

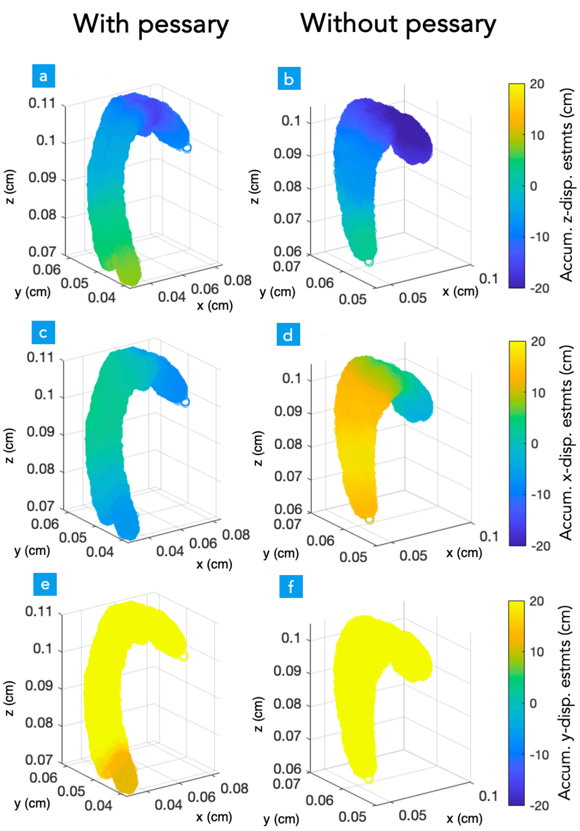


Figure 21: Accumulated displacement estimates from one patient. a, c, e) With pessary in situ. b, d, f) Without pessary in situ.

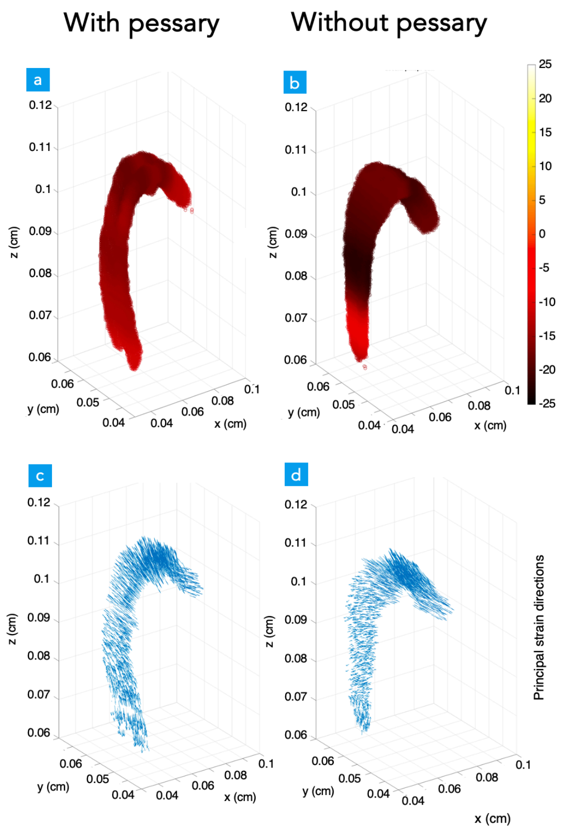


Figure 22: Accumulated principal strain from one patient a, c, e) With pessary in situ. b, d, f) Without pessary in situ.

ultrasound imaging without pessary in situ. The muscles might be more tense after pessary removal or need to slowly adapt to the new situation without the pessary, which could explain movement along the x-axis.

Based on this single attempt to evaluate the effect of a pessary versus no pessary present in the ultrasound, it is not possible to draw a conclusion. It is recommended to utilize ultrasound images without a pessary in situ until further research is conducted to investigate the influence of the pessary on strain calculations.

Aside from the possibility that the pessary may disrupt muscle tracking, there are other effects of the pessary for which there is currently limited understanding. The pessaries in this study were inserted prior to the first ultrasound scan. However, recent (unpublished) research at the University of Twente has discovered that the location of the pessary significantly differs immediately after insertion compared to a week later. Shortly after insertion, the pessary is positioned lower in the pelvic region and slightly displaces the pelvic floor muscles, while a week later, the pessary is located at a higher position in the vagina. The extent to which this affects PRM contraction has not been investigated yet, but this could have an impact on the results in this study. Therefore, further research is required.

Limitation 4 I First and follow-up TPUS

As previously discussed, first and follow-up ultrasound of all participants were assessed visually to evaluate the quality of contraction during these recordings. In the Pessary therapy group, a substantial part of the group exhibited sufficient contractions in both the first and follow-up ultrasounds.

Conversely, a remarkable observation can be seen in the Combined therapy group, where four out of seven women demonstrated minimal contractions in both the first and follow-up ultrasounds. Considering the expected improvement in contraction after PFMT and the increasing MPS values, these observations are deviating.

A clear explanation for these disparities is lacking. No significant difference exists in terms of POP-Q stage, comorbidities, ring type or size, except for menopausal status. Within the Pessary therapy group, all participants are post-menopausal, while in the Combined therapy group, only four out of seven participants have reached the post-menopausal phase. The menopause is associated with a decrease in oestrogen, which leads to a reduction in collagen and therefore a loss of muscle volume.[32,33] This may impact the effect of PFMT, despite the observation that both premenopausal and post-menopausal women showed suboptimal contractions within this dataset.

