Evaluation of ultrasound point shear wave elastography reliability in an elasticity phantom

Salahaden R. Sultan1, Amer Alghamdi2, Rawan Abdeen1, Fahad Almutairi1

1Department of Diagnostic Radiology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah; 2Department of Diagnostic Radiology, King Fahad Hospital, Al-Baha, Saudi Arabia

Purpose: To date, limited studies have specifically addressed the reliability of ultrasound point shear-wave elastography (pSWE). Therefore, the aim of the present study was to assess the reproducibility of ultrasound pSWE within and between operators using two ultrasound scanners.

Methods: iU22 and EPIQ7 ultrasound scanners were used to assess the reliability of pSWE measurements of four inclusions [L I (8 kPa), L II (14 kPa), L III (48 kPa), and L IV (80 kPa)] at a depth of 3.5 cm in an elasticity phantom using a curvilinear 5–1 MHz transducer. The intra-operator, inter-operator, and inter-scanner reproducibility of pSWE was assessed using intraclass correlation coefficients (ICCs). Bland-Altman plots were used to establish bias and limits of agreement (LoA) between measurements. The accuracy of pSWE from manufacturer values was determined using the one-sample t-test.

Results: Intra-operator agreement was excellent, with an ICC >0.90. The bias in measurements for operator A was −0.36±3.13 kPa (LoA, −6.47 to 5.75), and for operator B it was 1.97±6.29 kPa (LoA, −10.25 to 14.21). Inter-operator agreement was excellent, with an ICC of 0.95. The bias in measurements between operators was −0.42±5.00 kPa (LoA, −10.24 to 9.38). The inter-scanner agreement between EPIQ7 and iU22 was excellent, with an ICC of 0.96. The bias in measurements between scanners was 1.74±4.44 kPa (LoA, −6.95 to 10.45). There was significant overestimation for L I (17.75%) and L II (31.14%) and underestimation for L III (−15.28%) and L VI (−98.00%) relative to the manufacturer-reported values.

Conclusion: Phantom ultrasound pSWE was reproducible within and between operators, and between Philips ultrasound scanners; further studies using different ultrasound systems and transducers are required.

Keywords: Ultrasound; Point shear-wave elastography; pSWE; Elasticity phantom

Key points: This study assessed the reliability of ultrasound pSWE in an elasticity phantom. Our findings suggest that ultrasound pSWE is a reliable and reproducible method for quantitative assessment of tissue stiffness.

Introduction

Ultrasound elastography has been developed to provide quantitative information on tissue elasticity.
Ultrasound elastography techniques can be broadly categorised into two types: strain elastography (SE) and shear wave elastography (SWE). SE characterizes tissue strain in response to transducer compression, which induces a tissue displacement resulting in strain that is quantified as a percentage and displayed as an elastogram image; soft tissue tends to have high strain values (shown in blue), whereas stiff tissue tends to have low strain values (shown in red) [1]. In contrast, the SWE acoustic radiation force impulse technique is applied based on the emission of acoustic pulses into the sampled area, inducing tissue displacement and resulting in transversely oriented mechanical shear waves within the investigated tissue from the region of excitation [2,3]. There are two approaches to SWE: point SWE (pSWE), which provides a single elasticity measurement in a selected region, and two-dimensional SWE (2D-SWE), which provides a colour elastographic map of tissue superimposed over the ultrasound B-mode image in real-time; this feature provides the operator with an overview of the tissue stiffness throughout the area of interest, enabling a qualitative and quantitative evaluation of elasticity in selected regions of interest on the elastographic map [4]. Quantitative information from ultrasound SWE has been reported to be a valid and useful diagnostic tool for various diseases, including breast cancer, thyroid lesions, and liver fibrosis [5–7]. Both pSWE and 2D-SWE measure the velocity of the shear wave (C). The shear modulus (G) is G=pC^2, where p denotes the density of tissue (≈1,000 kg/m^3). An approximate formula is used to calculate the Young modulus (E) as E=2G(1+γ)=3G, where γ is the Poisson ratio (≈0.5 in tissue) [8].

