Original Article

Postprocessing of pelvic floor ultrasound data: How repeatable is it?

Hans P. DIETZ,[1] Rodrigo Guzman ROJAS[2] and Ka Lai SHEK[3]

[1]*Sydney Medical School Nepean, Nepean Hospital, Penrith, NSW, Australia,* [2]*Departamento de Ginecología y Obstetricia, Clínica Alemana de Santiago-Hospital Padre Hurtado, Universidad del Desarrollo, Santiago, Chile and* [3]*Liverpool Hospital, University of Western Sydney, Sydney, NSW, Australia*

Aims: Translabial 3D/4D pelvic floor ultrasound (PFUS) is increasingly used in the evaluation of pelvic floor disorders. Commonly, this involves the analysis of stored volume data sets by postprocessing. In this study, we aimed to assess the time requirement to reaching acceptable repeatability for commonly employed outcome measures in PFUS.

Methods: Between 2010 and 2013, 20 individuals from 11 countries underwent training in postprocessing of PFUS volume data sets. They undertook test–retest series ($n \geq 20$) between day 2 and day 15 of training. Outcome measures tested included levator hiatal area on Valsalva, descent of the bladder neck, bladder, uterus and rectal ampulla, and rectocele depth. After an initial training session of 10–20 cases, test–retest series were undertaken between the trainee and measurements obtained by the author or senior trainees.

Results: Trainees were obstetricians/gynaecologists in training ($n = 4$), obstetricians/gynaecologists or subspecialty trainees ($n = 13$), medical students ($n = 1$) and physiotherapists ($n = 2$). A total of 58 repeatability series were analysed, obtained between days 2 and 15 of training. When second or third retest series were necessary, there always was improvement in repeatability except for one series in one individual. Satisfactory repeatability (ICC > 0.7) was achieved by all trainees for all parameters required by them. Training lasted from 3 to 15 days, with means between 4 and 5.8 days.

Conclusions: Postprocessing analysis of commonly used PFUS parameters can be taught to an acceptable standard within 1 week. Most commonly used ultrasound parameters obtained by postprocessing for prolapse assessment can be taught to an acceptable standard of repeatability within one week.

Key words: cystocele, female pelvic organ prolapse, rectocele, repeatability, ultrasound, uterine prolapse.

Introduction

Translabial 3D/4D ultrasound is increasingly used in the investigation of pelvic floor disorders, especially in women suffering from female pelvic organ prolapse, faecal and urinary incontinence. The method allows simple, noninvasive evaluation of pelvic organ mobility and pelvic floor functional anatomy and is superior to other imaging methods due to high spatial and temporal resolutions.[1] A core advantage of 4D ultrasound over plain X-ray, computed tomography and magnetic resonance imaging is that the sonographic method allows the acquisition of sequences of volume data blocks which are easily archived and retrieved for analysis at a later date. This process was first described in 2004.[2] As there is no 3D or 4D DICOM (Digital Imaging and Communications in Medicine) standard yet, proprietary software is used to analyse data sets. Use of such software has greatly enhanced training and research capabilities in this field, which is evident from the published literature.

The most commonly used systems are those of the Voluson series (GE Kretz Ultrasound, Zipf, Austria), and the resulting data sets are analysed using versions of the software '4D View' (GE Kretz Ultrasound, Zipf, Austria). The analysis of stored data sets allows retrospective studies on large populations, requiring minimal resources.[3] However, this requires that staff be trained in the use of postprocessing software and in pattern recognition. While all published repeatability series using pelvic floor ultrasound data acquired with Voluson-type systems have demonstrated acceptable repeatability for all the standard measures of functional pelvic floor anatomy, information on the time required for training is lacking. One recent publication addresses the length of the learning process for multiple measures of hiatal functional anatomy in one individual,[4] showing that the measurement of all assessed

Correspondence: Hans P. Dietz, Sydney Medical School Nepean, Nepean Hospital, Penrith, NSW 2750, Australia. Email: hpdietz@bigpond.com

Conflict of Interests: HP Dietz has received unrestricted educational grants from GE Medical. KL Shek has received unrestricted educational grants from GE Medical.

Received 27 April 2014; accepted 23 July 2014.

