



Ultrasound imaging measures of vertebral bony landmark distances are weakly to moderately correlated with intervertebral disc height as assessed by MRI

Ulrike H Mitchell ¹, A Wayne Johnson,¹ Lauren Adams,¹ Tayva Sonnefeld,¹ Patrick J Owen ²

To cite: Mitchell UH, Johnson AW, Adams L, *et al*. Ultrasound imaging measures of vertebral bony landmark distances are weakly to moderately correlated with intervertebral disc height as assessed by MRI. *BMJ Open Sport & Exercise Medicine* 2022;**8**:e001292. doi:10.1136/bmjsem-2021-001292

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjsem-2021-001292>).

Accepted 18 March 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Exercises Sciences, Brigham Young University, Provo, Utah, USA

²Institute for Physical Activity and Nutrition, Deakin University, Burwood, Victoria, Australia

Correspondence to
Dr Ulrike H Mitchell;
rike_mitchell@byu.edu

ABSTRACT

Objectives To assess the validity and reliability of ultrasound-derived interbony landmark distances as a proxy for MRI-derived intervertebral disc (IVD) height.

Methods This is a cross-sectional criterion validity study. Twelve college-aged participants without current low back pain completed both MRI and ultrasound imaging of the lumbar spine in a prone position. Single-segment and multisegment distances between the spinous and mammillary processes at the lumbar segments (L2/L3, L3/L4, L4/L5) were measured twice using ultrasound and analysed digitally. Sagittal slices of the lumbar spine were taken via T1-weighted MRI and IVD height, and the overall distance between IVDs L2/L3 and L4/L5 was imaged once and measured twice.

Results There was moderate correlation between multilevel-based measurements (overall distance between L2 and L5, $r=0.677$, $p=0.016$) and the average across three levels ($r=0.596$, $p=0.041$) when using the spinous processes as bony landmarks. Single-segment measures were not significantly correlated (all: $p>0.092$). Accuracy and precision were better for the overall MRI-derived distance between the three IVDs from L2 and L5 MRI and the distance measured between the spinous processes L2–L5. There was excellent reliability within multiple measurements at each location, with intraclass correlation coefficient, ICC_(3,1), ranging from 0.93 to 0.99 (95% CI 0.82 to 0.99) for ultrasound and from 0.98 to 0.99 (95% CI 0.92 to 0.99) for MRI.

Conclusion Findings do not support the use of ultrasound imaging for estimating single-segment IVD height, yet it may be used to measure the change in distance over time with a certain degree of precision based on its excellent reliability.

INTRODUCTION

The spine height of a fully grown individual decreases with load-bearing and increases with unloading.¹ For example, during space flight, a spinal height increase of 2 inches is common,² which is largely due to water fluctuations occurring in the intervertebral

Key messages

What is already known on this topic

- Changes in lumbar spine height are related to water fluctuations which drastically change with ageing and disease processes.
- Ultrasound imaging has been used to measure lumbar spine height, but the results have not been validated against data obtained from MRI, or the data were not collected with the participant in a prone position.

What this study adds

- There is some validity in using ultrasound-derived multisegment but not single-segment distances to estimate related intervertebral disc height.

How this study might affect research, practice or policy

- The assessment of spinal height changes over time could become an additional tool for clinicians who use point-of-care ultrasound imaging.

disc (IVD), and places the astronaut at an increased risk of IVD trauma.³ Measuring IVD height and assessing changes thereof can provide valuable information on the hydration status of the IVD.² Decreases in IVD height occur with ageing and disease processes.⁴ Together with other radiological signs, such as the presence of osteophytes and endplate sclerosis, decreased IVD height is considered a sign of disc degeneration.⁵ In addition, loss of disc height has been correlated with low back pain,⁶ and therefore knowledge of this phenomenon could be of value for the treating clinician. Point-of-care ultrasound imaging is becoming more popular as a form of examination. It is used with 'the intent of clarifying uncertain clinical examination findings to enhance the quality and effectiveness of a physical therapy

intervention'.⁷ The results of the systematic review published in 2019⁸ endorse the addition of point-of-care ultrasound imaging to the assessment and treatment process.