However, in future research, it is important to include universally either pre-menopausal or post-menopausal women. Alternatively, it would be valuable to initiate a study where menopausal status is explicitly included as a parameter of investigation.

Limitation 5 I Retrospective dataset

In this study, a retrospective dataset is utilized. The dataset from which the participants were obtained encompasses women who underwent 4D TPUS during their first consultation at the

Gynaecology department. These women, aging between 19 and 83 years old, presented a diverse range of symptoms. From this dataset, it was attempted to obtain a uniform group of women, with the only difference being PFMT or not. As a result, the group of participants that met the inclusion and exclusion criteria was rather small, including only 17 eligible women.

Because of the retrospective nature of the dataset, information regarding the number of sessions followed by women in the Combined therapy group is not available for all participants. Additionally, there was a lack of information regarding the therapy adherence. Both are important aspects of the effectivity of PFMT and should be considered in future research. Therefore, a prospective RCT is recommended.

Limitation 6 I Software

Another limitation pertains to the strain analysis software. The utilization of the software represents a quantitative approach to understand the behaviour and function of the PRM. However, it should be noted that the software needs to be optimized, as it currently takes 12 hours to run one segmentation. This time-consuming process makes it impossible to use the software on a large-scale basis or in clinical setting.

4.2 Mean Principal strain

Besides deviating results from the accumulated displacement estimates and the accumulated principal strain, a variability in the MPS value is apparent, especially within the Pessary therapy group showing a large standard deviation. Furthermore, there is an observed increase in strain within the Combined therapy group, contrary to the initial hypothesis.

The hypothesis of this study was that the combination of PFMT and pessary wearing leads to a more substantial difference between the strain of the muscle in rest and the strain of the muscle during maximum contraction, compared to pessary wearing alone. In other words, it was expected that women in the Combined therapy group would exhibit improved muscle contraction and consequently reveal a more negative MPS after treatment, opposed to the Pessary therapy group.

The assumption that women in the Pessary therapy group achieve improved muscle contraction can be explained by the role of the pessary in maintaining organ position and reducing symptoms such as vaginal bulging. When the pelvic organs are no longer descending into the LH, it seems reasonable that women would experience improved contraction ability. However, this explanation does not account for the conflicting outcomes observed in the Combined therapy group.

4.3 Biometric parameters

For the biometric parameters, it was expected that the hiatal area in rest and the difference in AP-diameter between rest and maximum contraction would decrease according to existing literature.[1,16] However, the quantitative results present a different outcome.

Hiatal area

According to Table 3, the Combined therapy group reveals a significant increase of the hiatal area. A potential explanation for this could be the fact that women with POP may have both underactive and overactive pelvic floor muscles.[34] In case of an overactive pelvic floor, the pelvic floor muscles tend to be excessively tense. This makes relaxation training for the patient necessary, followed by instruction on functional use of the pelvic floor muscles. Consequently, it is reasonable that the resting area of the LH was initially reduced due to continuous pelvic floor tension. After PFMT, the tension in the pelvic floor is decreased, leading to relaxation and enlargement of the hiatal area in rest.

AP-diameter

For the AP-diameter, an increase was observed for both groups, although these observed differences are not statistically significant. It is important to note that the sample size was in fact too small for statistical analyses.

However, a possible explanation for the observed increase in AP-diameter is that TPUS represents a snapshot of the real behaviour of the pelvic floor muscles. It is imaginable that some women may not have comprehended the instructions adequately, leading to suboptimal muscle contraction or uncertainty regarding the specific muscles to engage.

Another potential explanation for the deviating results in both hiatal area and AP-diameter, concerns the accuracy of the measurements. The segmentation of the hiatal area and AP-diameter was performed according to the method described in the article by Kruger et al.[26] However, the exemplary ultrasound image depicted in their article provides clear boundaries of the hiatal area. Contrarily, the images employed in this study obtained lower resolution, giving challenges in accurate segmentation of the hiatal boundaries.

To avoid all these issues in future research, it is recommended to utilize a larger dataset to identify and reduce potential outliers. Additionally, consideration of the baseline resting tonus of all participants would help providing context and possible explanations for the observed results.

4.4 Correlation between parameters

According to Pearson's correlation coefficient, there is no significant correlation between the MPS, the area of the LH at rest, and the difference in the AP-diameter of the LH between rest and maximum contraction. This is not surprising, given the variation in outcomes.

When correlation between these parameters would exist, the added value of the MPS can be questioned. Biometric parameters are easier to determine compared to MPS, since running of the strain software is a time-consuming process. Therefore, it must be considered whether strain determination is feasible for clinical purposes. However, further research with larger sample size is necessary to evaluate the additional value of strain compared to other parameters.

4.5 Interobserver variability

As discussed earlier, correct segmentation of the PRM is crucial for correct calculation of the strain. Therefore, it is reasonable that segmentation by different observers could significantly impact the outcome of the results. When examining the outcomes of the comparison metrics from Table 5, it can be observed that the DSC ranges from 0.43 to 0.51, indicating a moderate level of overlap between the segmented regions. Furthermore, the 95% HD fluctuates between 6.59 and 7.42, indicating susceptibility of the segmentations to outliers. However, the average HD, ranging between 2.09 and 2.41 mm, implies proximity of the segmentation boundaries.