The recent expansion and development of ultrasound SWE and its known advantages of being non-invasive and inexpensive and providing real-time assessment of tissue and organ stiffness make it a preferable method for investigating various clinical conditions [9]. However, ultrasound SWE measurements may show variability unrelated to physiological changes of organ structure. This could be related to variations in imaging techniques such as depth of region of interest, the presence of artifacts, the use of different ultrasound imaging systems and operator-related errors [9–11].

The reliability of SWE measurements and the potential impact of operators and ultrasound scanners is of concern, as these factors may result in inconsistent pSWE measurements. Therefore, it is important to assess the reproducibility of pSWE measurements and to ensure low variability of measurements within and between operators and scanners pre-clinically. To date, there is a limited number of studies that have specifically addressed the reliability of ultrasound pSWE. Therefore, the aim of the present study was to assess the accuracy and variability of ultrasound pSWE measurements within and between operators using two ultrasound scanners.

**Materials and Methods**

**Study Design**

Two different generations of Philips Healthcare ultrasound imaging systems iU22 and EPIQ7 (Philips Healthcare, Bothell, WA, USA) were used to assess intra-operator, inter-operator, and inter-scanner variability and the accuracy of pSWE measurements in terms of the Young modulus (kPa) on an elasticity quality assurance phantom using a curvilinear 5–1 MHz transducer. The reason for assessing pSWE variability between these two generations of scanners was to determine whether consistent elasticity measurements can be provided. Two experienced certified clinical sonographers with efficient training on ultrasound SWE (operator A and operator B) were asked to take pSWE measurements at different visits. Both operators were blinded to each other’s measurements and to the reference stiffness values of the phantom lesions. This study does not include animals or human subjects; therefore, ethical approval was not required.

**Phantom**

pSWE measurements were performed on an elasticity quality assurance phantom Model 049, produced by Computerized Imaging Reference Systems, Inc. (CIRS, Norfolk, VA, USA) [12]. The phantom is made of eight spherical lesions: four superficial and four deep (Fig. 1). The superficial lesions are at a depth of 1.5 cm, with a volume ranging from 4.1 to 4.3 mL. The deep lesions are at a depth of 3.5 cm, with volume ranging from 4.1 to 4.3 mL. The phantom is made from Zerdine, a poly-acrylamide polymer with acoustic properties comparable to those of human tissue [13]. All the lesions have a known stiffness in terms of the Young modulus (Table 1).

**Data Acquisition**

Each operator was asked to take pSWE measurements at two
Reliability of ultrasound pSWE

different visits. On the first and second visits, 10 pSWE measurements were obtained from each of the four deep lesions in kPa using a curvilinear 5–1 MHz transducer with the EPIQ7 ultrasound imaging system [13,14]. Due to the poor image resolution at a depth of 1.5 cm with use of the curvilinear 5–1 MHz transducer, the superficial lesions within the phantom were not assessed in this study. The data were acquired starting from L-I to L-VI in longitudinal scanning with the transducer lifted and repositioned between acquisitions by applying minimal pressure on the phantom surface [15]. An additional third visit by operator A was made, in which the same protocol followed in the previous visits was repeated using a curvilinear 5–1 MHz transducer with the iU22 ultrasound imaging system. The ultrasound system controls were optimised before the elasticity measurements. The pSWE values of all four lesions were obtained using medium penetration mode with gain of 57% in EPIQ7 and 66% in iU22, and the image depth was set at 9 cm. The pSWE region of interest with a diameter of 1 cm was placed in the centre of the lesion at a depth of 3–4 cm (Fig. 2).