Various methods have been employed to study immediate changes in IVD height, such as MRI, stadiometry, X-ray and ultrasound imaging.^{1,9–12} MRI and X-rays can measure IVD height directly, while stadiometry and ultrasound imaging estimate IVD height indirectly.¹³ MRI is considered the reference standard when imaging the spine,¹⁴ yet is often not available or obtainable for different reasons. X-ray, while also yielding high-resolution images, exposes patients to potentially harmful ionising radiation.¹⁵ Ultrasound imaging does not provide exposure to ionising radiation and has the additional advantages of being financially cost-effective and portable. However, the method yields comparatively lower-resolution images and may be less accurate. Nevertheless, ultrasound imaging has been used to measure spinal segment height *in vivo*^{1,2,16–18} and *in vitro*.¹³ Distances between different bony landmarks, including lumbar transverse processes,¹⁷ mammillary processes,¹ anterior vertebral bodies² and spinous processes,^{16,18} have been used. Still, the results were either not validated against data obtained from MRI,^{1,2,17,18} or the data were not collected with the participant in prone position.^{1,2,16}

Ultrasound imaging of the lower back cannot be easily done with the patient in supine position due to inaccessibility of the spine and associated probe placement issues.

The specific aims of this study were to (1) assess the validity of estimating IVD height from ultrasound images by measuring single-segment and multisegment distances between spinous processes and mammillary processes and compare them with single-segment and multisegment MRI-derived IVD height, respectively; and (2) assess intrarater test–retest reliability and intrarater reliability of measuring distances between posterior bony landmarks using ultrasound imaging and IVD heights on MRI scans. We hypothesised that intermammillary distances are better proxies for IVD heights than interspinous processes distances.

METHODS

This is a cross-sectional criterion validity study. Participants were recruited from the university community and by word of mouth. We included individuals aged 18 years and older without current, that is, within the previous week, low back pain. Informed consent was obtained before data collection, and the rights of the participants were protected. Each participant was informed that their data would be submitted for publication and that their confidentiality would be protected. The exclusion criteria were (1) inability to lie on stomach for 30 min; (2) history of spinal surgery; (3) history of traumatic injury to the spine; (4) known scoliosis for which prior medical consultation was sought; (5) known claustrophobia; and (6) absolute contraindications to MRI.¹⁹

Imaging

Ultrasound and MRI were performed in a prone lying position to directly compare the measurements and keep the spine in a consistent position across imaging modalities. Due to the ferromagnetic characteristics of the ultrasound machine, the participant had to be moved out of the MRI room and repositioned for the ultrasound imaging procedure. A folded-up small pillow was placed under the hips to decrease lordosis of the lumbar spine. A 3-inch high rolled-up towel also supported the forehead, and the arms were resting at the side.

Ultrasound imaging

Ultrasound images (Logiq e, Probe, GE Healthcare, Chicago, Illinois, USA) were captured by placing the probe (12 MHz linear) in longitudinal orientation on the participant's skin with a water-based gel as a coupling agent between the probe and the skin. The frequency ranged from 8 MHz to 12 MHz as needed to enhance imaging of the bony structure. The spinous processes were first visualised. Separate images were taken of each of the following pairs of spinous processes: L2 and L3, L3 and L4, as well as L4 and L5. Keeping the probe in longitudinal orientation, it was moved about 2 cm laterally off midline to locate the posterior-most hyperechoic points, typically indicating the mammillary processes of L5–L2. Using the LogiqView mode, a panoramic image was captured. The probe was placed distally with the right edge of the sacrum in the field of view, the orientation marker on the probe pointed cephalad. Slowly the probe was glided cranially along with the mammillary processes until L2 was visualised. This was repeated for the mammillary processes on the contralateral side. Images of the spinous processes and intermammillary processes were recorded for processing after the assessment. This imaging process was performed twice for each side, with the probe briefly removed from the skin between each scan without repositioning the participant. No specific breathing instructions were given to the participant for the scan because ultrasound imaging is less prone to artefacts with restful breathing compared with MRI.²⁰

The same examiner, who has 1 year of experience with this imaging modality, collected all images and performed the postimaging measurements using the OsiriX (Pixmeo, Geneva, Switzerland) software program in a blinded fashion. The distances between the mammillary processes were measured by first drawing straight vertical lines through the apex (most posterior aspect) of each mammillary process (L2/L3, L3/L4 and L4/L5). Horizontal lines were drawn between each vertical line (see [figure 1](#)). Interspinous process distance was measured by drawing vertical lines through the apex of the spinous processes with a horizontal connecting line between the two (see [figure 2](#)). In rare cases, there were two apices on one lumbar spinous process. In these cases, we chose the most posterior one.