Although the DSC and 95% HD do not return impressive results, the average HD demonstrates acceptable outcomes. Given that the metrics were assessed based on only six segmentations, the presence of outliers is reasonable. Visual inspection of the segmentations of the three observers shows acceptable overlap either, when there was consensus about the avulsion site. More important is: What is the impact of the different observers on the MPS?

Observing Table 6, there seems no clear difference between good or poor overlap among the different observers. Looking at Figures 15 and 17, the MPS values for the segmentation without consensus was even more similar among the observers, than the MPS value for the segmentation with good overlap. This raises the question: how confidently can be relied on the strain outcome?

Due to the deviating outcomes concerning the accumulated displacement estimates, accumulated principal strain, and mean principal strain, no definitive conclusions can be drawn regarding the interobserver results. Examination of the software is needed to assess the accuracy of the strain results. However, considering the lack of consensus among observers, it is relevant to question the accuracy of the segmentations.

Segmentation of the PRM remains challenging, despite the pelvic floor muscles appearing brighter on ultrasound than surrounding tissue, due to higher collagen concentrations.[35] Also, standardized protocols are lacking. In this study, Observers 2 and 3 were instructed by Observer 1, who is an expert in the field of PRM segmentation. Despite this being a potential bias, Observers 2 and 3 did not always agree with Observer 1, raising the question of whose segmentation is correct. The answer remains inconclusive, and the variability observed highlights the human element involved in this process.

To overcome this problem, automatic segmentation by means of a deep learning network should be considered. This already exists for segmentation of the PRM, although it is not currently working on the newer DICOM files and is only trained by the segmentations of one observer. Therefore, it is important to train the network with new segmentations of different observers to overcome observer bias.

5. Conclusion

In conclusion, this retrospective study investigated the impact of adding PFMT to pessary wearing for the treatment of POP, as a first step towards to predict who will benefit from PFMT and who will not. The study focused on assessing the strain and biometric parameters of the PRM by 4D-TPUS. The key findings are:

- Hiatal area increased significantly post-treatment in the Combined therapy group.
- No significant changes in AP-diameter of the LH and MPS were observed between first and follow-up ultrasounds within or between groups.
- Observer segmentations showed moderate overlap, but similar MPS values, suggesting misleading results of the strain outcome.

However, the unexpected variability of the accumulated displacement estimates and accumulated principal strain, prevents the formulation of a conclusion regarding the additional value of PFMT to pessary wearing for the treatment of POP. Therefore, this study highlights the need for further research with a prospective dataset and larger study population. Besides, strain software and image tracking need to be optimized and automatic image segmentation via deep learning needs to be implemented.