© 2014 The Royal Australian and New Zealand College of Obstetricians and Gynaecologists

hiatal dimensions could be taught to an acceptable standard within 23 h of total training, confirming several other studies demonstrating good repeatability of levator hiatal dimensions.[4–9]

In this study, we intended to assess the time requirement to reaching acceptable repeatability for the most commonly employed outcome measures in pelvic floor ultrasound in 20 individuals from a variety of professional backgrounds and nationalities.

Materials and Methods

Between 2010 and 2012, 20 individuals from 11 countries underwent training in postprocessing of translabial ultrasound volume data sets at our tertiary urogynaecological unit for the purpose of prolapse assessment. None of the trainees had previously used postprocessing software. There was a varying degree of previous experience with diagnostic imaging. Trainees undertook test–retest series ($n \geq 20$) between day 2 and day 15 of training, depending on progress. Outcome measures tested included levator hiatal area, descent of the bladder neck, bladder, uterus and rectal ampulla, and rectocele depth on Valsalva. The tested measures depended on the parameters required for the trainee's selected field of study.

Ultrasound volume data sets used for retest series had been acquired supine and after voiding, using a Voluson 730 expert system with RAB 8-4 MHz transducer, as previously described.[2] The resulting 4D ultrasound data sets were subsequently investigated with the help of postprocessing software 4D View v 10.0 (GE Kretz Ultrasound, Zipf, Austria). Volume data sets were analysed blinded against all clinical data.

Bladder neck and pelvic organ descent was determined by measuring the position of the internal urethral meatus, the most dependent part of the bladder, the cervix and small bowel/rectal ampulla relative to the inferior margin of the symphysis pubis (see Fig. 1) as previously described.[10] Hiatal area was measured in the plane of minimal hiatal dimensions as previously described,[5] see Figure 2. After an initial training session of 10–20 cases assessed under direct supervision of one of the authors, test–retest series were undertaken between the trainee and measurements obtained by the first author or senior trainees, that is by individuals with at least one-year experience in translabial ultrasound. Repeatability of measurements was tested with intraclass correlations (ICC, single measurement, absolute agreement definition), on series of 20 women assessed on multiple days of training until satisfactory agreement was achieved. We used SPSS (SPSS 16; SPSS, Chicago, IL, USA).

Training occurred for a maximum of 8 hours per calendar day and involved analysis of stored volume data sets, some under direct supervision, and live scanning under direct supervision of the authors. The characteristics of women whose data were used for retest series varied widely as their volumes were taken from the respective study populations. All data were obtained in the context of IRB-approved individual research projects (SWAHS HREC (Nepean Hospital) 05-004, 05-029, 07-022, 07-063, NBMLHD HREC (Nepean Hospital) 09-03, 09-38, 09-42, 10-03, 10-05, 11-04, 11-12, 11-13, 11-35, 11-55, 12-15, 12-45, 12-71, Greenslopes Private Hospital HREC 10-09, Townsville HREC 84/04). In this retrospective study, we reviewed and summarised the teaching experience of those 20 individuals in order to demonstrate average time requirements for the teaching of offline analysis of commonly used parameters of pelvic floor functional anatomy.

Results

Trainees were obstetricians/gynaecologists in training ($n = 4$), obstetricians/gynaecologists or subspecialty trainees ($n = 13$), medical students ($n = 1$) and physiotherapists ($n = 2$). For 12 out of 20 trainees, the language of instruction (English) was their second or third language. They originated from the United Kingdom (5), the USA (4), Australia (3), South Africa (1), the Czech

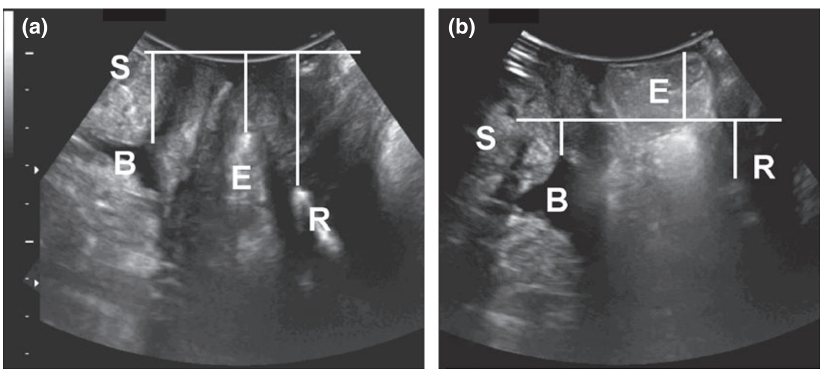

Figure 1 Prolapse assessment by translabial ultrasound, as seen in the midsagittal line. Image (a) shows appearances at rest, image (b) on maximal Valsalva. There is bladder descent to 1 cm above the symphyseal reference line and an enterocele to 2 cm below this line. B, bladder; E, enterocele; R, rectal ampulla; S, symphysis pubis.