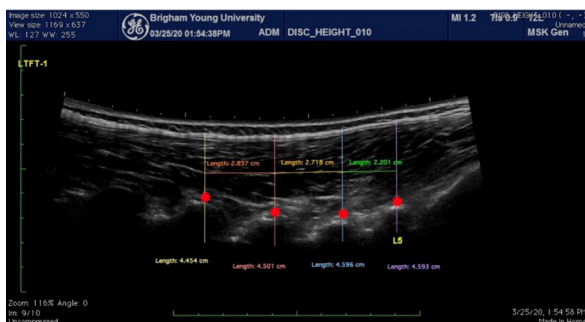


Figure 1 Straight vertical lines were drawn through the mamilary processes, signifying the apex of each facet joint (L2/L3, L3/L4 and L4/L5). Horizontal lines were then drawn between neighbouring vertical lines and measured. *Indicates the apex of the hyperechoic bone cortex, the location of the mamilary processes.

Magnetic resonance imaging

MRI was performed using the 3-Tesla Siemens MAGNETOM Tim Trio (Siemens Healthineers, Erlangen Germany) scanner with a four-channel flexible coil. The participant was placed prone on the imaging table. The inferior border of the 51.2 cm × 22 cm coil was positioned caudally on the posterior superior iliac spines and placed in a cranial direction along the long axis of the spine. The participant was instructed to breathe shallowly during the scan. A total of 36 sagittal slices were acquired using a fast, low-angle shot sequence using the following parameters: field of view 192×173×119 mm, repetition time 13 ms, echo time 2 ms, slices 36, flip angle 10°, resolution 1.5×1.5 mm on the plane, 3 mm through the plane and no fat suppression. The scan took approximately 38 s. After obtaining images in Digital Imaging and Communications in Medicine (DICOM) format, they were viewed on OsiriX (Pixmeo). A midsagittal plane view was used and the IVDs of interest (L2/L3, L3/L4 and L4/L5) were manually segmented from the two-dimensional T1-weighted structural scan images. Lines connecting the superior anterior and inferior anterior borders and the superior posterior and inferior posterior borders were drawn, representing the anterior and posterior height of the IVD. A horizontal line was drawn connecting the anterior and posterior lines. An additional horizontal line was

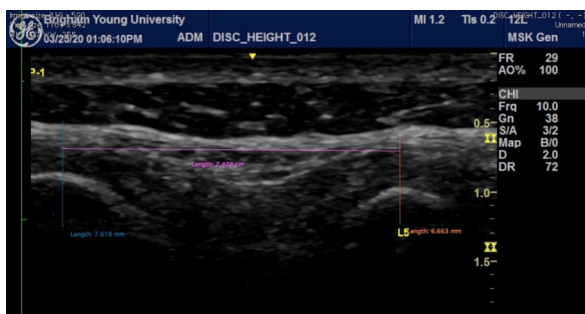


Figure 2 Vertical lines were drawn through the apices of the spinous processes of vertebrae L2, L3, L4 and L5. Horizontal lines were then drawn between neighbouring vertical lines and measured.

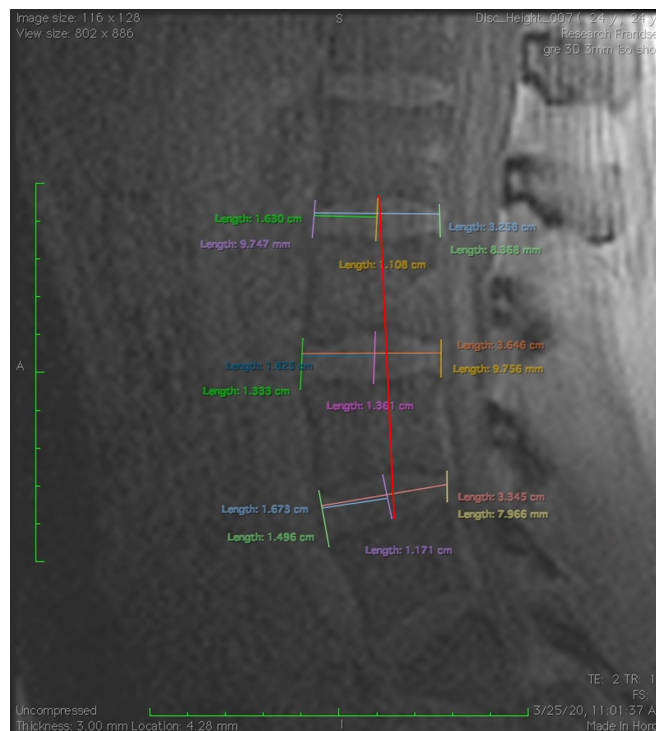


Figure 3 Anterior/posterior and mid-IVD height at segments L2/L3, L3/L4 and L4/L5 as well as the distance between the superior border of L2/L3 to the inferior border of L4/L5 were measured. The latter was done by drawing a straight line between the cephalad midpoint of IVD at segment L2/L3 and the caudal midpoint of IVD at segment L4/L5. IVD, intervertebral disc.

drawn to mark the halfway point of the full IVD width. We then drew a line perpendicular to the vertebral body at the halfway mark of the IVD.²¹ An additional vertical line was drawn from the superior midpoint at IVD segment L2/L3 to the inferior midpoint at IVD segment L4/L5 to capture the multisegment distance that includes three IVDs (figure 3).