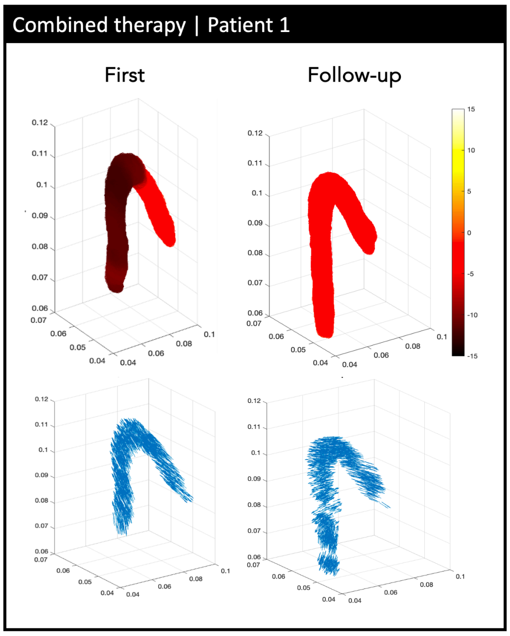
References

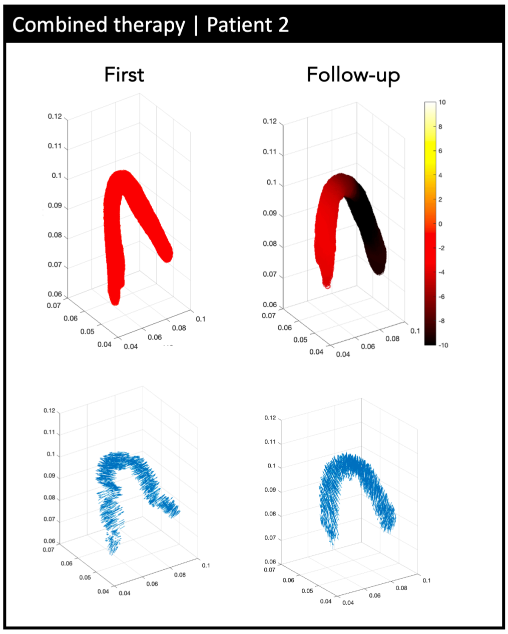
1. Sainz-Bueno, J. A. *et al.* Quantification of 3/4D ultrasound pelvic floor changes induced by postpartum muscle training in patients with levator ani muscle avulsion: A parallel randomized controlled trial. *Quant Imaging Med Surg* **12**, 2213–2223 (2022).
2. Giroux, M. *et al.* Effectiveness of pessary and pelvic floor physiotherapy alone and in combination for management of stress and mixed urinary incontinence: a literature review. *Gynecol Pelvic Med* **3**, 3–3 (2020).
3. Richter, H. E. *et al.* Continence pessary compared with behavioral therapy or combined therapy for stress incontinence: A randomized controlled trial. *Obstetrics and Gynecology* **115**, 609–617 (2010).
4. Li, C., Gong, Y. & Wang, B. The efficacy of pelvic floor muscle training for pelvic organ prolapse: a systematic review and meta-analysis. *Int Urogynecol J* **27**, 981–992 (2016).
5. Wang, T., Wen, Z. & Li, M. The effect of pelvic floor muscle training for women with pelvic organ prolapse: a meta-analysis. *Int Urogynecol J* **33**, 1789–1801 (2022).
6. Weintraub, A. Y., Glinter, H. & Marcus-Braun, N. Narrative review of the epidemiology, diagnosis and pathophysiology of pelvic organ prolapse. *International Braz J Urol* **46**, 5–14 (2020).
7. Bohlin, K. S., Ankardal, M., Nüssler, E., Lindkvist, H. & Milsom, I. Factors influencing the outcome of surgery for pelvic organ prolapse. *Int Urogynecol J* **29**, 81–89 (2018).
8. Shek, K. L. & Dietz, H. P. Assessment of pelvic organ prolapse: a review. *Ultrasound in Obstetrics and Gynecology* **48**, 681–692 (2016).
9. Das, S. *et al.* 3D Ultrasound Strain Imaging of Puborectalis Muscle. *Ultrasound Med Biol* **47**, 569–581 (2021).
10. Federatie Medisch Specialisten. *Richtlijn Prolaps.* https://richtlijnendatabase.nl/richtlijn/prolaps/prolaps_-_startpagina.html (2023).
11. Espiño-Albela, A., Castaño-García, C., Díaz-Mohedo, E. & Ibáñez-Vera, A. J. Effects of Pelvic-Floor Muscle Training in Patients with Pelvic Organ Prolapse Approached with Surgery vs. Conservative Treatment: A Systematic Review. *J Pers Med* **12**, (2022).
12. Wiegersma, M. *et al.* Predictors of Success for Pelvic Floor Muscle Training in Pelvic Organ Prolapse. *Phys Ther* **99**, 109–117 (2019).
13. Wallace, S. L., Miller, L. D. & Mishra, K. Pelvic floor physical therapy in the treatment of pelvic floor dysfunction in women. *Curr Opin Obstet Gynecol* **31**, 485–493 (2019).
14. Razzak, L. Chapter 2: Pathophysiology of Pelvic Organ Prolapse. in *Pelvic Floor Disorders* (InTech, 2018). doi:10.5772/intechopen.76629.
15. Basnet, R. Impact of pelvic floor muscle training in pelvic organ prolapse. *International Urogynecology Journal* **32**, 1351–1360 (2021).
16. Nyhus, M. *et al.* Effect of preoperative pelvic floor muscle training on pelvic floor muscle contraction and symptomatic and anatomical pelvic organ prolapse after surgery: randomized controlled trial. *Ultrasound in Obstetrics and Gynecology* **56**, 28–36 (2020).
17. Deegan, E. G., Stothers, L., Kavanagh, A. & Macnab, A. J. Quantification of pelvic floor muscle strength in female urinary incontinence: A systematic review and comparison of contemporary methodologies. *Neurourol Urodyn* **37**, 33–45 (2018).