 © 2014 The Royal Australian and New Zealand College of Obstetricians and Gynaecologists

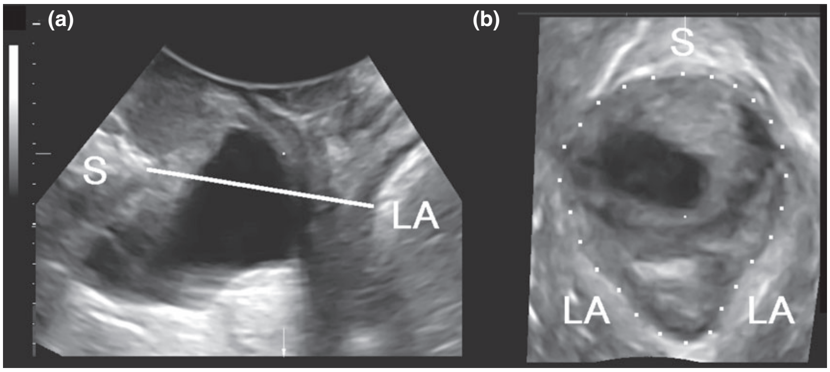

Figure 2 Determination of hiatal area in single-plane measurements. Panel a shows the mid-sagittal plane, with the oblique white line demonstrating the location of the plane of minimal dimensions, which is shown in Panel b. The hiatal area is indicated by a dotted line. LA, levator ani; S, symphysis pubis.

Republic (1), Canada (1), Germany (1), Chile (1), Sweden (1), Austria (1) and China (1). Depending on the requirements of their individual research projects, they measured levator hiatal area on Valsalva (n = 17), bladder neck descent (n = 7), cystocele descent (n = 8), uterine descent (n = 3), rectal descent (n = 6) and depth of a rectocele (n = 4). A total of 58 repeatability series were analysed and obtained between days 2 and 15 of training. When second or third retest series were necessary because of suboptimal initial results, there always was improvement in repeatability except for one series in one individual. The table shows results for the six tested parameters, giving data on 45 final series obtained by 20 individuals.

The commonest parameter analysed was hiatal area on Valsalva, which was assessed by 17 individuals. After an average of 5.8 days in training, an average intraclass correlation of 0.86 (range 0.70–0.99) was reached. Bladder neck descent, cystocele descent and descent of the rectal ampulla were measured by between 6 and 8 individuals, and all reached acceptable repeatability of an ICC of 0.65 or better within 10 days, with a mean ICC of between 0.76 and 0.89 after an average of 4.1–4.7 days. The lowest ICC values were obtained for rectocele depth and uterine descent which generally seem to be slightly more challenging for trainees (see Table 1 for time requirement of training and final ICC reached, including range and standard deviation).

Discussion

The advent of modern 3D/4D ultrasound systems has led to much wider use of imaging in urogynaecology over the last 10 years, both in research and in clinical practice. Since its first description in 2004,[2] 4D translabial ultrasound has become of great utility in pelvic floor medicine, especially for the investigation of urinary incontinence,[11,12] female pelvic organ prolapse[1,13] anal incontinence[14] and obstructed defecation.[15] The literature in this field continues to grow fast, and many of the studies undertaken in this field rely on the offline analysis of stored data sets with the help of proprietary software such as 4D View (GE Kretz), Qlab (Royal Philips, Amsterdam, the Netherlands) or BK 3D View (Bruel and Kjaer, Naerum, Denmark). Hence, the repeatability of measures is influenced not just by the quality of data acquisition at the time of the examination itself, but also by the quality of postprocessing. This may be seen as a potential disadvantage, but it is amply balanced by the ability to assess multiple aspects of functional anatomy at a later time. For example, a volume data set acquired to assess bladder and urethral mobility in a woman with

Table 1 Days in training and Intraclass correlation for measurements used in the assessment of pelvic organ support. Suboptimal series (ie those requiring additional training) are omitted