Statistical analyses were conducted using Stata V.16 statistical software. For assessing measurement precision using four ultrasound examinations per participant, it is recommended that at least nine participants are measured²²; hence, our study was capable of detecting reasonable estimates of reliability. The strength and direction of association between MRI and ultrasound-derived IVD height were assessed by Pearson correlation coefficient (weak: 0.10–0.39; moderate: 0.40–0.69; strong: >0.7).²³ Univariate linear regression was used to quantify the proportion of variance of MRI-derived IVD height explained by ultrasound-derived methods. Bland-Altman plots were generated to assess agreement between MRI and ultrasound imaging methods and to graphically depict the ultrasound–MRI difference as a function of the mean for each subject.²⁴ Analyses used average IVD height at segments L2/L3, L3/L4 and L4/L5 (ie, the average of the anterior, mid and posterior height), the average from across the three segments, and the total distance from the superior border of L2/L3 to the inferior

border of L4/L5 (L2–L5) as obtained from the MRI. An alpha level of 0.05 was adopted for statistical tests. Calculation of intraclass correlation coefficient (ICC) was performed using SPSS V.28 for consistency with random subjects and fixed raters, mixed model ICC (model_{3,k}). Intrarater test–retest reliability, that is, the same rater measures the first and second image, of the ultrasound and intrarater reliability, that is, the rater measures the same image twice, of MRI measurements were obtained by computing a two-way mixed-model ICC_(3,1) with absolute agreement. To identify testing error we calculated the SE of the measurement (SE_m) and per cent SE_m (%SE_m) and the minimum detectable difference at 95% (MDD at 95%) using the following equations: SE_m=SD (Sq rt 1–r_{icc}); 95% CI SE_m=muscle mean±(1.96×SE_m); MDD=SE_m×1.96×√2.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

RESULTS

Twelve asymptomatic participants (8 men: mean±SD age 24.1±1.7 years, height 182.0±7.9 cm, weight 84.6±23.6 kg; 4 women: age 23.5±2.2 years, height 171.0±6.3 cm, weight 63.0±8.0 kg) completed both MRI and ultrasound imaging of the lumbar spine within several minutes of each other.

The average distances for both ultrasound and MRI imaging are presented in table 1.

Comparison between methods

The total distance between the L2 and L5 spinous processes (ultrasound) moderately correlated with MRI-derived distance from the superior border of IVD at the L2/L3 level to the inferior border of IVD at the L4/L5 level (r=0.677, p=0.016). IVD height, when measured indirectly via spinous process distance using ultrasound

imaging, moderately correlated with MRI-derived IVD height when averaged across the three IVDs examined (r=0.596, p=0.041), yet not at each individual IVD (all, p≥0.092; table 2).

Linear regression revealed that the distance between the L2 and L5 spinous processes explained approximately 46% (R²=0.458) of the variance in MRI-derived distance from the superior border of L2/L3 IVD to the inferior border of L4/L5 IVD. Moreover, the average interspinous distance explained approximately 36% (R²=0.355) of the variance in average MRI IVD height. These values were greater than the variance explained by mammillary-based methods (all: <2.7%).

Bland-Altman analysis revealed a mean bias of –0.159 cm and wide limits of agreement (LoA) ranging from –1.288 to 0.970 for the overall distance between L2 and L5 measured by MRI and ultrasound imaging using spinous processes as bony reference (online supplemental material figure 1). The mean bias for the averages (average IVD height derived from levels L2/L3, L3/L4 and L4/L5 and average distances between spinous processes L2, L3, L4 and L5) was –2.123 cm, and the LoA were narrower, ranging from –2.518 cm to –1.729 cm.

Within-method repeatability

The methods of measuring the distance between the spinous processes and the mammillary processes and IVD height were reliable for ultrasound imaging and MRI, respectively (table 3).