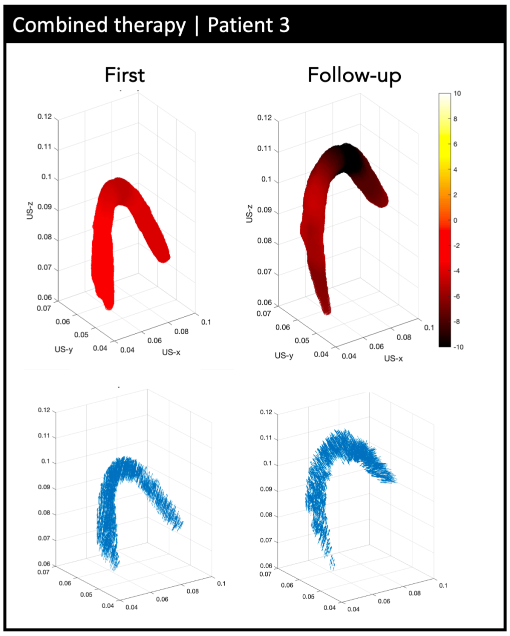
18. Hoff Braekken, I., Majida, M., Engh, M. E. & Bø, K. Morphological Changes After Pelvic Floor Muscle Training Measured by 3-Dimensional Ultrasonography A Randomized Controlled Trial. *Obstet Gynecol* **115**, 317–341 (2010).
19. Volløyhaug, I., Mørkved, S., Salvesen, O. & Salvesen, K. A. Assessment of pelvic floor muscle contraction with palpation, perineometry and transperineal ultrasound: A cross-sectional study. *Ultrasound in Obstetrics and Gynecology* **47**, 768–773 (2016).
20. Navarro Brazález, B. *et al.* The evaluation of pelvic floor muscle strength in women with pelvic floor dysfunction: A reliability and correlation study. *Neurourol Urodyn* **37**, 269–277 (2018).
21. Brady, B., King, G., Murphy, R. T. & Walsh, D. Myocardial strain: a clinical review. *Ir J Med Sci* **192**, 1649–1656 (2022).
22. Jelovsek, J. E., Maher, C. & Barber, M. D. Pelvic organ prolapse. *Lancet* **369**, 1027–1038 (2007).
23. Bø, K. *et al.* Are hypopressive and other exercise programs effective for the treatment of pelvic organ prolapse? *Int Urogynecol J* **34**, 43–52 (2023).
24. Hendriks, H. J. M. *KNGF-richtlijn Stress (urine-)incontinentie*. www.kngfrichtlijnen.nl. (2011).
25. van den Noort, F., Grob, A. T. M., Slump, C. H., van der Vaart, C. H. & van Stralen, M. Automatic segmentation of puborectalis muscle on three-dimensional transperineal ultrasound. *Ultrasound in Obstetrics and Gynecology* **52**, 97–102 (2018).
26. Kruger, J. A., Heap, S. W., Murphy, B. A. & Dietz, H. P. How best to measure the levator hiatus: Evidence for the non-Euclidean nature of the 'plane of minimal dimensions'. *Ultrasound in Obstetrics and Gynecology* **36**, 755–758 (2010).
27. Dietz, H. P., Shek, C. & Clarke, B. Biometry of the pubovisceral muscle and levator hiatus by three-dimensional pelvic floor ultrasound. *Ultrasound in Obstetrics and Gynecology* **25**, 580–585 (2005).
28. Webb, J. M. *et al.* Comparing deep learning-based automatic segmentation of breast masses to expert interobserver variability in ultrasound imaging. *Comput Biol Med* **139**, (2021).
29. Taha, A. A. & Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med Imaging* **15**, (2015).
30. Zou, K. H. *et al.* Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Acad Radiol* **11**, 178–189 (2004).
31. Sammarco, A. G. *et al.* Interactions among pelvic organ protrusion, levator ani descent, and hiatal enlargement in women with and without prolapse. *Am J Obstet Gynecol* **217**, 614.e4-614.e7 (2017).
32. Alves, F. K. *et al.* A pelvic floor muscle training program in postmenopausal women: A randomized controlled trial. *Maturitas* **81**, 300–305 (2015).
33. Tinelli, A. *et al.* Age-related pelvic floor modifications and prolapse risk factors in postmenopausal women. *Menopause* **17**, 204–212 (2010).
34. Grimes, W. R. & Stratton, M. Pelvic Floor Dysfunction. https://www.ncbi.nlm.nih.gov/books/NBK559246/ (2023).
35. Limbeek-van den Noort, F. van. Automatic analysis of transperineal ultrasound images. (2021).

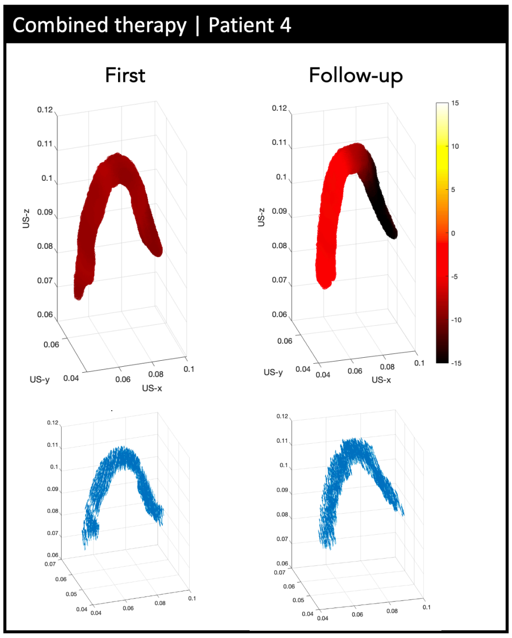
Appendix

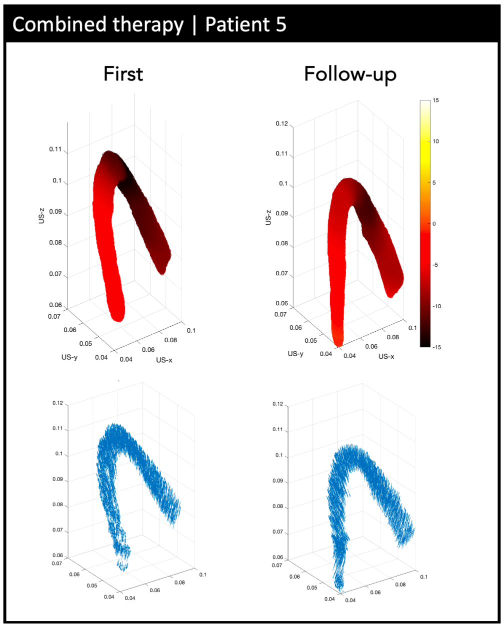
Appendix I: Accumulated principal strain of the Combined therapy group

