REVIEW ARTICLE

Diagnostic value of pelvic floor ultrasonography for diagnosis of pelvic organ prolapse: a systematic review

Yuanjing Gao[1] · Zhiwei Zhao[2] · Yanlin Yang[1] · Meiqin Zhang[1] · Jiang Wu[3] · Yali Miao[1]

Abstract

Introduction and hypothesis Pelvic organ prolapse (POP) is a common medical condition universally. In addition to physical examination, experts have increasingly turned their attention to ultrasound in diagnosing POP for its low cost and dynamic imaging. The aim of this paper is to provide an overview of the methods of pelvic floor ultrasound in diagnosing POP, which has been lacking up till now.

Methods We included original papers comparing the outcome of the Pelvic Organ Prolapse Quantification system and ultrasound, published from 2008 to present in English, using electronic databases (MEDLINE, EMBASE, CENTRAL, PUBMED). All stages of the review were conducted in parallel by two reviewers.

Results Fifteen papers were included. We found that current methods have advantages and limitations. The main methods are to measure levator hiatus-related parameters and distances between the lowest point of the pelvic organs and reference lines during Valsalva maneuver, contraction, and at rest.

Conclusions Pelvic floor ultrasound is valuable in diagnosing POP, yet suffers from a weakness in precision compared with physical examination. From the existing research, we found that the differences in baseline data such as weight, height, ethnicity, etc., may affect the cutoffs of the above-mentioned parameters. Further research is required to find one appropriate cutoff for each parameter, even if it is necessary to set group values for every parameter according to varying situations.

Keywords Pelvic floor ultrasound · Pelvic organ prolapse · Clinical examination · POP-Q

Introduction

Female pelvic organ prolapse (POP) is defined as descent of the female pelvic organs (uterus, bladder, rectum or small intestine), forming herniation of the anterior or posterior vaginal wall or vaginal apex in the vagina [1]. It is generally

accepted that POP is a highly prevalent disorder all over the world [2–5]. On the one hand, more and more women are at risk. Women's lifetime risk of needing prolapse surgery is one tenth to one fifth, with over 300,000 surgeries in the USA every year [6]. The number of women who suffer from POP is forecast to rise by 46%, up to 4.9 million, by 2050 [7]. On the other hand, the diagnosis of POP is nontrivial. Around 8.33% women in the UK claim symptoms of POP [2]. These symptoms include a visible or palpable vaginal bulge, or the feeling of vaginal protrusion, stress urinary incontinence, fecal incontinence, chronic straining to defecate, chronic back pain, etc. [8]. Among them, the first one the most particular symptom of all compartments [9, 10]. However, 41–50% of patients are diagnosed with certain degrees of prolapse by clinical examination and only 3–6% are diagnosed by symptoms [11]. Therefore, investigating the diagnosis of POP is urgent and critical.

The traditional diagnosis of POP includes symptom evaluation and clinical examination. With regard to symptom evaluation, in spite of there being some questionnaires, such as the Pelvic Floor Distress Inventory-short form 20 (PFDI-20) and

Yuanjing Gao and Zhiwei Zhao contributed equally to this work

✉ Jiang Wu
jw@scu.edu.cn

✉ Yali Miao
miaoyali2006@163.com

[1] Department of Obstetrics and Gynecology, Key Laboratory of Birth Defects and Related Diseases of Women and Children of MOE, West China Second University Hospital, West China Campus, Sichuan University, No.17 People's South Road, Chengdu 610041, China

[2] West China School of Basic Medical Sciences & Forensic Medicine, Sichuan University, Chengdu, China

[3] Deep Underground Space Medical Center, West China Hospital, Sichuan University, No. 37 Guoxuexiang, Chengdu, China

the Pelvic Floor Impact Questionnaire-short form-7 (PFIQ-7), there is not yet any standard questionnaires to assess symptoms in detail [10]. Thus, the evaluation results may be impacted by many additional factors owing to people's subjective judgment.

With regard to clinical examination, two methods are commonly used. The Pelvic Organ Quantification (POP-Q) system [1, 12], recommended by the International Continence Society, was used to assess the extent of POP. Six points are set, scattered on the vaginal wall, then the distance between the hymen and these points are measured on maximal Valsalva maneuver. Simultaneously, the total vaginal length, the genital hiatus, and the length of the perineal body need to be noted. In general, Ba, C, and Bp stand for anterior, central, and posterior compartments respectively. These outcomes should be rated in stages according to Table 1. The Baden–Walker Halfway Scoring System [13] is the other method shown in Table 1. Moreover, adding 1 cm strategically would result in an increase in the severity of assessment. However, inter-agreement is not perfect with this system. Pham et al. [14] counted the frequency of ways describing the degree of POP by clinical surgeons who are members of the American Urogynecologic Society (AUGS) and the International Continence Society (ICS). Over three quarters of surgeons claimed that they were inclined to use POP-Q; around 14% of surgeons use the Baden–Walker System. The remaining surgeons use descriptive words (9%) or another system (1.7%).

Most noteworthy is that some researchers have intended to rectify the POP-Q to sufficiently identify symptomatic prolapse in recent years. Wiegersma et al. [15] testified the poor relation between the outcome of the present POP-Q clinical examination system and prolapse symptom occurrence. Then, Dietz and Mann [16] reported that the POP-Q system should be revised. Prolapse of the anterior and posterior compartments of less than 1 cm ought to be regarded as normal and stage 1 uterine prolapse seemed clinically significant. This

introduced a definition called clinically significant prolapse: that of the anterior and posterior compartments indicates POP-Q stage 2 or higher and that of the center (apical) compartment was defined as stage 1 or higher. This difference was produced by the higher probability of symptom occurrence of the central compartment compared with the anterior and posterior compartments. Some researchers also see POP-Q stage 2 and higher as indicating a clinically significant prolapse of all compartments.

Pelvic floor ultrasound on diagnosing POP

In the past 30 years, it has become accepted by more and more scholars and surgeons that clinical examination and symptom evaluation alone are inadequate for grasping knowledge of the whole pelvic floor anatomy and function because they focus on surface anatomy rather than tridimensional structural abnormalities. Therefore, radiological techniques such as magnetic resonance imaging (MRI) have been an alternative for gynecologists and urologists to acquire more detailed information about pelvic floor structures to estimate the severity of prolapse, multicompartmental prolapse, and for making treatment plans. However, because of its high cost, limited availability, and poor dynamic imaging, it is too limited for widespread use, while ultrasound meets these requirements; thus, increasingly researchers have recently realized that ultrasound could be used in the detection of the field in pelvic floor diseases after Dietz et al. first pointed that ultrasound could be used in pelvic floor dysfunction (PFD) [17]. Pelvic floor ultrasound includes endovaginal, endoanal, transperineal ultrasound imaging (TPUS) and translabial ultrasound imaging (TLUS). The commonest means of diagnosing POP is the latter. In this paper, we use TPUS consistently. Current guidelines characterize the site of prolapse as the anterior vaginal wall, vaginal apex/uterine prolapse (apical prolapse), and posterior vaginal wall [18]. Some other researchers are inclined to

Table 1 Stages of Baden–Walker system and Pelvic Organ Prolapse Quantification (POP-Q) system measurements

Baden–Walker system		Pelvic organ prolapse quantification system	
Grade	Description	Stage	Description
0	Normal position for each respective site, no prolapse	0	No prolapse
1	Descent halfway to the hymen	I	Greater than 1 cm above the hymen
2	Descent to the hymen	II	1 cm or less proximal or distal to the plane of the hymen
3	Descent halfway past the hymen	III	Greater than 1 cm below the plane of the hymen, but protruding no farther than 2 cm less than the total vaginal length
4	Maximal possible descent for each site	IV	Eversion of the lower genital tract is complete

Reproduced with permission from: Onwude [59]

classify the sites of organ descent as anterior, central, and posterior compartments [19]. To reduce misunderstanding, we use this latter system uniformly.

Generally, ultrasound physicians adopt two methods of rating the severity or assess the possibility of POP using 2D/3D ultrasound. The first one is to quantify the vertical maximal distance between the horizontal line (H line) placed through the inferior margin of the symphysis pubis and prolapsed organ on the midsagittal plane (Fig. 1). The second one is to measure the anteroposterior diameter and area or other indexes of levator hiatus, sketched as the area surrounded by the pubovisceral muscle, symphysis pubis, and inferior pubic ramus at the level of minimum hiatal dimensions on the axial plane, to estimate the degree of hiatus enlargement (Fig. 2) [20]. Furthermore, the puborectalis avulsion could be detected when loss of continuity is noticed between the muscle and pelvic sidewall on no less than one slice (Fig. 3). Besides, with the development of 4D ultrasound, gaining a rendered volume of the levator ani hiatus is another method (Fig. 4).