Statistical Analysis

pSWE measurement reproducibility was assessed using the intraclass correlation coefficient (ICC): ICC <0.50, poor agreement; ICC 0.50–0.75, moderate agreement; ICC 0.76–0.90, good agreement; ICC >0.90, excellent agreement [16]. A Bland-Altman plot was used to establish bias and limits of agreement (LoA) between pSWE measurements [17]. The coefficient of determination (R²; range, 0% to 100%) was used to determine the percentage of variance in pSWE measurements around the fitted regression line: R²<50%, poor variance in measurements; R² 50%–70%, moderate variance in measurements; R²>70%, excellent variance in measurements [18]. The mean of the coefficient of variation (CV) from all lesions was used to determine intra-operator measurement variability in each visit: <15%, low variability; 16%–25%, moderate variability; >25%, high variability [19]. The accuracy of the elasticity measurements relative to the manufacturer-reported values was determined using the one-sample t-test. The level of significance was set at P<0.05. The statistical analysis was performed using IBM SPSS Statistics version 21 (IBM Corp., Armonk, NY, USA) and GraphPad Prism 7 (GraphPad, La Jolla, CA, USA).

Results

In total, 240 pSWE images with elasticity measurements in kPa were

Table 1. Manufacturer’s elasticity measurements of deep lesions within the phantom

| Lesion type | Young modulus (kPa, ±5% SD) |
|-------------|-----------------------------|
| L I         | 8±3                         |
| L II        | 14±4                        |
| L III       | 45±5                        |
| L IV        | 80±8                        |

SD, standard deviation.

Fig. 2. Point shear-wave elastography (pSWE) of a lesion in an elasticity quality assurance phantom Model 049. A. pSWE measurements were obtained in kPa using EPIQ7 with a gain of 57%. B. pSWE measurements were obtained in kPa using iU22 with a gain of 66%. In both systems, the image depth was set at 9 cm and the pSWE region of interest with a diameter of 1 cm was placed in the centre of the lesion at a depth of 3–4 cm.
The inter-operator agreement of pSWE measurement using EPIQ7 was excellent, with an ICC of 0.98 (95% CI 0.96–0.99) and \( R^2 = 0.93 \) obtained. The two operators acquired 40 pSWE measurements from the four deep lesions (10 pSWE measurements per lesion) on each visit.

**Intra-operator Agreement**

The intra-operator agreement of pSWE measurement at two different visits (pSWE n=40 for each operator/visit) was excellent for operators A and B using EPIQ7, with ICC values of 0.98 (95% confidence interval [CI], 0.96 to 0.99; \( P<0.001 \)) (Fig. 3A) and 0.92 (95% CI, 0.85 to 0.96; \( P<0.001 \)) (Fig. 3B), respectively. The bias in measurements for operator A was \(-0.36\pm3.13 \) kPa (LoA, \(-6.47 \) to 5.75) (Fig. 3C), and that for operator B was \(1.97\pm6.29 \) kPa (LoA, \(-10.25 \) to 14.21) (Fig. 3D). For both operators, the percentage of variance in pSWE measurements was excellent, with \( R^2 \) equalling 93% and 80% for the measurements made by operators A and B, respectively, in terms of distribution around the fitted regression line (Fig. 3A–D). The mean CV for operator A was low (14% and 12% variability at visits 1 and 2, respectively). For operator B, the mean CV percentage was low at visit 1 and moderate at visit 2, with variability of 13% and 16%, respectively, between the pSWE measurements.

**Inter-operator Agreement**

The inter-operator agreement of pSWE measurement using EPIQ7 (pSWE n=80 for each operator) was excellent, with an ICC of...
Reliability of ultrasound pSWE

0.95 (95% CI, 0.93 to 0.97; P<0.001) (Fig. 4A). The bias in measurements was -0.42±5.00 kPa (LoA, −10.24 to 9.38) (Fig. 4B). The percentage of variance in the pSWE measurements was excellent, with $R^2=83\%$ in terms of the distribution of measurements around the fitted regression line (Fig. 4A, B).