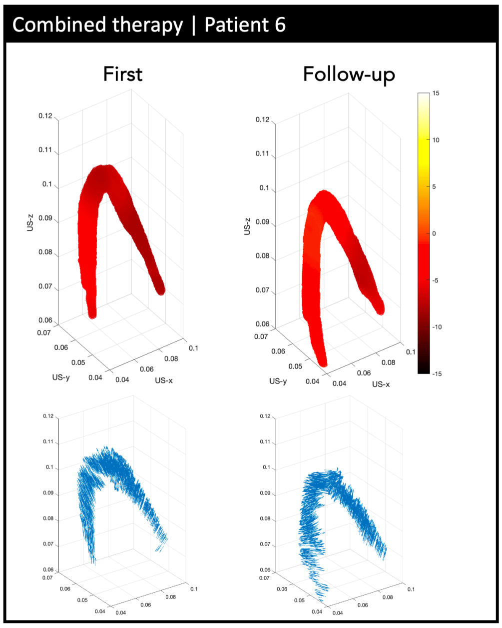


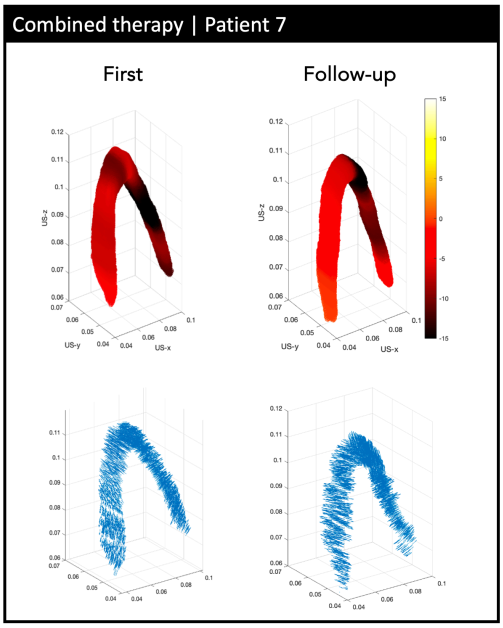




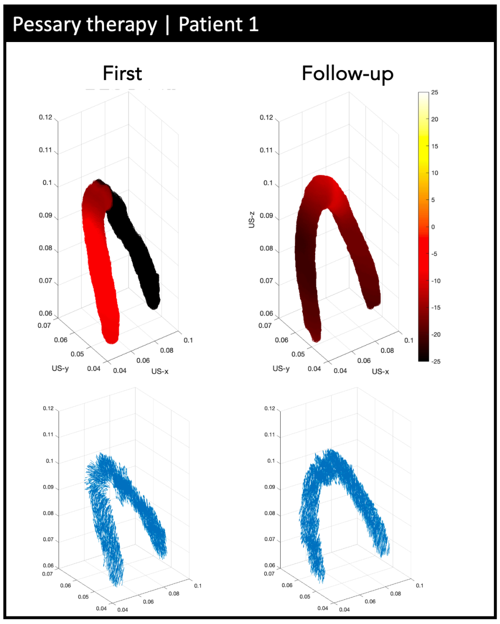


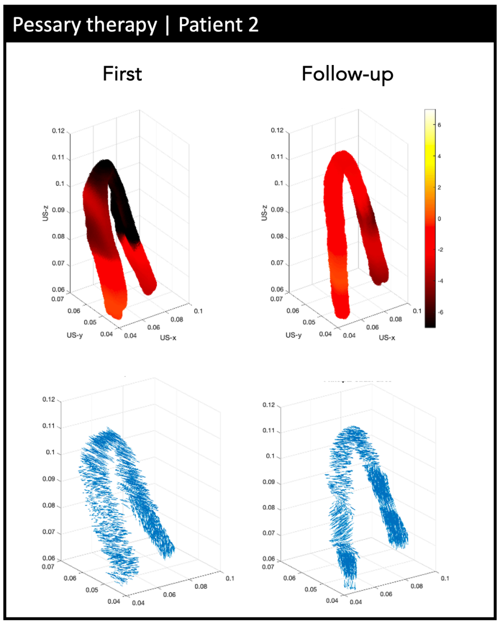

Combined therapy | Patient 5
First
Follow-up

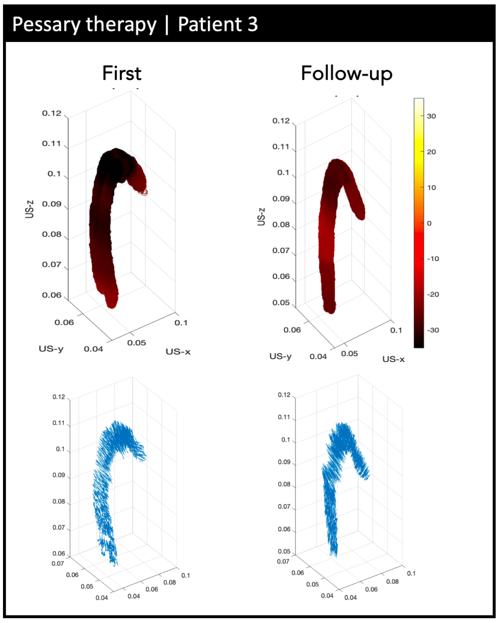

Combined therapy | Patient 6
First
Follow-up

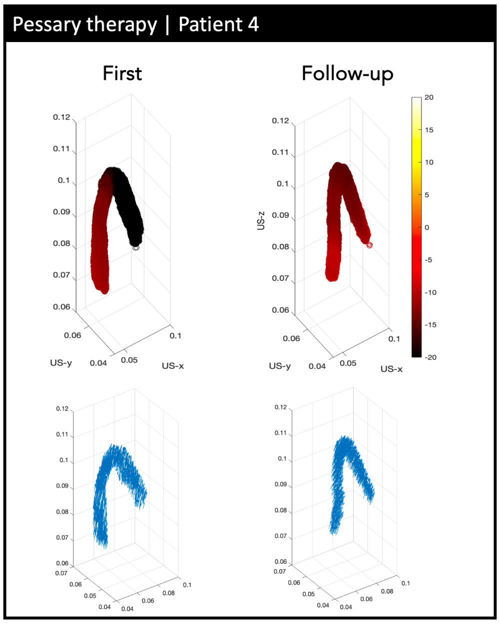

Combined therapy | Patient 7
First
Follow-up

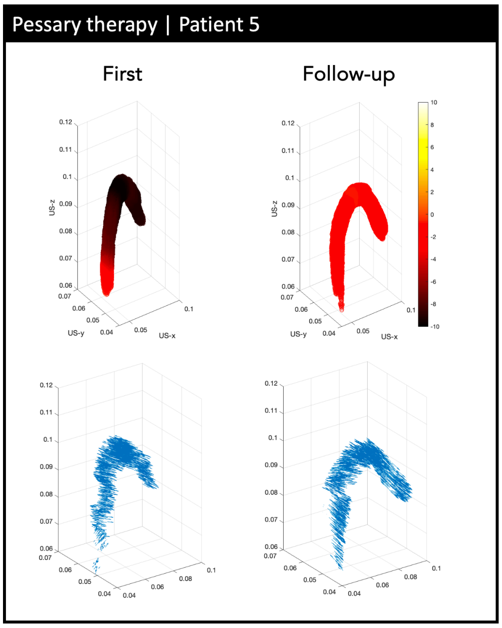
Appendix II: Accumulated principal strain of the Pessary therapy group