Various studies have been conducted by diagnosing POP with the help of ultrasound and indeed improving accuracy. To our knowledge, there is no overview of all the methods. This systematic review is a detailed and comprehensive summary of the current methods used in pelvic floor ultrasound for POP diagnosis, and offers an objective and fair comparison of its advantages and disadvantages. The aim is to evaluate how current ultrasound techniques aid POP diagnosis and to provide potential future directions for utilizing ultrasound. In addition, this paper proposes the possible reasons for different biases in various methods, and points out the direction for further
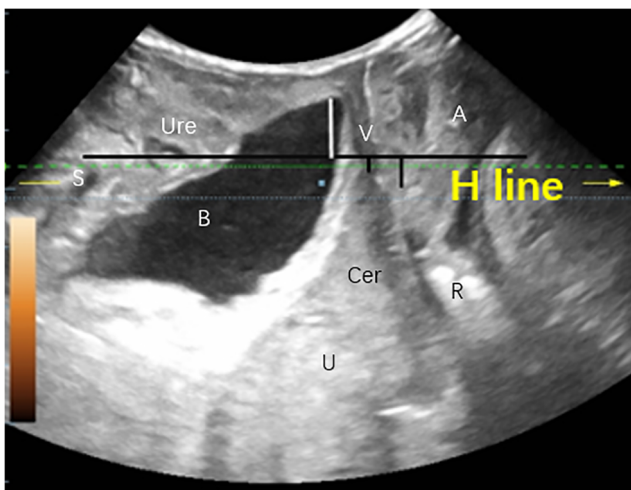
optimization of diagnostic methods in the future. We also analyzed the effects of BMI, height, race, pregnancy, and changes in normal values of pelvic floor data at different ages in normal women on the diagnostic reference threshold for ultrasound. It is beneficial for the ultrasound physicians to understand the current field of development, improve the methods, and promote the diagnostic value.
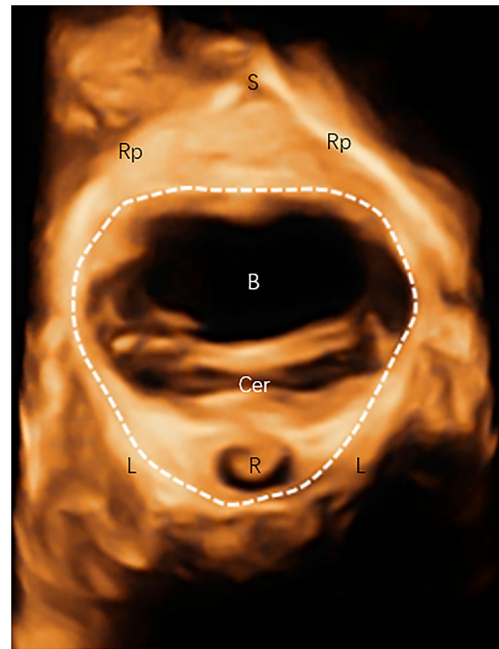

Fig. 2 The axial view of the levator hiatus surrounded by pubovisceral muscle, symphysis pubis, and inferior pubic ramus at the level of minimum hiatal dimensions on the axial plane, with the hiatal area indicated by the *dashed line*. S symphysis pubis, Rp ramus pubis, B bladder, Cer cervix, R rectum, L levator ani muscle

Fig. 1 The midsagittal plane image obtained on maximal Valsalva on the 3D transperineal ultrasound. The H line is the *black line*. The *vertical lines* indicate the maximal descent of the bladder (*B*), uterus (*U*), and rectal ampulla (*R*) relative to the symphysis pubis (*S*). Ure urethra, V vagina, Cer cervix, A anal canal, R rectum
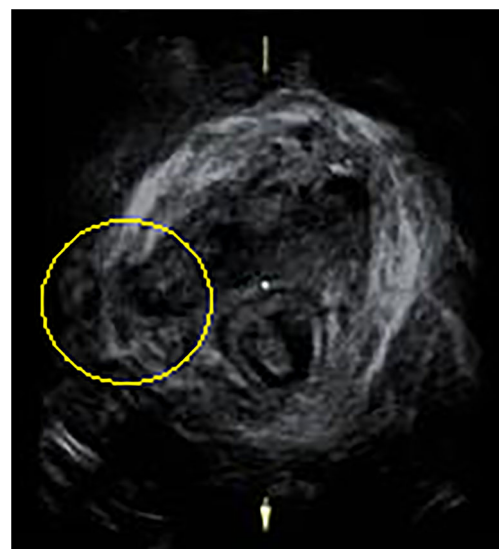

Fig. 3 Avulsion injury detected on 3D transperineal ultrasound image of women with pelvic organ prolapse (POP). The *yellow dashed circle* includes the avulsion injury structure
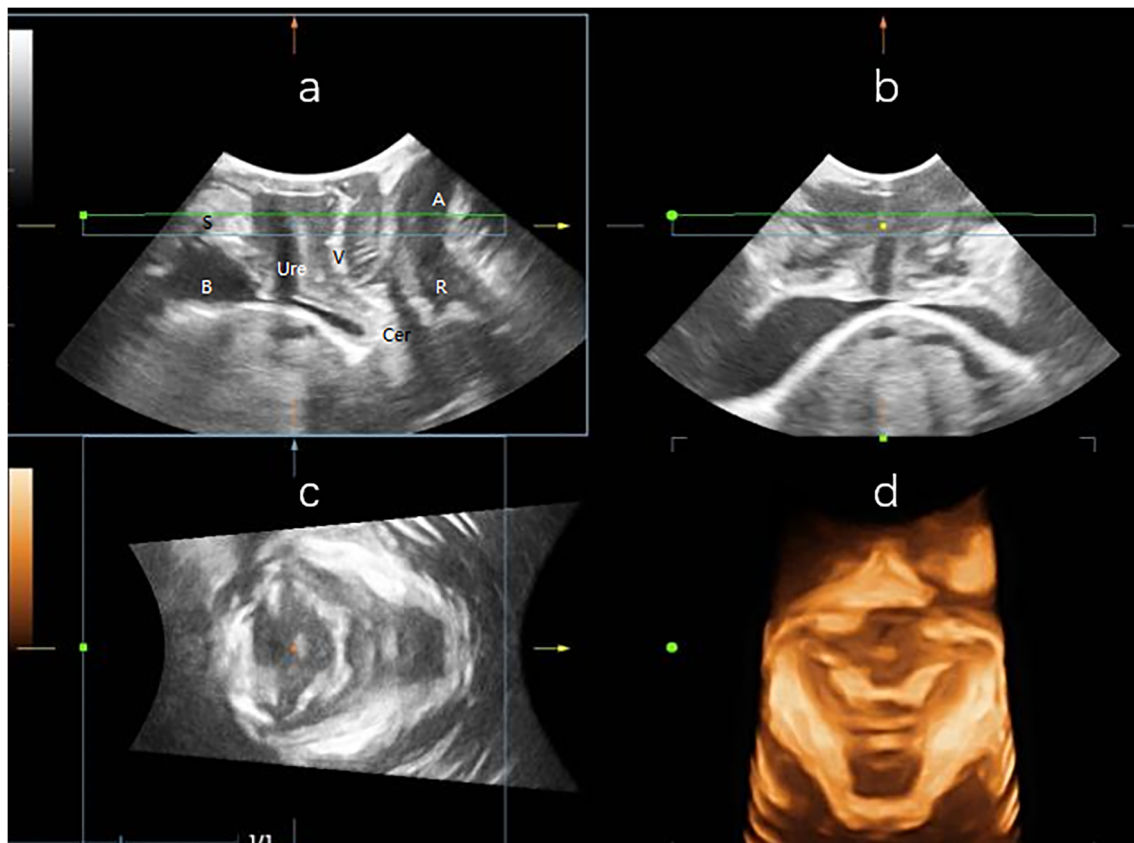
Fig. 4 **a**: The midsagittal plane, **b**: coronal plane, **c**: axial plane on 3D pelvic floor ultrasound, **d**: rendered volume on 3D pelvic floor ultrasound. **S**: symphysis pubis, **Ure**: urethra, **B**: bladder, **V**: vagina, **Cer**: cervix, **A**: anal canal, **R**: rectum

Materials and methods

This systematic literature search was performed by the first two authors: Gao and Zhao. The electronic databases PubMed, Medline, Embase, and CENTRAL were searched from 2008 to 14 June 2019. Our inclusion criteria were as follows: written in English; using ultrasound to diagnose POP; comparison of diagnosing outcomes with POP-Q; full text is available. To obtain all the articles that meet the inclusion criteria of this system evaluation, we developed the following strategy, which is shown in the Appendices.

All articles were screened by Zhao and Gao separately according to the inclusion criteria. If we had different opinions of certain papers, Miao, who is an experienced pelvic floor specialist, would make a final decision. References of the relevant retrieved articles were cross-checked to find additional articles that had been neglected in the database search.

The full-text articles were assessed to collect data on study design, sample size, baseline data (parity, age, setting, BMI), POP staging, inclusion and exclusion criteria of the study population, ultrasound index, and modified Quality Assessment of Studies of Diagnostic Accuracy Studies (QUADAS-2) outcomes.

The quality of selected studies was evaluated according to the modified QUADAS-2 checklist.

Results

Out of initially thousands of papers, the final 17 are included and the selection process is shown in Fig. 5. The quality evaluation overview of the selected articles recorded in a QUADAS-2 form is shown in Appendix Table 7. No articles were considered to have a high risk of bias.

Levator hiatus and POP

Many researchers have found that the size of the levator hiatus has a positive association with the severity of POP. Therefore, ultrasound physicians have tended to find a substantial reference value for diagnosing POP. We included seven authors measuring related parameters of levator hiatus. They are the anteroposterior (AP) diameter of the levator hiatus and levator hiatus area (HA) or the left–right diameter of the levator hiatus (LR) as the main measurement indexes. Patients' behavior during the examination are also marked after the name. For