DISCUSSION

The results of our study indicate that there is some validity in using ultrasound-derived multisegment bony distances (between mammillary processes and between spinous processes) as a measure of multisegment IVD height. Single-segment measurements, however, showed little validity. More specifically, we observed significant moderate correlations (r=0.60–0.68) for the

Table 1 Ultrasound and MRI measurements

Ultrasound measurements between bony landmarks (cm)																
Distance between	Intermammillary distance, right side					Intermammillary distance, left side					Spinous processes					
	L4/L5 (1)	L3/L4 (1)	L2/L3 (1)	AVG (3)	L2/L5 (3)	L4/L5 (1)	L3/L4 (1)	L2/L3 (1)	AVG (3)	L2/L5 (3)	L4/L5 (1)	L3/L4 (1)	L2/L3 (1)	AVG (3)	L2/L5 (3)	
Mean	2.55	2.90	3.21	2.89	8.67	2.69	2.93	3.21	2.95	8.84	2.61	2.91	2.94	2.81	8.44	
SD	0.29	0.28	0.38	0.33	0.81	0.27	0.32	0.33	0.26	0.63	0.37	0.28	0.38	0.18	0.75	

MRI-measured IVD height (cm)																
Disc levels	L4/L5A	L4/L5M	L4/L5P	L4/L5AVG	L3/L4A	L3/L4M	L3/L4P	L3/L4AVG	L2/L3A	L2/L3M	L2/L3P	L2/L3AVG	AvLx L5	L2–L5		
	Mean	0.84	0.92	0.53	0.76	0.75	0.86	0.50	0.70	0.64	0.75	0.48	0.63	0.70	8.31	
SD	0.17	0.15	0.10	0.12	0.12	0.14	0.09	0.10	0.17	0.14	0.13	0.12	0.11	0.30		

A, anterior; AVG, average; AvLx, average IVD height derived from the L2/L3, L3/L4 and L4/L5 levels; IVD, intervertebral disc; M, middle; P, posterior.

Table 2 Correlations between MRI and ultrasound methods

	Ultrasound, distance between the right corresponding mammillary processes	Ultrasound, distance between the left corresponding mammillary processes	Ultrasound, distance between the corresponding spinous processes
MRI, L2–L5	0.046 (–0.542 to 0.604), p=0.887	–0.008 (–0.579 to 0.569), p=0.981	0.677 (0.169 to 0.901), p=0.016*
MRI, AvLx	0.057 (–0.534 to 0.611), p=0.860	–0.166 (–0.675 to 0.451), p=0.607	0.596 (0.034 to 0.872), p=0.041*
MRI, L2/L3	0.008 (–0.569 to 0.579), p=0.981	–0.417 (–0.800 to 0.206), p=0.177	0.508 (–0.093 to 0.838), p=0.092
MRI, L3/L4	0.139 (–0.472 to 0.660), p=0.666	0.181 (–0.438 to 0.684), p=0.574	0.358 (–0.271 to 0.773), p=0.253
MRI, L4/L5	–0.171 (–0.678 to 0.447), p=0.595	–0.081 (–0.626 to 0.517), p=0.804	0.365 (–0.264 to 0.776), p=0.244

Data are Pearson correlation coefficient, (95% CI) and p value.

L2–L5: total distance from the posterior superior border of L2/L3 to the inferior border of L4/L5.

*P<0.05.

AvLx, average disc height across L2/L3, L3/L4 and L4/L5.

multivertebral segment-based measurements (L2–L5) and the average across these three segments, while single-segment measurements were not significantly correlated with their respective MRI-derived IVD height. This suggests that averaging or combining distances likely helps account for variations between individual vertebral segments.

Our hypothesis that intermammillary distances are better proxies for IVD heights than interspinous processes distances was not confirmed. Considering that any distance of the same landmark measured between adjacent vertebrae includes a part or the whole height of the vertebral body, we recognise that this problem is the same for any bony landmark. However, measuring the distance between two different bony landmarks in

neighbouring vertebrae, for example the mammillary process of the inferior vertebra to the transverse process of the superior vertebra, could potentially lessen this problem.

Bland-Altman

To further understand the relationship between ultrasound-derived and MRI-derived methods we used to estimate lumbar IVD height, we also employed Bland-Altman plots. The precision of the measurements (ie, the narrower the LoA, the more precise the measurements) and degree of agreement between the two methods can be gleaned from the Bland-Altman plots. The LoA in our study ranged from –1.288 cm to 0.970 cm for MRI-derived overall distance between three IVDs (L2–L5)