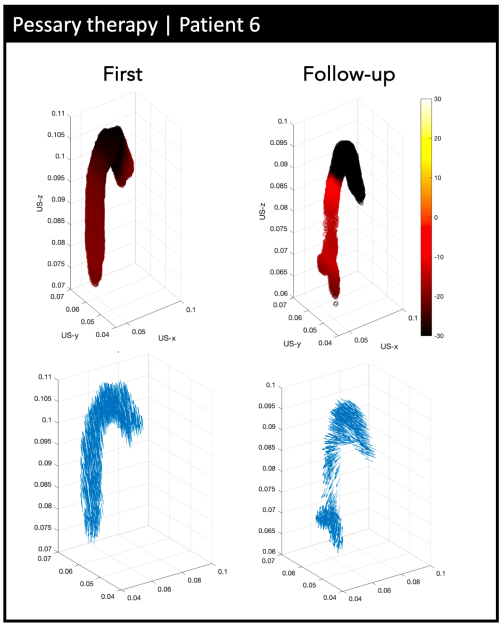


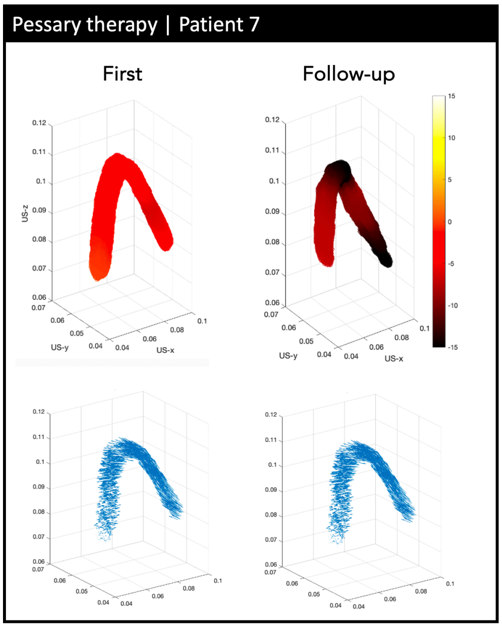


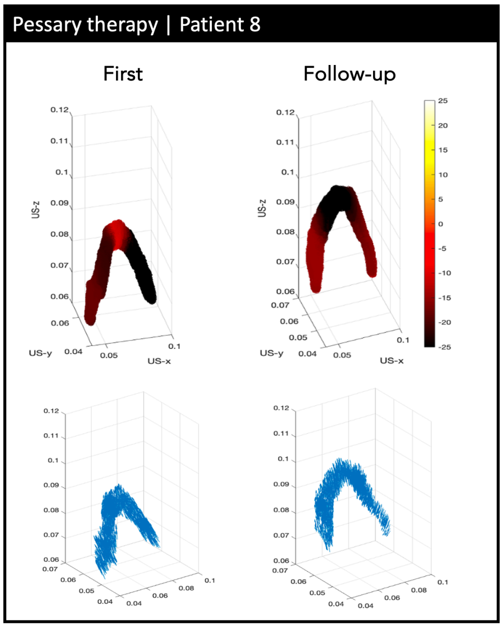


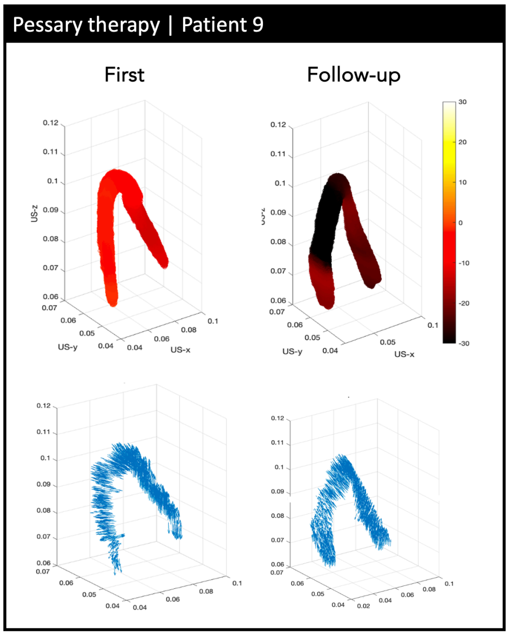


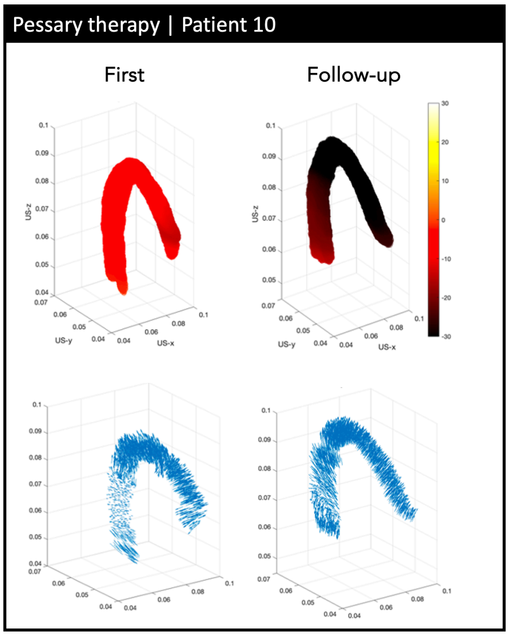

Pessary therapy | Patient 5
First
Follow-up
US-z
US-y
US-x


Pessary therapy | Patient 6
First
Follow-up
US-z
US-y
US-x


Pessary therapy | Patient 7
First
Follow-up
US-z
US-y
US-x


Pessary therapy | Patient 8
First
Follow-up
US-z
US-y
US-x


Pessary therapy | Patient 9
First
Follow-up
US-z
US-y
US-x


Pessary therapy | Patient 10
First
Follow-up
US-z
US-y
US-x

Appendix III: Visual analysis

Combined therapy

Pt.	Avulsion	First/FU	Contraction	Tracking	MPS-value at max. contraction (%)
1	Complete unilateral	First	Good	Accurate	-10,03
		FU	Good	Accurate	-3,64
2	No avulsion	First	Poor	Accurate	-1,43
		FU	Moderate	Moderately accurate	-8,94
3	Complete bilateral	First	Poor	Accurate	-3,94
		FU	Moderate	Accurate	-7,16
4	Partial unilateral left & complete unilateral right	First	Poor	Moderately accurate	-8,69
		FU	Moderate	Moderately accurate	-5,32
5	No avulsion	First	Poor	Accurate	-9,86
		FU	Poor	Accurate	-7,71