Fig. 5 Selection process of studies included in the systematic review. *n* number of articles
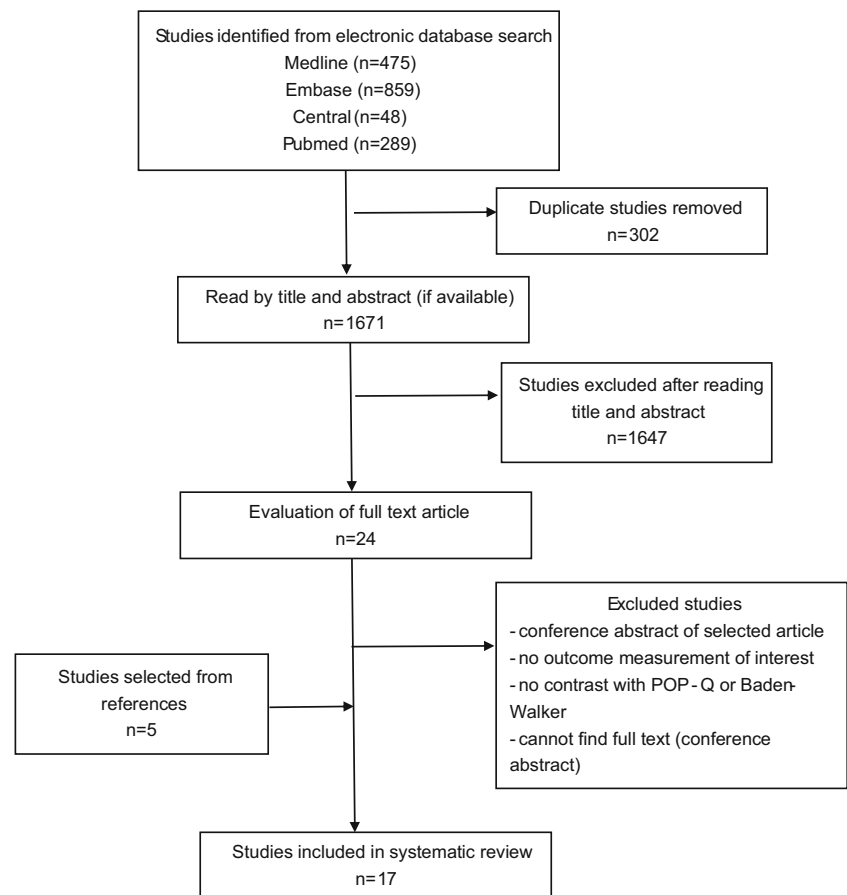
example, HAval, HArest, and HAtract means HA on Valsalva maneuver, at rest, and on pelvic floor contraction respectively.

Majida et al. [21] found a significant positive association between the stage of prolapse and HArest, HAval (both $p < 0.001$). There was a moderate agreement between HArest and HAval ($r_p = 0.62$). Clinically significant anterior compartment prolapse has been found to be strongly related to HAval and HArest ($p < 0.001$), but central and posterior prolapse has no association with HAval and HArest ($p = 0.152$ and $p = 0.406$). Detailed continuous data have been listed in Table 2. A limitation of this study is the composition of participants. This study comprised women with and without prolapse symptoms, who were originally included, to analyze the effects of pelvic floor muscle exercise on POP of the initial stage. Therefore, most of them had not intended to pursue medical treatment; thus, their data should not be compared with the main population seeking help for POP-related problems.

Ying et al.[22] measured the size of the levator hiatus in the process of various patients' maneuvers, by taking three evaluating metrics as reference: HA, AP, and left-to-right diameter of the levator hiatus (LR). Detailed data are summarized in Table 3. For both nulliparous women and women with POP, all the above-mentioned metrics increased from contraction,

through rest, to Valsalva. In addition, the HA, AP, and LR values of women with POP were significantly higher than those of nulliparous women ($p < 0.001$). The shape of the levator hiatus of women with POP is more circular than that of nulliparas. Besides, they observed whether or not the line named levator hiatus axis matching the midpoint of the inner edge of the symphysis pubis with the puborectalis overlaps with the pelvic floor axis. The intersection angle is obvious between the two axes if they fail to overlap and appeared in some women with POP. In 36 women in the POP group (72%) the levator hiatus axis departed from the pelvic floor axis, whereas in the nulliparous women, the two axes overlapped (Fig. 6). In 18 women in the POP group (36%) avulsion developed in the puborectalis. But we cannot overlook some of the limitations to this study. Two age groups are too broad; thus, it was hard to attribute the whole increase to POP. However, we also realized that this limitation existed in almost all the research studies, except for those that selected every participant from the normal population.

Wen and Zhou [23] found that the AP diameter had an excellent linear correlation ($r = 0.814$, $p < 0.001$) with HAval. This result testified that several parameters of HA were consistent to some extent. Furthermore, these parameters were significantly related to POP-Q stages ($p < 0.01$). Receiver

Table 2 Area of the levator hiatus at rest (HArest) and on Valsalva maneuver (HAval) according to POP-Q stages (from Majida et al. [21])

Stage	HArest (cm^2), $N = 149$		HAval (cm^2), $N = 149$		HAval ≥25 cm^2
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n (%)
0	2 (0)	16.1	2 (0)	21.2	0
1	43 (0)	20.1 (18.7–21.5)	39 (4)	26.2 (23.8–28.5)	20 (51.3)
2	76 (0)	22.9 (21.9–23.9)	71 (5)	30.1 (28.3–31.9)	55 (77.4)
3	27 (0)	26.7 (24.2–29.2)	20 (7)	35.4 (31.6–39.1)	17 (85.0)
4	1 (0)	23.1	0 (1)	–	0
p value	<0.001		<0.001		

operating characteristic (ROC) helped them to identify 6.0 cm for APval and 20cm^2 for HAval as a cutoff for determining clinically significant POP (AP: sensitivity 73%, specificity 52%; HA: sensitivity 76%, specificity 54%).

However, Albrich et al. [24] acquired 27.53 cm^2 for HAval as a cutoff (sensitivity, 70%; specificity, 69%) for diagnosing clinically significant prolapse with the aid of the Youden Index. The area under the curve (AUC) of HAval in women suffering from POP was 0.755 (95%CI, 0.696~0.814). The levator avulsions were detected in 20.7% women. In addition, they recorded baseline data about parity mode. Twenty women had undergone vaginal deliveries, there were 17 nulliparas, 7 women had undergone cesarean deliveries and no vaginal deliveries. The limitation is that they did not exclude patients who had had pelvic floor muscle training. The final data contained women who had undergone training to different degrees and did not divide participants into two groups according to before and after training to investigate.

Some scholars tried to second-process these parameters to gain a combination or simpler values. Dou et al. [25] put forward a two-step method to improve on the traditional method by measuring levator hiatus dimension to diagnose clinically significant POP. They analyzed the ultrasound indexes of 323 women such as HArest, HAval, APrest, APval, LRval, and the distance between HArest and HAval (ΔHArest-val). The ROC curve analysis for HAval yielded an AUC of 0.79 (95%CI, 0.73~0.86), which is the largest of all the parameters. The cutoff was decided to be 19.5 cm^2(sensitivity, 80%; specificity, 70%; Youden Index, 0.51). Moreover, using 25 cm^2 as the cutoff for HAval resulted in lower sensitivity and higher

specificity (sensitivity, 27%; specificity, 83%). Inspired by this outcome, the author used a two-step method that selected patients whose HA was less than 25 cm^2 into the second step to compare their parameters except for HAval. Finally, they discovered that a combination of HArest = 10 cm^2 (sensitivity, 95%; specificity, 30%) and ΔHArest-val = 5 cm^2 (sensitivity, 81%; specificity, 64%) was the best second-step parameter. This method had a better effect (sensitivity, 87%; specificity, 70%) than the one-step method. This paper offers a feasible idea for us to combine several efficient parameters to improve diagnostic accuracy.

Wen et al. introduced the concept of the Z score to second-process the ultrasound index obtained from patients. In 2016 [26], he provided two formulas: Z-HAval = (measured value – 17.19)/2.98, Z-APval = (measured value – 55.65)/5.48. The 90% reference range of Z-HAval was from −1.8 to +1.8, and that of Z-APval was from −2.0 to +2.0. The ROC showed that the cutoff of HAval for diagnosing POP-Q stage 2 was 20 cm^2 (sensitivity, 79%; specificity, 66%). That of APval was 6.0 cm (sensitivity, 68%; specificity, 65%). These two cutoffs were equivalent to Z-HAval of 1.0 and Z-APval of 1.0 accordingly. The cutoff of HAval and APval for diagnosing POP-Q stage 3 was 24 cm^2 (sensitivity, 85%; specificity, 83%) and 6.3 cm (sensitivity, 77%; specificity, 80%). These two values were equal to Z-HAval of 2.0 and Z-APval of 1.5.

In 2018 [27], Wen et al. updated the coefficient in the formula: Z-HAval = (measured value – 17.15)/3.11. The authors obtained a cutoff for HAval of 20.26 cm^2,

Table 3 Measured parameters in the POP group and the nulliparous group (mean ± SD) (from Ying et al. [22])

Group	AP (cm)			LR (cm)			HA (cm^2)		
	Rest	Valsalva	Contraction	Rest	Valsalva	Contraction	Rest	Valsalva	Contraction
Nullipara	4.31 ± 0.72	4.57 ± 0.74	4.08 ± 0.74	3.67 ± 0.58	3.98 ± 0.61	3.44 ± 0.52	11.22 ± 2.61	14.50 ± 4.47	9.18 ± 2.65
POP	5.32 ± 0.52	5.92 ± 0.54	4.96 ± 0.53	4.43 ± 0.51	4.89 ± 0.50	4.07 ± 0.52	17.01 ± 2.66	22.76 ± 3.72	15.07 ± 2.72
p values	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

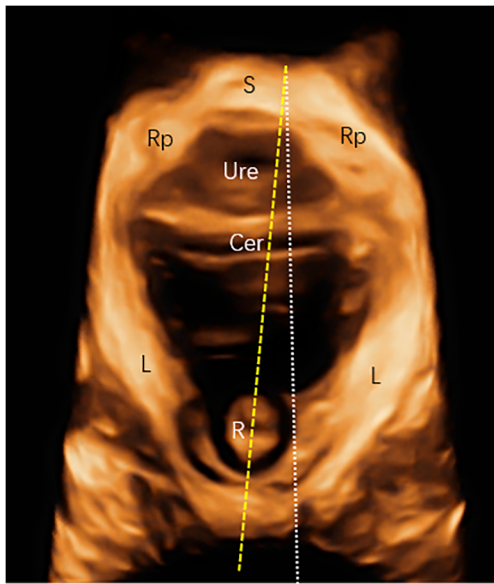*AP* anteroposterior, *LR* left-to-right, *HA* hiatus area

Fig. 6 Levator hiatus axis departs from the pelvic floor axis in women with POP. The *white dashed line* is the pelvic floor axis. The *yellow dotted line* is the levator hiatus axis. *S* symphysis pubis, *Rp* ramus pubis, *Cer* cervix, *R* rectum, *L* levator ani muscle

corresponding to Z-HAval of 1.0 (against POP stage 2 and higher: sensitivity 77% and specificity 60%; against substantial POP on ultrasound: sensitivity 84% and specificity 75%). In addition, the results revealed an excellent correlation between the HArest and HAval by Pearson's correlation coefficient analysis.