Table 3 Distances between spinal segments and reliability indicators

Ultrasound	Segment	Intermammillary distance (right and left, averaged)				Interspinous process distance			
		L4–L5	L3–L4	L2–L3	L2–L5	L4–L5	L3–L4	L2–L3	L2–L5
	Mean (cm)	2.62	2.92	3.21	8.75	2.61	2.91	2.94	8.44
	SD	0.28	0.29	0.35	0.72	0.37	0.28	0.38	0.75
	SEm	0.07	0.08	0.09	0.12	0.05	0.03	0.02	0.18
	%SEm	2.71	2.63	2.94	1.33	2.10	1.12	0.83	2.16
	MDD at 95%	0.20	0.21	0.26	0.32	0.15	0.09	0.07	0.51
	ICC _(3,1)	0.94	0.93	0.93	0.97	0.98	0.99	1.00	0.94
	ICC 95% CI	0.85 to 0.97	0.84 to 0.97	0.82 to 0.97	0.94 to 0.99	0.92 to 0.99	0.95 to 1.00	0.98 to 1.00	0.86 to 0.98
Average intervertebral disc height									
MRI	Segment	L4–L5	L3–L4	L2–L3	L2–L5				
	Mean (cm)	0.76	0.68	0.62	8.29				
	SD (cm)	0.11	0.15	0.12	0.29				
	SEm	0.02	0.02	0.01	0.03				
	%SEm	1.49	2.27	1.86	1.73				
	MDD at 95%	2.21	3.51	2.12	0.40				
	ICC _(3,1)	0.98	0.98	0.99	0.99				
	ICC 95% CI	0.96 to 0.99	0.96 to 0.99	0.96 to 0.99	0.96 to 0.99				

ICC, intraclass correlation coefficient; MDD, minimal detectable difference; SEm, SE of the measurement.

compared with the ultrasound-derived distance between spinous processes L2–L5. The LoA for the average MRI-derived disc height and the average ultrasound-derived spinous process distance were smaller (range: -2.518 cm to -1.729 cm) and should therefore be considered more precise. The former LoA were not narrow enough to conclude that the method of indirectly measuring IVD height was precise. The latter indicated that the ultrasound imaging measurements were within 0.4 cm of the MRI-derived values. We suggest that the 0.4 cm difference between ultrasound imaging measurements and MRI-derived IVD height could be explained by the respective anatomical starting and stopping points of the two measurement locations: the MRI measurements started at the inferior border of the vertebral body of L2 and ended at the superior border of L5, while the ultrasound measurement began at the apex of the L2 spinous process, which corresponds more closely to the middle of the body of L2 and ended at the apex of L5 spinous process, which again corresponds more closely to the middle of the vertebral body. Thus, based on the differences in the measured structures, it is not surprising that the distances measured are dissimilar, although consistently so.

The degree of agreement between the ultrasound and MRI measures can be extrapolated from the bias shown in the Bland-Altman plot (online supplemental material figure 1; only the significant correlations are shown). The x-axes on the Bland-Altman plots represent the total (online supplemental material figure 1A) and average (online supplemental material figure 1B) distances between L2 and L5 as measured by ultrasound (spinous processes) and by MRI (three IVDs and two vertebral bodies) ($\text{distance} = (\text{MRI distance} + \text{ultrasound distance}) / 2$) for each participant. The y-axes on the Bland-Altman plots represent the absolute differences (total and average distance difference, respectively) between the two values ($\text{distance difference} = \text{MRI distance} - \text{ultrasound distance}$) for each participant. The mean bias for the data representing the overall distance between segments L2 and L5 was -0.159 cm, while it was -2.123 cm for the average distance. Since the mean biases are smaller than 0, it indicates that the ultrasound-derived measures are consistently larger than the MRI measures.

The positive correlation between IVD height and the discrepancy between ultrasound and MRI measurements seen in both Bland-Altman graphs indicate that the discrepancy between MRI and ultrasound imaging values becomes larger with increasing multisegment distance. This could be a function of increasing IVD or/and increasing vertebral body heights. Regardless, clinically this implies that ultrasound measurements tend to underestimate the overall distance (and overestimate to a lesser degree the average distance) between three IVDs in people with shorter lumbar spines and overestimate it in people with longer lumbar spines.

Linear regression

The linear regression analysis sheds further light on the relationship between the ultrasound-derived and MRI-derived methods we used to estimate lumbar IVD height. Linear regression revealed that the ultrasound-derived multisegment distance between the L2 and L5 spinous processes accounted for about 46% of the variance of the data obtained via MRI. This finding supports a relationship between the two measurement methods and modalities, although a moderate one. However, we believe that this result is more a function of the correlation between lumbar vertebral body height when measured with ultrasound imaging and compared with our reference standard obtained via MRI. This is based on the fact that, on average, vertebral bodies make up more than 60% of the distance from segments L2–L5.²⁵

Reliability and repeatability

The second aim of this study was to assess intrarater test-retest reliability and intrarater reliability. Our methods, ultrasound imaging (captured twice, measured once) and MRI (captured once, measured twice), showed excellent reliability, with ICC between 0.93 and 1.00. Interestingly, the ICC for the single-segment measurements of intermammillary distances was lower than the multisegment measurements; this was reversed for measurements using interspinous process distances (both using ultrasound imaging).