Inter-scanner Agreement

The inter-scanner agreement of pSWE measurements between the two ultrasound imaging systems (EPIQ7 and iU22, n=80 pSWE measurements from each machine by operator A) was excellent, with an ICC of 0.96 (95% CI, 0.92 to 0.97; P<0.001) (Fig. 5A). The bias in measurements was 1.74±4.44 kPa (LoA, −6.95 to 10.45) (Fig. 5B). The percentage of variance in the pSWE measurements was excellent, with $R^2=87\%$ in terms of the distribution of measurements around the fitted regression line (Fig. 5A, B).

Accuracy of pSWE Measurements

The pSWE measurements of all lesions (pSWE n=60 per lesion)

---

**Fig. 4.** Inter-operator and Bland-Altman agreement of point shear-wave elastography (pSWE) measurements between operators A and B using EPIQ7.

A. Intraclass correlation coefficient (ICC) and variance of pSWE between operators A and B (ICC, 0.95 and $R^2=0.83$) were obtained. B. Bland-Altman assessment of pSWE between operators A and B (bias, -0.42 kPa; limits of agreement [LoA], −10.24 to 9.38) was obtained.

**Fig. 5.** Inter-scanner and Bland-Altman agreement of point shear-wave elastography (pSWE) measured by operator A using EPIQ7 and iU22.

A. Intraclass correlation coefficient (ICC) and variance of pSWE between EPIQ7 and iU22 (ICC, 0.96 and $R^2=0.87$) were obtained. B. Bland-Altman assessment of pSWE between EPIQ7 and iU22 (bias, 1.74 kPa; limits of agreement [LoA], −6.95 to 10.45) was obtained.
were significantly different from the manufacturer-reported values. There was an overestimation of mean pSWE values for L I (17.75%, P<0.001) and L II (31.14%, P<0.001) and an underestimation of mean pSWE values for L III (−15.28%, P<0.001) and L VI (−98.00%, P<0.001).

Discussion

The recent expansion and development of ultrasound elastography underscores the need for quantitative studies to address the variability of ultrasound pSWE measurements. To achieve this goal, the present study investigated the reproducibility of ultrasound pSWE measurements within and between operators using two different generations of Philips Health Care ultrasound imaging devices (i.e., EPIQ7 and iU22) with an elasticity phantom. Intra-operator and inter-operator agreement were excellent, with low to moderate variability. Similarly, excellent agreement in the pSWE measurements was found between the two ultrasound scanners. These findings suggest that ultrasound pSWE is a reliable and reproducible method for quantitative assessment of tissue stiffness.

Further work is required to investigate the reproducibility of ultrasound pSWE in healthy subjects and humans under pathological conditions.

The excellent reproducibility of pSWE measurements seen within and between operators and scanners in the present study was also reported by Mulabecirovic et al. (2018) [20], who made pSWE measurements in kPa using a curvilinear 5–1 MHz transducer with a Philips iU22 device on a liver fibrosis phantom with lesions of elasticity ranging from values found in normal to cirrhotic liver tissues. Similar findings of high pSWE reproducibility were also seen using linear 9–4 MHz and curvilinear 4–1 MHz transducers with a Siemens Acuson S3000 ultrasound system on soft and hard cylindrical phantoms at depths of 1, 2.4, and 4 cm [21]. Another study in which the same phantom in the present study was used reported that pSWE and 2D-SWE had good and excellent intra-operator reproducibility, respectively, and both showed excellent inter-operator reproducibility [22]. Other studies have assessed the reproducibility of different SWE platforms with phantoms containing lesions with stiffness ranging from 3.6 to 18.4 kPa, and have reported that pSWE measurements showed excellent intra-operator and inter-operator reproducibility, but had the higher CV values than transient elastography and 2D-SWE imaging techniques [11,21]. These findings from phantom studies suggest that ultrasound pSWE is a reliable imaging method with reproducibility ranging from good to excellent.