The works of Wen et al. introduced simpler parameters for processing the raw data obtained during ultrasound measurement, which is convenient for clinically more elegant staging, but the widely accepted reference values for HA and AP has not been determined. Even the data of women in the region of China have been changing over the past 2 years, but this practice is still suggestive.

Different from others, one author focused on whether or not HAtract is an appropriate parameter for diagnosing POP.

Nyhus et al. [19] tried to investigate the connection between pelvic contraction capability and POP. They set the 75th percentile as the upper limit of the normal range, which provided a lowest cutoff of abnormal HAval of 42 cm^2. Then, the authors put forward a formula: $\Delta AP/ area = 100*$ ((measurement$_{rest}$-measurement$_{contraction}$)/measurement$_{rest}$ and compared these values of normal women and women with clinically significant prolapse. The data of 555 women were finally included. The Δ AP for any compartment was 25.7% in women without POP, whereas it was 22.8% in women with POP ($p < 0.001$) and Δarea was 33.8% and 29.6% respectively ($p < 0.001$). This implied that women with clinically significant prolapse in any compartment have weaker pelvic floor muscles than women without. This conclusion showed a statistically significant difference in the anterior and central compartments, and not in the posterior compartment. However,

this paper provided raw data without advanced processing to testify the value of the hiatus area in diagnosing POP.

Majida et al. [21] and Ying et al. [22] provided exact and original data, but not reference values. However, some variation discipline between severity of POP and the dimension of HA could be found in Tables 2 and 3. Inspired by theirs and other previous works, increasing numbers of researchers have tried to define cutoffs using statistic tools. Even some considered devising new indexes with advanced processing of the original parameter values, to improve their accuracy or make values simpler. Limited by single-center works and an insufficient sample number, none of them could provide authoritative reference values for levator hiatus-related parameters.

Reference lines of pelvic floor ultrasound and POP

All the lines on the ultrasound image mentioned can be seen in Fig. 7. The horizontal line (H line) is a horizontal reference line that passes the inferior-most point of the symphysis pubis, which was first introduced by Dietz et al. [17]. The midpubic line (MPL) refers to the line drawn through the central point and parallel to the longitudinal axis of the symphysis pubic bone, which was an ordinary reference line in the MRI evaluation of POP [28]. The pubococcygeal line (PCL) is a straight reference line passing the inferior point of the symphysis pubis and the anorectal junction.

Volloyhaug et al. [29] recruited 590 parous women from the same population as the study performed by Nyhus et al. [19]. They were requested to take TPUS examinations using the H reference line and the POP-Q clinical examinations. The authors applied a substantial significant prolapse standard on ultrasound imaging (bladder descent more than 10 mm, rectal ampulla more than 15 mm below the H line, and the cervix more than 15 mm
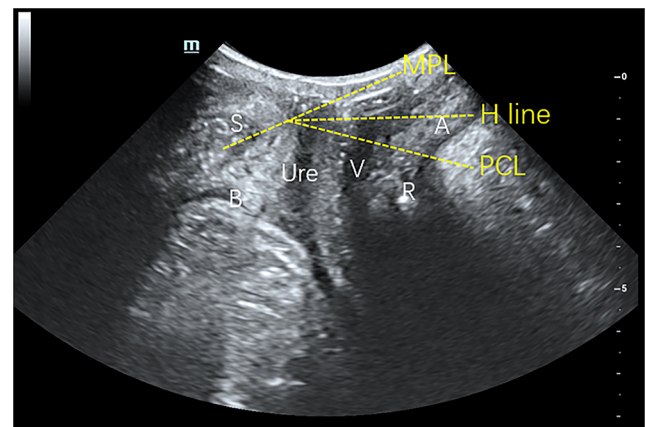

Fig. 7 The midsagittal image of the pelvic floor at rest shows the position of the H line, midpubic line (*MPL*), pubococcygeal line (*PCL*). The names of these three lines are marked in *yellow* beside the line. *S* symphysis pubis, *Ure* urethra, *B* bladder, *V* vagina, *A* anal canal, *R* rectum

above the H line) in diagnosing clinically significant prolapse. Finally, they gained the outcomes of using ultrasound to diagnose POP (anterior: sensitivity 62%, specificity 91%; central: sensitivity 85%, specificity 87%; posterior: sensitivity 38%, specificity 84%). Dietz et al. [30] and Lone et al. [31] carried out similar research. Dietz's results were overall: sensitivity 90%, specificity 64%; anterior: sensitivity 90%, specificity 64%; central: sensitivity 60%, specificity 77%; posterior: sensitivity 93%, specificity, 47%. Lone did not mention the criterion of clinical examination clearly; we think that POP-Q stage 1 and severer are seen as prolapse in this research. Besides, their diagnosing criterion on ultrasound has a slightly different point from Volloyhaug: the cervix more than 10 mm below the H line (anterior: sensitivity 59.0%, specificity 100.0%; central: sensitivity 39.3%, specificity 96.2%; posterior: sensitivity 69.0%, specificity 94.9%). It seems that more papers need to be analyzed to acquire a comprehensive result.

Najjari et al. [32] suggested a classification system to quantify cystocele (anterior compartment) by TPUS. They measured the distances between the furthest descending point of the bladder and the MPL and H lines respectively. All ultrasound data were compared with the outcome of POP-Q. It revealed that when ultrasound physicians shifted the probe, the viewing angle changed (Fig. 8). Then, the distance (D_{RV}) between the lowest point of the bladder at rest (P_R) and Valsalva (P_V) changed when using the horizontal line. But the distance is constant when using the MPL and these distances had been put

down representing for the mobility of patients's bladder during Valsalva maneuver. Patients had been separated into three groups (group I, >1 cm above the MPL; group II, <1 cm above the MPL; group III, below the MPL). The distinction among three groups is significantly great ($p < 0.00001$). The ultimate results signified that the lower bladder falls during the Valsalva maneuver, the greater distance it descends when they sorted patients into POP-Q stages 0 and 1 and POP-Q stages 2–4. The former range corresponded to group I and the latter range corresponded to groups II and III. It reached the kappa-coefficient value equal to 0.65. Additionally, they claimed that they were the first team to estimate the interrater agreement of TPUS in the assessment of cystoceles. $\kappa = 1.00$ for the classification of three groups between two examiners using the MPL as the reference line indicated an excellent correlation. This result could tell us that taking MPL as reference value is superior to taking H line, because using the former could reduce measurement deviation resulting from the shifting angle of the probe and perineal body. In theory, this discipline maybe suitable for three-compartment prolapse, but more research is still needed to test it.

Arian et al. [33] tried to compare the PCL with the H line. In this study, researchers distributed the 40 Iranian women with a history of POP into groups by two modes of classification depending on POP-Q in three compartments respectively. The first mode is to divide participants into with or without POP groups, the second is to divide them into POP-Q stage 3 and higher or stage 0 and 1 groups. The authors designed a plan to contrast the out-

Horizontal line
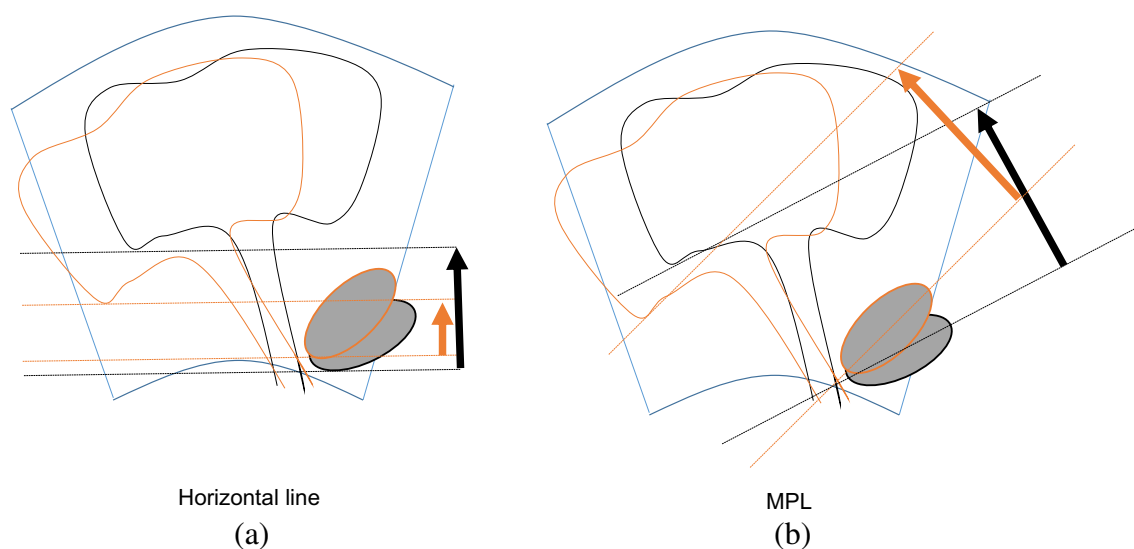
(a)

MPL

(b)

Fig. 8 Diagnosing cystocele by using **a** the H line and **b** the MPL as the reference line on transperineal ultrasound (from Najjari et al. [32]). *Red lines* and *black lines* in **a** and **b** are reference lines before and after shifting the probe. The length of the *arrows* represents the distance between the lowest point of the bladder and the reference line

comes of diagnosing POP by using the PCL and H line separately. They obtained the consequence that in the anterior and central compartments, it is effective to distinguish stages 0, 1, and 2 from stages 3 and 4 (PCL—anterior: sensitivity, 92.3%, specificity 85.2%; central: sensitivity 100%, specificity 91.7%; H line—anterior: sensitivity 84.6%, specificity 92.6%; central: sensitivity 100%, specificity 97.2%). In the posterior compartment, it is comparably effectual to distinguish stages 0 and 1 from stage 2 and higher (PCL—posterior: sensitivity 100.0%, specificity 54.5%; H line—posterior: sensitivity 100.0%, specificity 63.6%). This conclusion is correct in adopting both reference lines. There is comparable stronger correlation between ultrasound and POP-Q by using the H line as the reference line. In addition, the authors found that almost all the mismatches were women negative on POP-Q but positive on TPUS in three compartments. This indicates that ultrasound is maybe over-sensitive in POP diagnosis, which would make patients unnecessarily nervous. But from another perspective, ultrasound is maybe a better screening tool, but not good for patients who have undergone therapy.