6	Complete unilateral	First	Moderate	Accurate	-4,60
		FU	Moderate	Accurate	-1,74
7	No avulsion	First	Good	Moderately accurate	-7,79
		FU	Good	Accurate	-5,52

Pessary therapy

Pt.	Avulsion	First/FU	Contraction	Tracking	MPS-value at max. contraction (%)
1	No avulsion	First	Good	Moderately accurate	-21,83
		FU	Good	Accurate	-14,18
2	No avulsion	First	Moderate	Accurate	-4,01
		FU	Moderate	Moderately accurate	-4,04
3	Complete unilateral	First	Good	Accurate	-31,81
		FU	Good	Moderately accurate	-25,12
4	Complete bilateral	First	Good	Inaccurate	-19,72
		FU	Moderate	Moderately accurate	-13,12
5	Partial unilateral	First	Good	Moderately accurate	-8,88
		FU	Moderate	Moderately accurate	-6,82
6	Partial unilateral left & complete unilateral right	First	Moderate	Inaccurate	-23,61
		FU	Good	Inaccurate	-24,32
7	No avulsion	First	Poor	Accurate	-3,23
		FU	Moderate	Moderately accurate	-9,91
8	No avulsion	First	Moderate	Inaccurate	-20,83
		FU	Moderate	Inaccurate	-20,45
9	No avulsion	First	Poor	Moderately accurate	-5,57
		FU	Moderate	Inaccurate	-22,97
10	No avulsion	First	Good	Accurate	-8,03
		FU	Good	Accurate	-30,62