	Days in training		Intraclass correlation		
Parameter	Mean	Range	Mean	Range	SD
Hiatal area on Valsalva (n = 17)	5.8	3–15	0.86	0.70–0.99	0.09
Bladder neck descent (n = 7)	4.7	4–6	0.81	0.71–0.94	0.12
Cystocele descent (n = 8)	4.1	3–6	0.89	0.78–0.99	0.07
Uterine descent (n = 3)	4.0	3–5	0.74	0.51–0.97	0.23
Rectal descent (n = 6)	4.7	3–10	0.76	0.66–0.89	0.08
Rectocele depth (n = 4)	5.0	3–10	0.75	0.65–0.80	0.07

© 2014 The Royal Australian and New Zealand College of Obstetricians and Gynaecologists

stress urinary incontinence may also serve to investigate symptoms of obstructed defecation or faecal incontinence, potentially avoiding another diagnostic procedure. However, it is recognised that the established diagnostic approach to those two conditions involves defecation proctography and endo-anal ultrasound, and translabial 4D imaging is by no means generally accepted as equivalent.

Regardless of immediate clinical utility, postprocessing of ultrasound volume data has greatly enhanced teaching and research capabilities of urogynaecological units, especially as regards the assessment of prolapse and pelvic floor functional anatomy, and it has the capacity to simplify quality assurance and audit projects. From the available literature, it seems evident that most of the described parameters can be obtained with a high degree of repeatability.

This current study adds to the literature by showing that the most commonly used pelvic floor ultrasound parameters obtained by postprocessing for the assessment of pelvic organ support can reliably be taught to an acceptable standard of repeatability within 1 week. However, it is also evident that there is a spectrum of individual competence and that certain measures are harder to learn – such as those relating to the assessment of the central and posterior compartment.

Several weaknesses of this study need to be mentioned. The study design was retrospective and opportunistic, resulting in widely varying numbers of retest series for different measures, from $n = 17$ for hiatal dimensions on Valsalva to only 3 for uterine descent. In addition, the professional and linguistic background of trainees varied widely. However, we would contend that this makes our data more widely applicable. Another weakness is the absence of qualitative diagnoses commonly applied in functional pelvic floor imaging, such as the diagnosis of a 'true rectocele',[15,16] a levator avulsion[17,18] or a (residual) defect of the anal sphincter.[14,19] These parameters may require further investigation in future. In addition, other numeric anatomical measures such as the levator–urethra gap,[20] detrusor wall thickness,[21] urethral rotation or the retrovesical angle were not assessed either, leaving room for further work on the teaching aspects of pelvic floor ultrasound. Finally, it has to be conceded that postprocessing of data sets is only one aspect of the repeatability or reliability of such measures. Surely, ultrasound volume data acquisition itself also adds variability, especially when it involves the performance of manoeuvres such as a Valsalva or a pelvic floor contraction, most obviously for posterior compartment assessment. The latter is largely due to different degrees of stool filling and varying stool quality, and it is understood that a Valsalva manoeuvre does not mimic the actual process of defecation. Siafarikas *et al.* recently investigated the consistency and learning curve of volume data acquisition itself and found that training was easily accomplished within a short time frame of a few days, reducing the likelihood that this is a major factor.[4] In a recent study in our unit, we investigated the repeatability of multiple standard measures by comparing two independently performed tests in 106 patients at a mean interval of 73 days. Both volume data acquisition and offline analysis were performed blinded to all other imaging data. Comparison of six numerical measures of organ descent and hiatal dimensions yielded ICC of between 0.44 for descent of the rectal ampulla and 0.93 for hiatal area on Valsalva, confirming the data presented in this paper, and kappa between 0.91 for levator avulsion and 0.73 for true rectocele.[22]

In conclusion, offline analysis of ultrasound volume data sets acquired by translabial ultrasound can reliably be taught to an acceptable standard of repeatability within 1 week in most individuals.

References

1 Dietz HP. Pelvic Floor ultrasound in prolapse: what's in it for the surgeon? *Int Urogynecol J* 2011; **22**: 1221–1232.

2 Dietz H. Ultrasound imaging of the pelvic floor: 3D aspects. *Ultrasound Obstet Gynecol* 2004; **23**: 615–625.

3 Dietz H. Pelvic floor ultrasound: a review. *Am J Obstet Gynecol* 2010; **202**: 321–334.

4 Siafarikas F, Staer-Jensen J, Braekken I *et al.* Learning process for performing and analyzing 3D/4D transperineal ultrasound imaging and interobserver reliability study. *Ultrasound Obstet Gynecol* 2013; **41**: 312–317.