Additionally, Lone et al. [34] noticed the substantial association between different reference lines of POP-Q and 2D pelvic floor ultrasound. Thus, they tried to refine the first method mentioned above by adding the offset measured from the curved array of the probe to the reference line (H line) drawn parallel to the inferior–posterior margin of the pubic symphysis to the value of the distance from the furthest slipping point of the organ to the reference line. They chose Ba, Bp, and C on POP-Q compared with corresponding points on 2D ultrasound imaging and there is a statistically significant correlation between them. The proportion of correct diagnoses for bladder, bowel, and middle compartment prolapse were 59.6, 61.5, and 32.6% respectively which is different from previous and most studies. It is interesting that the improved correlation provided by the offset has little effect on anterior and middle compartments, but has an obvious effect (from 0.67 to 0.71) on the posterior compartment. The authors gave as an explanation for this phenomenon that the rectal angle and ampullae are closer to the H line. Moreover, the authors assumed that the pressure provided by the ultrasound transducer may induce underestimation of severe prolapse.

These three reference lines have not been compared together in one study; thus, we could only analyze advantages and disadvantages for them in theory. It is inappropriate to contrast their efficiency across studies because deficient studies were carried out. Taking the MPL may reduce bias by operators to some extent, because the outcome is unchangeable for the same person in the similar disease condition when ultrasound

physicians change shifting angles of the probe and the perineal body. According to Arian's work, the H line maybe a little better than the PCL, but the difference is not significant. Therefore, we cannot ascertain which reference line is the best; more research needs to be done.

Three reference lines on MRI are slightly different from the lines of the same names on TPUS. The H line$_{MRI}$ is a line connecting the inferior edge of the symphysis pubic and the posterior wall of the anal canal at the level of the impression of the puborectal sling. The PCL$_{MRI}$ is a straight line that passes through the inferior point of the pubic bone and the last visible coccygeal joint. The MPL$_{MRI}$ is the same line as the MPL$_{ultra}$.

Broekhuis et al. [35] contrasted the diagnostic effectiveness of several reference lines and points on MRI and TPUS. They chose the H line$_{MRI}$, PCL$_{MRI}$, and MPL$_{MRI}$ on MRI; the Ba, C, Bp points on clinical examination and MRI (points were evaluated using the MPL); the H line$_{ultra}$ on TPUS. Ninety-seven women had complete MRI and 61 women also had TPUS. The results showed that MRI had a better diagnostic compliance rate than ultrasound. In the anterior compartment, a good correlation for all three reference lines on MRI was seen (r_s, 0.61–0.66) and a moderate correlation for Ba on MRI and for the H line$_{ultra}$ on TPUS ($r_s = 0.49$ and $r_s = 0.58$). In the central compartment, the effect for all reference lines, except for PCL$_{MRI}$ ($r_s = 0.40$) and C on MRI, were poor ($r_s < 0.40$). The cervix was hardly detected on TPUS, containing the descent. In the posterior compartment, moderate correlation could be seen on MRI for MPL$_{MRI}$ and Bp ($r_s = 0.49, 0.49$) and a poor correlation for the remaining reference lines and points. With regard to the overall result, it was shown that MRI has a slightly better effect than ultrasound. For anterior and central compartment diagnoses, PCL is best on MRI, whereas for posterior compartment diagnosis, the MPL is more effective but is still not very good. Furthermore, the authors had mentioned that imaging techniques appeared to exaggerate the degree of POP and induced to over-treatment. They also thought that these imaging techniques seemed to have limited ability in providing extra value for the anterior compartment and offered poor value for the central and posterior compartments. This opinion is opposite to previous views. The limitation of this study is that many of the participants had undergone pelvic surgery.

Other parameters and POP

Bray et al. [36] recruited 243 women with symptomatic prolapse and collected the data of their vaginal wall thickness (VWT) from six sites compared with the POP-Q stage of these patients from a tertiary referral center for

urogynecology in the UK. VWT of women with POP-Q stage 3 is significantly higher than those with POP of a lower stage. Furthermore, the data pointed out that menopause status did not affect the VWT. However, this paper lacked some crucial baseline data such as body mass index (BMI) and parity number.

Wen et al. [37] also designed a clinical investigation related to the measurement of the AP vaginal canal. The rendered axial plane chosen for measuring the AP of the vaginal canal was decided to be at the level of uterine prolapse diagnosed by the first measuring method mentioned in the introduction of the TPUS part of this paper. They tried to settle the problem of how to detect concealed prolapse in the volume-rendering mode with 4D pelvic floor ultrasound. The uterine cervix is difficult to discriminate from isoechoic vaginal wall on the ultrasound image, but a hyperechoic line at the top of the vagina is the mark of the edge of the cervix. Although concealed prolapse occurs when the woman has an enterocele or rectocele that is hyperechoic gas or an acoustic shadow, which would conceal the vaginal canal and uterine cervix. Under this circumstance, it is hopeless to gauge uterine descent in the midsagittal plane. The authors set the widened vaginal AP as being above the 95th centile in normal women; then they gained the cutoff of 10 mm. Valid data from 233 women denoted that the widened AP of the vaginal canal (more than 10 mm) and an eye sign (an eye-shaped structure in the position of the central compartment with a wide AP of the vaginal canal) on ultrasound could be used to detect uterine prolapse of POP-Q stages 1 or higher (sensitivity 91.4%; specificity 88.6%). Forty-five women with concealed uterine prolapse could be diagnosed by the same parameter (sensitivity 100%; specificity 72.2%, calculated by us). Clinical diagnosis and ultrasound images had a good agreement($\kappa = 0.78$); thus, the volume-rendering mode was an efficient alternative for detecting the concealed uterine prolapse. Additionally, there was a significant difference in vaginal AP between normal women and women with POP (5.6 vs 17.8 mm). Some limitations still exist. The cutoff of uterine descent was not persuasive in that not all the participants are patients with urogynecological diseases.

Discussion

A variety of statistic indexes mentioned in 15 studies, such as median value, standard deviation, 95% confidence interval, AUC of the ROC, sensitivity and specificity for a specific cutoff processed from original data, Pearson's correlation, and Spearman's rank correlation, make it hard to decide an appropriate reference value for each parameter.

Anterior compartment descent includes cystocele and anterior enterocele. Posterior compartment descent includes rectocele and enterocele. Ultrasound is good at distinguishing among these conditions, which is so difficult on clinical examination [33]. Furthermore, ultrasound could identify levator avulsion, which is widely thought to be associated with POP [38–41]. Different levator avulsion types may lead surgeons to take variable surgery proposals. Therefore, ultrasound could provide additional information about internal herniation, rectal intussusception, and some other diseases to help to design cure and surgical plans. But some researchers noticed that TPUS had a higher misdiagnosis rate compared with clinical examination [29] and it did not change the medical management before and after surgery [31]; thus, it should not substitute clinical assessment.

More than one study obtained the result that using TPUS to diagnose POP is not efficient in the posterior compartment, but is relatively better in the anterior and central compartments [17, 21, 29, 34, 35]. A possible explanation for the terrible correlation in the posterior compartment in the studies above may be that posterior vaginal wall prolapse does not have the same meaning as rectocele, but many doctors may confuse them. It is so difficult to distinguish rectocele from enterocele by clinical examination because its reference mark is the hymen/introitus/perineum. However, the reference lines of ultrasound are related to the symphysis pubis. The use of different reference mark leads to enormous discrepancy. During Valsalva maneuver, patients' perineal descent could have a greater effect on the result of clinical examination in the posterior compartment than in the anterior and central compartments [29]. Another likely explanation is that the development of the anterior rectal wall is more ventral than caudal [35]. However, both hypotheses have not been confirmed yet.

Moreover, until now, almost no studies controlled height as a variation when collecting the baseline data. However, the reference marks of ultrasound are related to bone structure, so that pelvic morphology affects ultrasound diagnosis. Meanwhile, some papers [42–44] discover that the taller women are, the larger pelvis they have. Therefore, we reasonably speculate that height may have some potential and indirect effects on ultrasound diagnosis of POP.

In addition, we all know that pregnancy and mode of vaginal delivery affect the incidence of POP [20, 45, 46], but to our knowledge, there are no studies indicating that

pregnancy and vaginal delivery would affect the normal upper limit (women with vaginal delivery or have pregnancy but without POP) of parameters related to POP on TPUS, which would require us to set different reference values for nulliparas and multiparas. Therefore, this is maybe the reason why few studies regard these data as baseline. Maybe it makes sense to compare normal values of women who have undergone vaginal delivery and those who have not.