Minimum detectable difference

The MDD is a useful clinical variable that indicates the amount of change needed to have a significant difference in a given population. Our results for MDD at 95% (table 3) suggest that the ultrasound imaging protocol can detect a single-segment distance change of 2–3 mm, depending on the lumbar segment, when using the mammillary processes as bony landmarks. When using the spinous processes as bony landmarks, this value is even smaller, ranging between 0.7 mm and 1.5 mm. We submit that these values are small enough to suggest this measuring tool is reliable. To confirm this proposition we also assessed the possible measuring error and relative measuring error, as indicated by the %SEM. The values for our %SEM range between 2.6 and 2.9. This indicates that we can expect a measuring error of up to 2.9% or 0.7 mm on a measured distance of 2.7 cm (eg, between the L4/L5 mammillary processes), which is considered to be good.²⁶ Our results lead us to conclude that our ultrasound imaging protocol is reliable in making multisegment measurements, although not to estimate IVD height. Of note, our results apply to asymptomatic individuals rather than those with low back pain, so generalisation should be made with caution.

Limitations

Despite our effort to design a carefully controlled study, it also had several potential limitations. Our sample size is relatively small but larger than the sample size of other

studies.^{16 18 27} Our population sample included only healthy young adults who have a relatively small likelihood of having IVD or vertebral body pathology. In the case of presence of corner osteophytes and Schmorl's nodes, the MRI-derived IVD height measurement protocol would have to be adjusted. We only assessed the validity and reliability of segments L2/L3, L3/L4 and L4/L5. We did not consider any segments above and below. While ultrasound imaging cannot be performed in a supine position, it is the preferred position for MRI. The lordotic angle will differ between these two positions and thus affect the anterior and posterior IVD dimensions (as measured on MRI) and the interbony distances (as measured via ultrasound imaging). Lastly, selection of MRI images may not correspond to image selection in ultrasound imaging. This is most pertinent to our interspinous process distance measurements since the MRI-derived and ultrasound-derived images should be collected in the midsagittal plane.

CONCLUSIONS

This study demonstrates that there is some validity in using ultrasound-derived multisegment bony distances as a measure of multisegment IVD height. However, single-segment intermammillary and interspinous process distances are not good estimates of the actual respective single-segment lumbar IVD height. Nevertheless, because the ultrasound-derived data demonstrate excellent intrarater test–retest reliability, these measurements could be investigated for their ability to assess changes in lumbar spine height over time. This is especially likely for the multisegment measurements using either measurement method because their measurement repeatability is excellent as indicated by their high ICC, low SEM, low %SEM and low MDD. The assessment of spinal height changes over time could therefore become an additional tool for clinicians who use point-of-care ultrasound imaging.

Twitter Patrick J Owen @PatrickOwenPhD

Acknowledgements We want to thank the participants of this study.

Contributors UHM: responsible for the overall content as guarantor. AWJ: study design, data collection, interpretation of data, writing and approval of the final version of the manuscript. LA: data collection, methods development, writing and approval of the final version of the manuscript. TS: data collection, writing and approval of the final version of the manuscript. PJO: statistical analysis, interpretation of data, writing and approval of the final version of the manuscript.

Funding This work was supported by internal department funding and generous donations from the BYU MRI research facility.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study involves human participants and was approved by the Brigham Young University Institutional Review Board (approval number X18138). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request to the corresponding author (Ulrike H Mitchell, Department of Exercise Sciences, Provo, Utah).