5 Dietz H, Shek K, Clarke B. Biometry of the pubovisceral muscle and levator hiatus by three-dimensional pelvic floor ultrasound. *Ultrasound Obstet Gynecol* 2005; **25**: 580–585.

6 Van Veelen G, Schweitzer K, Van Der Vaart C. Reliability of pelvic floor measurements on three- and four-dimensional ultrasound during and after first pregnancy: implications for training. *Ultrasound Obstet Gynecol* 2013; **42**: 590–595.

7 Yang J, Yang S, Huang W. Biometry of the pubovisceral muscle and levator hiatus in nulliparous Chinese women. *Ultrasound Obstet Gynecol* 2006; **26**: 710–716.

8 Chan S, Cheung R, Lee L *et al.* Pelvic floor biometry in Chinese primiparous women 1 year after delivery: a prospective observational study. *Ultrasound Obstet Gynecol* 2014 Apr; **43**(4): 466–474.

9 Siafarikas F, Staer-Jensen J, Braekken IH *et al.* Learning process for performing and analyzing 3D/4D transperineal ultrasound imaging and interobserver reliability study. *Ultrasound Obstet Gynecol* 2013 Mar; **41**(3): 312–317.

10 Dietz HP, Haylen BT, Broome J. Ultrasound in the quantification of female pelvic organ prolapse. *Ultrasound Obstet Gynecol* 2001; **18**: 511–514.

11 Dietz HP. Pelvic Floor ultrasound in incontinence: what's in it for the surgeon? *Int Urogynecol J* 2011; **22**: 1085–1097.

12 Pirpiris A, Shek K, Kay P, Dietz H. Urethral mobility and urinary incontinence. *Ultrasound Obstet Gynecol* 2010; **36**: 507–511.

13 Steensma AB, Oom DMJ, Burger C, Schouten W. Assessment of posterior compartment prolapse: a comparison of evacuation proctography and 3D transperineal ultrasound. *Colorectal Dis* 2010; **12**: 533–539.

 © 2014 The Royal Australian and New Zealand College of Obstetricians and Gynaecologists

14 Oom DMJ, West CR, Schouten WR, Steensma AB. Detection of anal sphincter defects in female patients with fecal incontinence: a comparison of 3-dimensional transperineal ultrasound and 2-dimensional endoanal ultrasound. *Dis Colon Rectum* 2012; **55**: 646–652.

15 Dietz H. Female pelvic floor dysfunction—an imaging perspective. *Nat Rev Gastroenterol Hepatol* 2012; **9**: 113–121.

16 Beer-Gabel M, Teshler M, Barzilai N *et al.* Dynamic transperineal ultrasound in the diagnosis of pelvic floor disorders: pilot study. *Dis Colon Rectum* 2002; **45**: 239–248.

17 Dietz H, Bernardo M, Kirby A, Shek K. Minimal criteria for the diagnosis of avulsion of the puborectalis muscle by tomographic ultrasound. *Int Urogynecol J* 2011; **22**: 699–704.

18 Notten K, Weemhoff M, Schweitzer K *et al.* Diagnosing levator defects on 3D transperineal ultrasound compared to MR imaging. *Ultrasound Obstet Gynecol* 2012; **40** (Suppl 1): 19–20.

19 Guzman Rojas R, Shek K, Langer S, Dietz H. Prevalence of anal sphincter injury in primiparous women. *Ultrasound Obstet Gynecol* 2013; **42**: 461–466.

20 Dietz H, Abbu A, Shek K. The levator urethral gap measurement: a more objective means of determining levator avulsion? *Ultrasound Obstet Gynecol* 2008; **32**: 941–945.

21 Khullar V, Cardozo LD, Salvatore S, Hill S. Ultrasound: a noninvasive screening test for detrusor instability. *Br J Obstet Gynaecol* 1996; **103**: 904–908.

22 Tan L, Kamisan Atan I, Guzman Rojas R, Dietz HP. The repeatability of sonographic measures of functional pelvic floor anatomy. *Neurourol Urodyn* 2014; **33**: in print.

© 2014 The Royal Australian and New Zealand College of Obstetricians and Gynaecologists