Some reviews focused on the relationship between prolapse symptoms and the presence of prolapse on anatomy. They reached the similar conclusion that prolapse symptoms [10, 47, 48] have not always corresponded to the outcome of POP-Q. It has been reported that pelvic floor muscle training could alleviate symptoms in women with mild prolapse [49]. Therefore, we assumed that incomplete dysfunction of the pelvic floor muscles is a likely explanation for the discordance between symptoms and anatomy change. Incorrect Valsalva maneuver is another explanation, as not all authors emphasized that they requested participants to repeat the Valsalva maneuver twice or more. The third possible reason may be that we usually perform ultrasound examination with the patient in a supine position. However, prolapse bothered women more in a standing position. This effect must not be neglected. Hence, some researchers have explored determining cutoffs of ultrasound indexes in the standing position [50, 51]. They reached the same conclusion that it certainly led to a slightly higher cutoffs, but the difference was not crucially significant. That is to say, the outcomes gathered from a supine position match those gathered from a standing position.

Besides, we noticed that all studies, except for those papers in which all the data were gathered by one ultrasound physician, emphasized that inter-agreement between the two or several operators is good. For a feasible method of diagnosis, the repeatability of outcomes is pivotal, and TPUS meets this requirement [52–54]. Obesity [55] is a critical factor that may affect the measuring baseline and weaken the penetration and resolution of ultrasound to thus reduce the accuracy of the examination. Therefore, almost every paper had recorded the BMI of patients enrolled in the studies. In some research [22], ultrasound physicians set BMI over a concrete value, depending on the criterion of obesity in their country or area, as an exclusion criterion.

We also noticed that there are still no global multicenter clinical studies to discuss the diagnostic value of TPUS for POP. There is a great deal of Anatomical variation of the pelvic floor in women from different ethnicities [56–58]. A mutual conclusion we can draw from

these papers is that black women have the largest levator hiatuses, whereas white women have smaller ones. As for Asian women, theirs are of the smallest size. South Asians have greater pelvic organ mobility than Caucasians on ultrasound and are less likely to have levator avulsions than the other two ethnic groups. In the future, it will be necessary to set different cutoff values of the same ultrasound parameter to diagnose women from around the world.

Finally, no study has been conducted to follow the possible developments of the ultrasound parameters of normal females before and after delivery over a long period. Labor and aging are considered to be two key factors that bring about POP. One weakness is that, without multiple normal parameter data to compare, it is inaccurate to provide direct cutoffs. Another weakness is the cross-sectional study design, with no possibility of following the development of anatomical prolapse and symptoms over time.

Conclusion

Pelvic floor ultrasound could provide extra information for POP, and these messages could be useful in treatment protocol design. But it is unwise to replace clinical examination with ultrasound. We must admit its limitations and inaccuracy. Among present ultrasound indexes, hiatus dimension is extremely popular with ultrasound physicians. Based on proposed reference lines, we cannot reach a definitive answer because of deficient data. It is hopeful that a fitting formula will be defined combined with some effective ultrasound indexes to diagnose POP. But before this, differences in weight, height, and ethnicity should be taken into consideration. Overall, more multi-center clinical research needs to be done.

Compliance with ethical standards

Conflicts of interest None.

Appendices

Appendix 1

Table 4   Measured levator hiatus area (HA)-related parameters to diagnose POP

Number	Reference	Final number of participants	POP-Q (number)	Setting	Age (years), mean	BMI, mean	Parity, mean	Inclusion criteria	Exclusion criteria	Index
1	Majida et al. [21]	149	0 (2)/1 (43)/2 (76)/3 (27)/4 (1)	Norway	49.1	25.3	2.3	At least 1 year beyond childbirth	Previous history of surgery for POP, breastfeeding, backache, pelvic cancer, untreated urinary tract infection, neurological disorders, psychiatric disorders	HArest/val
2	Ying et al. [22]	50 + 50	≥2 / 2 (12)/3 (20)/4 (18)	China	23 (nulli) 64 (POP)	21.92 (nulli) 23.23 (POP)	2 POP	Required to undergo prolapse repair	BMI≥28 and pelvic tumor with the diameter≥3.0 cm (such as uterine myoma, ovarian tumor and fallopian tube tumor).	APrest/val/tract HArest/val/tract LRrest/val/tract Pelvic floor axis
3	Wen and Zhou[23]	294	0 (121)/1 (49)/2 (78)/3 (42)/4(4) ≥2 (124)	China	47	23.12	2	Seen for lower urinary tract symptoms or POP	Previous history of surgery or physiotherapeutic interventions for PFD	APval HAval
4	Albrich et al. [24]	246	≥2 (121)	Germany	66	—	2.1	Intend to send to urogynecological clinic	Previous history of surgery for incontinence or POP	HAval Levator avulsion
5	Dou et al. [25]	323	≥2 A (102) A & M (24) M (68) P (24)	China	59	—	2	Seen for medical help by the urogynecological clinic	Previous history of surgery for PFD and unable to perform adequate Valsalva	HArest/val, APrest/val, LRval ΔHArest-val
6	Wen et al. [26]	102 nulliparae 64 POP	0 (102) 1 (23) ≥2 (41) ≥3 (13)	China February 2015 to December 2015	31 (nulli) 43 (POP)	20.34 (nulli) 22.03 (POP)	—	Nulliparae: no symptoms or clinical evidence of prolapse	Incomplete volume data	APval HAval
7	Wen et al. [27]	134 nulliparae 159 POP	0 (134)/1 (46)/2 (62)/3 (43)/4 (8)	China January 2016 to May 2017	32 (nulli) 48 (POP)	21.20 (nulli) 22.53 (POP)	—	Standard of nulliparae was the same as number 7. With complete clinical examination data and ultrasound imaging	Previous history of surgery or interventions for PFD	APval HAval
8	Nyhus et al. [19]	608	≥2 (275) A (172) M (29) P (154)	Norway	47.9	25.6	2.1	Delivered their first children at Trondheim University Hospital of selected duration	Without measuring the HAval	ΔAP ΔHA

Reference5: ΔAP / $HA = 100 \times$ ((measurement rest - measurement contraction)/measurement rest)

HA levator hiatus area, *AP* anteroposterior diameter, *LR* left to right diameter, *rest* at rest, *val* on Valsalva, *tract* on contraction

Appendix 2

Table 5 Measured reference line-related parameters to diagnose POP

Number	Reference	Final number of participants	POP-Q (number)	Setting	Age (years), mean	BMI, mean	Parity, mean	Inclusion criteria	Exclusion criteria	Index
9	Najjari et al. [32]	120	≥2 (no data)	Germany	58.2	–	–	–	–	MPL H line
10	Arian et al. [33]	40	>0 A (32) M (22) P (28) >2 A (13) M (4) P(6)	Iran	49.9	–	3.4	Compatible history of POP	Unable to perform Valsalva	PCL H line
11	Broekhuis et al. [35]	97	Based on value of selected points on POP-Q	Netherland	54.6	26.1	2	97 women underwent MRI for the investigation of PFD, 61 women underwent TPUS.	Over-anxious patients, poor image quality	MPL PCL H line
12	Volloyhaug et al. [29]	590	≥2 (256) A (161) M (26) P (146)	Norway	47.9	25.7	2.2	Parous women, complete datasets	Poor image quality	H line
13	Lone et al. [34]	46 (no POP) 51 (POP)	Based on value of selected points on POP-Q	UK	49.5	29.3	2	Underwent clinic examination	Clinically obvious prolapse that is extruding outside the introitus (16)	H line
14	Lone et al. [31]	53 (no POP) 105 (POP)	Based on value of selected points on POP-Q	UK	49.5	29.3	2	Experiment: had symptoms of POP and UI Control: had other gynecological symptoms	With pelvic masses	H line
15	Dietz et al. [30]	825	≥2 (646) A (477) M (221) P (464)	Germany	56	29	2	Symptoms of the lower urinary tract and pelvic floor dysfunction	No completion of clinical examination and ultrasound	H line

MPL midpubic line, *H line* horizontal line, *PCL* pubococcygeal line

Appendix 3

Table 6 Other measured parameters to diagnose POP

Number	Reference	Final participant number	POP-Q	Setting	Age (years), mean	BMI, mean	Parity, mean	Inclusion criteria	Exclusion criteria	Index
16	Bray et al. [36]	243	0 A (0) M (182) P (0) 1 A (164) M (14) P (163) 2 A (23) M (17) P (33) 3 A (55) M (22) P (45)	UK	58.2	–	–	With symptomatic prolapse	Pregnancy, unable to position themselves correctly, previous history of surgery and interventions for PFD	Vaginal wall thickness
17	Wen et al. [37]	233	0 (43)/1 (55)/2 (64)/3 (53)/4 (18) ≥2 A (88) P (62) ≥1 M (93)	China	47	22.3	2	Seen for POP	Previous history of hysterectomy, incomplete clinical data, poor volume acquisition	AP diameter of vagina

MPL midpubic line, *H line* horizontal line, *PCL* pubococcygeal line

Appendix 4: search strategy

a) Database: EBM Reviews—Cochrane Central Register of
    Controlled Trials, May 2019
       Search strategy:

----

1   exp Visceral Prolapse/ or exp Pelvic Organ Prolapse/ or exp Uterine Prolapse/ or exp Rectal Prolapse/ (448)

2   pelvic or   gan prolapse.mp. (1140)

3   cystocele.mp. or exp Cystocele/ (208)

4   rectocele.mp. or exp Rectocele/ (162)

5   exp Ultrasonography/ or ultrasonography.mp. (18971)

6   echography.mp. (5808)

7   ultrasonic imaging.mp. (26)

8   1 or 2 or 3 or 4 (1477)

9   5 or 6 or 7 (22435)

10   8 and 9 (48)

b) Database: Embase, 1974 to 13 June 2019
Search strategy:

----

1   exp uterus prolapse/ or exp apical prolapse/ or exp anterior vaginal wall prolapse/ or exp pelvic organ prolapse/ or exp vaginal vault

prolapse/ or exp rectum prolapse/ or exp pelvic floor prolapse/ or exp intestinal prolapse/ or exp posterior vaginal wall prolapse/ or exp urethral

prolapse/ (24850)