Ultrasound imaging is operator-dependent; thus, the operator’s experience level may play an important role in lowering the variability of pSWE measurements within and between operators. Seliger et al. (2017) [22] also showed that the use of different regions of interest and measurement techniques can affect the reproducibility of pSWE measurements. Higher lesion stiffness has been reported as a source of variability in pSWE measurements [23–25]. This was noted in the present study, as more widespread pSWE measurements from the fitted regression line were noted for values of >30 kPa and ≥25 kPa for operators A and B, respectively (Fig. 3A, B). In addition, the analysis herein revealed significant differences between the measured pSWE and the manufacturer-reported values for all lesions. This finding is consistent with a number of studies assessing the accuracy of pSWE on CIRS phantoms [14,19,24]. This could be related to transducer pressure during pSWE measurements [26]; although minimal compression is used, it remains difficult to determine the actual pressure being applied during measurements. These factors should be considered in research investigating the reproducibility of pSWE imaging techniques in phantoms. Further potential factors such as breathing, health conditions, and tissue stiffness between subjects can alter the reproducibility of pSWE and should be considered in future human studies.

There are several limitations of this study. The stiffness of the superficial lesions was not assessed in this study due to the poor image resolution using the curvilinear 5–1 MHz transducer. Furthermore, pSWE of the phantom lesions was conducted using two different Philips ultrasound scanners. Further studies on the same phantom assessing pSWE variability between superficial and deep lesions, as well as the effect of depth using transducers with different frequencies and ultrasound imaging systems from different manufacturers, are required.

In conclusion, this study showed that pSWE is a reproducible imaging method for the assessment of tissue stiffness within and between operators, as well as between Philips ultrasound scanners (i.e., EPIQ7 and iU22), for lesions at a depth of 3.5 cm in an elasticity quality assurance phantom (Model 049, CIRS). There were significant differences between the measured pSWE and values reported by the phantom manufacturer for all deep lesions. Further studies should investigate variability in pSWE using the same phantom to compare superficial and deep lesions, as well as the effect of depth, using ultrasound imaging systems from different manufacturers.

ORCID: Salahaden R. Sultan: https://orcid.org/0000-0001-9981-6138; Amer Alghamdi: https://orcid.org/0000-0002-0445-4689; Rawan Abdeen: https://orcid.org/0000-0002-7578-4909; Fahad Almutairi: https://orcid.org/0000-0002-0494-5461

Author Contributions

Conceptualization: Sultan SR, Abdeen R, Almutairi F. Data acquisition: Sultan SR, Alghamdi A. Data analysis or interpretation: Sultan SR,
Almutairi F. Drafting of the manuscript: Sultan SR, Alghamdi A, Abdeen R, Almutairi F. Critical revision of the manuscript: Almutairi F. Approval of the final version of the manuscript: all authors.

Conflict of Interest
No potential conflict of interest relevant to this article was reported.