2   cystocele.mp. or exp cystocele/ (3417)

3   rectocele.mp. or exp rectocele/ (3065)

4   pelvic organ prolapse.mp. (15110)

5   ultrasonography.mp. or exp echography/ (765551)

6   ultrasonic imaging.mp. or exp ultrasound/ (171599)

7   echography.mp. (392409)

8   1 or 2 or 3 or 4 (26130)

9   5 or 6 or 7 (941877)

10   8 and 9 (2387)

11   limit 10 to (full text and female and english language and yr="2008 -Current") (859)

12   limit 10 to (full text and female and english language and yr="2008 - 2018") (821)

13   limit 10 to (full text and female and english language and yr="2019 -Current") (38)

c) Database: Ovid MEDLINE, 1946 to June week 2 2019
   Search strategy:

----

1   exp Visceral Prolapse/ or exp Rectal Prolapse/ or prolapse.mp. or exp Pelvic Organ Prolapse/ or exp Uterine Prolapse/ (27056)

2   cystocele.mp. or exp Cystocele/ (1213)

3   rectocele.mp. or exp Rectocele/ (1268)

4   echography.mp. or exp Ultrasonography/ (418644)

5   ultrasonic imaging.mp. or exp Ultrasonography/ (417034)

6   1 or 2 or 3 (27704)

7   4 or 5 (418775)

8   6 and 7 (3888)

9   limit 8 to (english language and female and full text and yr="2008 -Current") (475)

d) Database: PubMed, 1946 to 14 June 2019
   Search strategy:

Search Strategy: (289)

((((((((rectocele[Title/Abstract])   OR   rectal   prolapse[Title/Abstract])   OR   rectum   prolapse[Title/Abstract]))   OR

((cystocele[Title/Abstract])   OR   bladder   prolapse[Title/Abstract]))   OR   uterine   prolapse[Title/Abstract])   OR   pelvic   organ

prolapse[Title/Abstract]))   AND   ((ultraso*[Title/Abstract])   OR   ((ultrasonography[Title/Abstract])   OR   pelvic   floor

ultrasound[Title/Abstract]))   AND (full text[sb] AND "last 10 years"[PDat] AND Humans[Mesh])

Appendix 5

Table 7 Quality assessment

Study	Patient selection				Ultrasound diagnosis			Diagnosing golden standard			Process			Overall bias
	Random or consecutive sample	Case–control study avoided	Inappropriate exclusion avoided	Bias	Description	Did doctors perform ultrasound without knowing the results of POP-Q?	Bias	Description	Blindness of measuring	Bias	Reasonable time interval between two methods	All participants analyzed	Bias	
Majida et al. [21]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Ying et al. [22]	–	☺	☺	L	☺	–	L	☺	–	L	☺	☺	L	L
Wen et al. [26]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Albrich et al. [24]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Dou et al. [25]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☹	H	L
Wen and Zhou [23]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Wen et al. [27]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Nyhus et al. [19]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Najjari et al. [32]	☺	☺	☺	L	☺	–	L	☺	–	L	☺	☺	L	L
Arian et al. [33]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Broekhuis et al. [35]	☺	☺	☺	H	☺	☹	L		☺	L	☺	☺	L	L
Volloyhaug et al. [29]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Lone et al. [34]	☺	☺	☺	L	☺	☺	L	–	☺	L	☺	☺	L	L
Lone et al. [31] (same participants as [34])	☺	☺	☺	L	☺	☺	L	–	☺	L	☺	☺	L	L
Dietz et al. [30]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L
Bray et al. [36]	☺	☺	☺	H	☺	☹	L	☺	☹	H	☺	☺	L	L
Wen et al. [37]	☺	☺	☺	L	☺	☺	L	☺	☺	L	☺	☺	L	L

Is the risk control good or bad?

☺ : good control

– : unclear

☹ : bad control

Is the risk high or low?

L: low risk

H: high risk

References

1. Haylen BT, Maher CF, Barber MD, Camargo S, Dandolu V, Digesu A, et al. An International Urogynecological Association (IUGA)/ International Continence Society (ICS) joint report on the terminology for female pelvic organ prolapse (POP). Int Urogynecol J. 2016;27(4):655–84. https://doi.org/10.1007/s00192-016-3003-y.

2. Cooper J, Annappa M, Dracocardos D, Cooper W, Muller S, Mallen C. Prevalence of genital prolapse symptoms in primary care: a cross-sectional survey. Int Urogynecol J. 2015;26(4):505–10. https://doi.org/10.1007/s00192-014-2556-x.

3. Smith FJ, Holman CD, Moorin RE, Tsokos N. Lifetime risk of undergoing surgery for pelvic organ prolapse. Obstet Gynecol. 2010;116(5):1096–100. https://doi.org/10.1097/AOG. 0b013e3181f73729.

4. Wu JM, Vaughan CP, Goode PS, Redden DT, Burgio KL, Richter HE, et al. Prevalence and trends of symptomatic pelvic floor disorders in U.S. women. Obstet Gynecol. 2014;123(1):141–8. https:// doi.org/10.1097/aog.0000000000000057.

5. Yuk JS, Lee JH, Hur JY, Shin JH. The prevalence and treatment pattern of clinically diagnosed pelvic organ prolapse: a Korean National Health Insurance Database-based cross-sectional study 2009-2015. Sci Rep. 2018;8(1):1334. https://doi.org/10.1038/ s41598-018-19692-5.

6. Wilkins MF, Wu JM. Lifetime risk of surgery for stress urinary incontinence or pelvic organ prolapse. Minerva Ginecol. 2017;69(2):171–7. https://doi.org/10.23736/s0026-4784.16. 04011-9.

7. Wu JM, Hundley AF, Fulton RG, Myers ER. Forecasting the prevalence of pelvic floor disorders in U.S. Women: 2010 to 2050. Obstet Gynecol. 2009;114(6):1278–83. https://doi.org/10.1097/ AOG.0b013e3181c2ce96.

8. Barber MD. Symptoms and outcome measures of pelvic organ prolapse. Clin Obstet Gynecol. 2005;48(3):648–61.

9. Barber MD. Pelvic organ prolapse. BMJ. 2016;354:i3853. https:// doi.org/10.1136/bmj.i3853.

10. Manonai J, Wattanayingcharoenchai R. Relationship between pelvic floor symptoms and POP-Q measurements. Neurourol Urodyn. 2016;35(6):724–7. https://doi.org/10.1002/nau.22786.

11. Barber MD, Maher C. Epidemiology and outcome assessment of pelvic organ prolapse. Int Urogynecol J. 2013;24(11):1783–90. https://doi.org/10.1007/s00192-013-2169-9.

12. Madhu C, Swift S, Moloney-Geany S, Drake MJ. How to use the pelvic organ prolapse quantification (POP-Q) system? Neurourol Urodyn. 2018;37(S6):S39–s43. https://doi.org/10.1002/nau.23740.

13. Persu C, Chapple CR, Cauni V, Gutue S, Geavlete P. Pelvic organ prolapse quantification system (POP-Q)—a new era in pelvic prolapse staging. J Med Life. 2011;4(1):75–81.

14. Pham T, Burgart A, Kenton K, Mueller ER, Brubaker L. Current use of pelvic organ prolapse quantification by AUGS and ICS members. Female Pelvic Med Reconstr Surg. 2011;17(2):67–9. https://doi.org/10.1097/SPV.0b013e318207c904.

15. Wiegersma M, Panman CM, Kollen BJ, Berger MY, Lisman-van Leeuwen Y, Dekker JH. Is the hymen a suitable cut-off point for clinically relevant pelvic organ prolapse? Maturitas. 2017;99:86–91. https://doi.org/10.1016/j.maturitas.2017.02.012.

16. Dietz HP, Mann KP. What is clinically relevant prolapse? An attempt at defining cutoffs for the clinical assessment of pelvic organ descent. Int Urogynecol J. 2014;25(4):451–5. https://doi.org/10. 1007/s00192-013-2307-4.

17. Dietz HP, Haylen BT, Broome J. Ultrasound in the quantification of female pelvic organ prolapse. Ultrasound Obstet Gynecol. 2001;18(5):511–4. https://doi.org/10.1046/j.0960-7692.2001. 00494.x.

18. Iglesia CB, Smithling KR. Pelvic organ prolapse. Am Fam Physician. 2017;96(3):179–85.

19. Nyhus MO, Salvesen KA, Volloyhaug I. Association between pelvic floor muscle trauma and contraction in parous women from a general population—a cross sectional study. Ultrasound Obstet Gynecol. 2019;53(2):262–8. https://doi.org/10.1002/uog.19195.

20. Delancey JO, Hurd WW. Size of the urogenital hiatus in the levator ani muscles in normal women and women with pelvic organ prolapse. Obstet Gynecol. 1998;91(3):364–8.

21. Majida M, Braekken I, Bo K, Benth J, Engh M. Anterior but not posterior compartment prolapse is associated with levator hiatus area: a three- and four-dimensional transperineal ultrasound study. BJOG. 2011;118(3):329–37. https://doi.org/10.1111/j.1471-0528. 2010.02784.x.

22. Ying T, Li Q, Xu L, Liu F, Hu B. Three-dimensional ultrasound appearance of pelvic floor in nulliparous women and pelvic organ prolapse women. Int J Med Sci. 2012;9(10):894–900. https://doi. org/10.7150/ijms.4829.

23. Wen L, Zhou Q. Can we evaluate hiatal ballooning by measuring the anteroposterior diameter with 2-dimensional translabial ultrasonography? J Ultrasound Med. 2018;37(4):1001–6. https://doi.org/ 10.1002/jum.14445.

24. Albrich SB, Welker K, Wolpert B, Steetskamp J, Porta S, Hasenburg A, et al. How common is ballooning? Hiatal area on 3D transperineal ultrasound in urogynecological patients and its association with lower urinary tract symptoms. Arch Gynecol Obstet. 2017;295(1):103–9. https://doi.org/10.1007/s00404-016-4200-0.

25. Dou C, Li Q, Ying T, Yan Y, Wang X, Hu B. Determining "abnormal" levator hiatus distensibility using three-dimensional transperineal ultrasound in Chinese women. Front Med. 2018;12(5):572–9. https://doi.org/10.1007/s11684-017-0561-4.

26. Wen L, Zhang J, Zeng S, Zhou Q. Using Z-scores to evaluate levator hiatal dimensions with four-dimensional translabial ultrasound. J Obstet Gynaecol Res. 2017;43(12):1840–7. https://doi. org/10.1111/jog.13454.

27. Wen L, Liu M, Zhao B, Qing Z. Evaluating Z scores to quantify levator hiatal distensibility by 3-dimensional ultrasonography in nulliparas and women with pelvic organ prolapse. J Ultrasound Med. 2018;37(10):2363–9. https://doi.org/10.1002/jum.14590.

28. Lakeman MM, Zijta FM, Peringa J, Nederveen AJ, Stoker J, Roovers JP. Dynamic magnetic resonance imaging to quantify pelvic organ prolapse: reliability of assessment and correlation with clinical findings and pelvic floor symptoms. Int Urogynecol J. 2012;23(11):1547–54. https://doi.org/10.1007/s00192-012-1772-5.

29. Volloyhaug I, Rojas RG, Morkved S, Salvesen KA. Comparison of transperineal ultrasound with POP-Q for assessing symptoms of prolapse. Int Urogynecol J. 2019;30(4):595–602. https://doi.org/ 10.1007/s00192-018-3722-3.

30. Dietz HP, Kamisan Atan I, Salita A. Association between ICS POP-Q coordinates and translabial ultrasound findings: implications for definition of 'normal pelvic organ support'. Ultrasound Obstet Gynecol. 2016;47(3):363–8. https://doi.org/10.1002/uog.14872.

31. Lone F, Sultan AH, Stankiewicz A, Thakar R. The value of preoperative multicompartment pelvic floor ultrasonography: a 1-year prospective study. Br J Radiol. 2014;87(1040):20140145. https:// doi.org/10.1259/bjr.20140145.

32. Najjari L, Hennemann J, Larscheid P, Papathemelis T, Maass N. Perineal ultrasound as a complement to POP-Q in the assessment of cystoceles. Biomed Res Int. 2014;2014:740925. https://doi.org/10. 1155/2014/740925.

33. Arian A, Ghanbari Z, Chegini N, Hosseiny M. Agreement of ultrasound measures with POP-Q in patients with pelvic organ prolapse. Iran J Radiol. 2018;15(4):6. https://doi.org/10.5812/iranjradiol. 68461.

34. Lone FW, Thakar R, Sultan AH, Stankiewicz A. Accuracy of assessing pelvic organ prolapse quantification points using dynamic 2D transperineal ultrasound in women with pelvic organ prolapse. Int Urogynecol J. 2012;23(11):1555–60. https://doi.org/10.1007/s00192-012-1779-y.

35. Broekhuis SR, Kluivers KB, Hendriks JC, Futterer JJ, Barentsz JO, Vierhout ME. POP-Q, dynamic MR imaging, and perineal ultrasonography: do they agree in the quantification of female pelvic organ prolapse? Int Urogynecol J Pelvic Floor Dysfunct. 2009;20(5):541–9. https://doi.org/10.1007/s00192-009-0821-1.

36. Bray R, Derpapas A, Fernando R, Khullar V, Panayi DC. Does the vaginal wall become thinner as prolapse grade increases? Int Urogynecol J. 2017;28(3):397–402. https://doi.org/10.1007/s00192-016-3150-1.

37. Wen L, Li Y, Qing Z, Liu M. Detection of concealed uterine prolapse in the volume-rendering mode of 4-dimensional translabial ultrasound: a retrospective observational study. J Ultrasound Med. 2019;38(7):1705–11. https://doi.org/10.1002/jum.14855.

38. Dietz HP. Clinical consequences of levator trauma. Ultrasound Obstet Gynecol. 2012;39(4):367–71. https://doi.org/10.1002/uog.11141.

39. Dietz HP, Franco AV, Shek KL, Kirby A. Avulsion injury and levator hiatal ballooning: two independent risk factors for prolapse? An observational study. Acta Obstet Gynecol Scand. 2012;91(2):211–4. https://doi.org/10.1111/j.1600-0412.2011.01315.x.

40. Rostaminia G, Peck JD, Quiroz LH, Shobeiri SA. Characteristics associated with pelvic organ prolapse in women with significant levator ani muscle deficiency. Int Urogynecol J. 2016;27(2):261–7. https://doi.org/10.1007/s00192-015-2827-1.

41. Vergeldt TF, Notten KJ, Weemhoff M, van Kuijk SM, Mulder FE, Beets-Tan RG, et al. Levator hiatal area as a risk factor for cystocele recurrence after surgery: a prospective study. BJOG. 2015;122(8):1130–7. https://doi.org/10.1111/1471-0528.13340.

42. Holland EL, Cran GW, Elwood JH, Pinkerton JH, Thompson W. Associations between pelvic anatomy, height and year of birth of men and women in Belfast. Ann Hum Biol. 1982;9(2):113–20.

43. Munabi IG, Mirembe F, Luboga IG. Human pelvis height is associated with other pelvis measurements of obstetric value. Anat J Afr. 2015;4(1):457–65.

44. Munabi IG, Byamugisha J, Luboobi L, Luboga SA, Mirembe F. Relationship between maternal pelvis height and other anthropometric measurements in a multisite cohort of Ugandan mothers. Pan Afr Med J. 2016;24:257. https://doi.org/10.11604/pamj.2016.24.257.9889.

45. Zhu YC, Deng SH, Jiang Q, Zhang Y. Correlation between delivery mode and pelvic organ prolapse evaluated by four-dimensional pelvic floor ultrasonography. Med Sci Monit. 2018;24:7891–7. https://doi.org/10.12659/msm.911343.

46. Dietz HP, Scoti F, Subramaniam N, Friedman T, Shek KL. Impact of subsequent pregnancies on pelvic floor functional anatomy. Int Urogynecol J. 2018;29(10):1517–22.

47. Mouritsen L, Larsen JP. Symptoms, bother and POPQ in women referred with pelvic organ prolapse. Int Urogynecol J Pelvic Floor Dysfunct. 2003;14(2):122–7. https://doi.org/10.1007/s00192-002-1024-1.

48. Ellerkmann RM, Cundiff GW, Melick CF, Nihira MA, Leffler K, Bent AE. Correlation of symptoms with location and severity of pelvic organ prolapse. Am J Obstet Gynecol. 2001;185(6):1332–7; discussion 1337-1338. https://doi.org/10.1067/mob.2001.119078.

49. Panman C, Wiegersma M, Kollen BJ, Berger MY, Lisman-Van Leeuwen Y, Vermeulen KM, et al. Two-year effects and cost-effectiveness of pelvic floor muscle training in mild pelvic organ prolapse: a randomised controlled trial in primary care. BJOG. 2017;124(3):511–20. https://doi.org/10.1111/1471-0528.13992.

50. Rodriguez-Mias NL, Subramaniam N, Friedman T, Shek KL, Dietz HP. Prolapse assessment supine and standing: do we need different cutoffs for "significant prolapse"? Int Urogynecol J. 2018;29(5):685–9. https://doi.org/10.1007/s00192-017-3342-3.

51. Braverman M, Kamisan Atan I, Turel F, Friedman T, Dietz HP. Does patient posture affect the ultrasound evaluation of pelvic organ prolapse? J Ultrasound Med. 2019;38(1):233–8. https://doi.org/10.1002/jum.14688.

52. Dietz HP, Rojas RG, Shek KL. Postprocessing of pelvic floor ultrasound data: how repeatable is it? Aust N Z J Obstet Gynaecol. 2014;54(6):553–7. https://doi.org/10.1111/ajo.12250.

53. Majida M, Braekken IH, Umek W, Bo K, Saltyte Benth J, Ellstrom Engh M. Interobserver repeatability of three- and four-dimensional transperineal ultrasound assessment of pelvic floor muscle anatomy and function. Ultrasound Obstet Gynecol. 2009;33(5):567–73. https://doi.org/10.1002/uog.6351.

54. Tan L, Shek KL, Atan IK, Rojas RG, Dietz HP. The repeatability of sonographic measures of functional pelvic floor anatomy. Int Urogynecol J. 2015;26(11):1667–72. https://doi.org/10.1007/s00192-015-2759-9.

55. Giri A, Hartmann KE, Hellwege JN, Velez Edwards DR, Edwards TL. Obesity and pelvic organ prolapse: a systematic review and meta-analysis of observational studies. Am J Obstet Gynecol. 2017;217(1):11–26.e13. https://doi.org/10.1016/j.ajog.2017.01.039.

56. Abdool Z, Dietz HP, Lindeque BG. Ethnic differences in the levator hiatus and pelvic organ descent: a prospective observational study. Ultrasound Obstet Gynecol. 2017;50(2):242–6. https://doi.org/10.1002/uog.17297.

57. Abdool Z, Dietz HP, Lindeque BG. Interethnic variation in pelvic floor morphology in women with symptomatic pelvic organ prolapse. Int Urogynecol J. 2018;29(5):745–50. https://doi.org/10.1007/s00192-017-3391-7.

58. Derpapas A, Ahmed S, Vijaya G, Digesu GA, Regan L, Fernando R, et al. Racial differences in female urethral morphology and levator hiatal dimensions: an ultrasound study. Neurourol Urodyn. 2012;31(4):502–7. https://doi.org/10.1002/nau.21181.

59. Onwude JL. Genital prolapse in women. BMJ Clin Evid. 2012;3:817