References
1. Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-analysis. Clin Gastroenterol Hepatol 2007;5:1214-1220.
2. Bercoff J, Tanter M, Muller M, Fink M. The role of viscosity in the impulse diffraction field of elastic waves induced by the acoustic radiation force. IEEE Trans Ultrason Ferroelectr Freq Control 2004;51:1523-1536.
3. Sigrist RMS, Liau J, Kaffas AE, Chammas MC, Willmann JK. Ultrasound elastography: review of techniques and clinical applications. Theranostics 2017;7:1303-1329.
4. Lee SM, Kim MJ, Yoon JH, Hong W, Ha HI, Lee K, et al. Comparison of point and 2-dimensional shear wave elastography for the evaluation of liver fibrosis. Ultrasonography 2020;39:288-297.
5. Youk JH, Gweon HM, Son EJ. Shear-wave elastography in breast ultrasonography: the state of the art. Ultrasonography 2017;36:300-309.
6. Kim HJ, Kwak MK, Choi IH, Jin SY, Park HK, Byun DW, et al. Utility of shear wave elastography to detect papillary thyroid carcinoma in thyroid nodules: efficacy of the standard deviation elasticity. Korean J Intern Med 2019;34:850-857.
7. Fu J, Wu B, Wu H, Lin F, Deng W. Accuracy of real-time shear wave elastography in staging hepatic fibrosis: a meta-analysis. BMC Med Imaging 2020;20:16.
8. Wells PN, Liang HD. Medical ultrasound: imaging of soft tissue strain and elasticity. J R Soc Interface 2011;8:1521-1549.
9. Shiina T, Nightingale KR, Palmeri ML, Hall TJ, Bamber JC, Barr RG, et al. WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 1: basic principles and terminology. Ultrasound Med Biol 2015;41:1126-1147.
10. Hudson JM, Milot L, Parry C, Williams R, Burns PN. Inter- and intra-operator reliability and repeatability of shear wave elastography in the liver: a study in healthy volunteers. Ultrasound Med Biol 2013;39:950-955.
11. Palmeri M, Nightingale K, Fielding S, Rouze N, Deng Y, Lynch T, et al. RSNA QIBA ultrasound shear wave speed Phase II phantom study in viscoelastic media. In: 2015 IEEE International Ultrasonics Symposium; 2015 Oct 21-24; Taipei, Taiwan. Piscataway, NJ: Institute of Electrical and Electronics Engineers Inc., 2015.
12. Computerized Imaging Reference Systems [Internet]. Norfolk, VA: Computerized Imaging Reference Systems, 2016 [cited 2021 Apr 16]. Available from: https://www.cirsinc.com/products/ultrasound/zerdine-hydrogel/elasticity-qa-phantom/.
13. Cournoy S, Fagan AJ, Browne JE. Review of ultrasound elastography quality control and training test phantoms. Ultrason Med Biol 2012;20:16-23.
14. Mulabecirovic A, Vesterhus M, Giljia OH, Havre RF. In vitro comparison of five different elastography systems for clinical applications, using strain and shear wave technology. Ultrasound Med Biol 2016;42:2572-2588.
15. Hollerieth K, Gassmann B, Wagenpeil S, Kemmer S, Heemann U, Stock KF. Does standoff material affect acoustic radiation force impulse elastography?: a preclinical study of a modified elastography phantom. Ultrasonography 2018;37:140-148.
16. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-163.
17. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015;25:141-151.
18. Moore DS, Notz WI, Flinger MA. The basic practice of statistics. 6th ed. New York: Freeman and Company, 2013.
19. Bland M. An introduction to medical statistics. 3rd ed. Oxford: Oxford University Press, 2000.
20. Mulabecirovic A, Mjelle AB, Giljia OH, Vesterhus M, Havre RF. Repeatability of shear wave elastography in liver fibrosis phantoms-Evaluation of five different systems. PLoS One 2018;13:e0189671.
21. Dillman JR, Chen S, Davenport MS, Zhao H, Urban MW, Song P, et al. Superficial ultrasound shear wave speed measurements in soft and hard elasticity phantoms: repeatability and reproducibility using two ultrasound systems. Pediatr Radiol 2015;45:376-385.
22. Seliger G, Chaoui K, Kunze C, Dridi Y, Jenderka K, Wiencek A, et al. Intra- and inter-observer variation and accuracy using different shear wave elastography methods to assess circumscribed objects - a phantom study. Med Ultrason 2017;19:357-365.
23. Lee SM, Chang W, Kang HJ, Ahn SJ, Lee JH, Lee JM. Comparison of four different shear wave elastography platforms according to abdominal wall thickness in liver fibrosis evaluation: a phantom study. Med Ultrason 2019;21:22-29.
24. Tozaki M, Saito M, Joo C, Yamaguchi M, Isobe S, Ogawa Y, et al. Ultrasoundographic tissue quantification of the breast using acoustic radiation force impulse technology: phantom study and clinical application. Jpn J Radiol 2011;29:598-603.
25. Chang S, Kim MJ, Kim J, Lee MJ. Variability of shear wave velocity using different frequencies in acoustic radiation force impulse (ARFI) elastography: a phantom and normal liver study. Ultraschall Med 2013;34:260-265.
26. Byenfeldt M, Elvin A, Fransson P. Influence of probe pressure on ultrasound-based shear wave elastography of the liver using Comb-Push 2-D technology. Ultrasound Med Biol 2019;45:411-428.