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been

peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Ulrike H Mitchell <http://orcid.org/0000-0002-2298-8756>

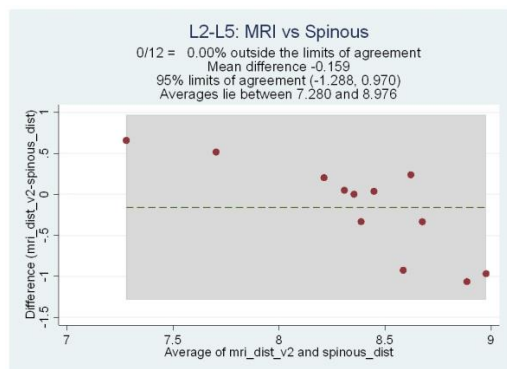
Patrick J Owen <http://orcid.org/0000-0003-3924-9375>

REFERENCES

- Poortmans V, Brismée J-M, Poortmans B, *et al.* Assessment of lumbar spine height following sustained lumbar extension posture: comparison between musculoskeletal ultrasonography and Stadiometry. *J Manipulative Physiol Ther* 2016;39:586–93.
- Marshburn TH, Hadfield CA, Sargsyan AE, *et al.* New heights in ultrasound: first report of spinal ultrasound from the International space Station. *J Emerg Med* 2014;46:61–70.
- Johnston SL, Campbell MR, Scheuring R, *et al.* Risk of herniated nucleus pulposus among U.S. astronauts. *Aviat Space Environ Med* 2010;81:566–74.
- Elfering A, Semmer N, Birkhofer D, *et al.* Risk factors for lumbar disc degeneration: a 5-year prospective MRI study in asymptomatic individuals. *Spine* 2002;27:125–34.
- Akeda K, Yamada T, Inoue N, *et al.* Risk factors for lumbar intervertebral disc height narrowing: a population-based longitudinal study in the elderly. *BMC Musculoskelet Disord* 2015;16:344.
- de Schepper EIT, Damen J, van Meurs JBJ, *et al.* The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. *Spine* 2010;35:531–6.
- Whittaker JL, Ellis R, Hodges PW, *et al.* Imaging with ultrasound in physical therapy: what is the Pt's scope of practice? A competency-based educational model and training recommendations. *Br J Sports Med* 2019;53:1447–53.
- Andersen CA, Holden S, Vela J, *et al.* Point-Of-Care ultrasound in general practice: a systematic review. *Ann Fam Med* 2019;17:61–9.
- Kourtis D, Magnusson ML, Smith F, *et al.* Spine height and disc height changes as the effect of hyperextension using stadiometry and MRI. *Iowa Orthop J* 2004;24:65–71.
- Karadimas EJ, Siddiqui M, Smith FW, *et al.* Positional MRI changes in supine versus sitting postures in patients with degenerative lumbar spine. *J Spinal Disord Tech* 2006;19:495–500.
- Dvorák J, Panjabi MM, Chang DG, *et al.* Functional radiographic diagnosis of the lumbar spine. flexion-extension and lateral bending. *Spine* 1991;16:562–71.
- Allaire BT, DePaolis Kaluza MC, Bruno AG, *et al.* Evaluation of a new approach to compute intervertebral disc height measurements from lateral radiographic views of the spine. *Eur Spine J* 2017;26:167–72.
- Sobczak S, Dugailly P-M, Gilbert KK, *et al.* Reliability and validation of in vitro lumbar spine height measurements using musculoskeletal ultrasound: a preliminary investigation. *J Back Musculoskelet Rehabil* 2016;29:171–82.
- Madan SS, Rai A, Harley JM. Interobserver error in interpretation of the radiographs for degeneration of the lumbar spine. *Iowa Orthop J* 2003;23:51–6.
- Fazel R, Krumholz HM, Wang Y, *et al.* Exposure to low-dose ionizing radiation from medical imaging procedures. *N Engl J Med* 2009;361:849–57.
- Chleboun GS, Amway MJ, Hill JG, *et al.* Measurement of segmental lumbar spine flexion and extension using ultrasound imaging. *J Orthop Sports Phys Ther* 2012;42:880–5.
- Ledsome JR, Lessoway V, Susak LE, *et al.* Diurnal changes in lumbar intervertebral distance, measured using ultrasound. *Spine* 1996;21:1671–5.
- Tozawa R, Katoh M, Aramaki H, *et al.* Absolute and relative reliability of lumbar interspinous process ultrasound imaging measurements. *J Phys Ther Sci* 2016;28:2210–3.



- 19 Dill T. Contraindications to magnetic resonance imaging: non-invasive imaging. *Heart* 2008;94:943–8.
- 20 Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. *J Magn Reson Imaging* 2015;42:887–901.
- 21 Abdollah V, Parent EC, Battié MC. Reliability and validity of lumbar disc height quantification methods using magnetic resonance images. *Biomed Tech* 2019;64:111–7.
- 22 Glüer CC, Blake G, Lu Y, *et al.* Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* 1995;5:262–70.
- 23 Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg* 2018;126:1763–8.
- 24 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
- 25 Zhou SH, McCarthy ID, McGregor AH, *et al.* Geometrical dimensions of the lower lumbar vertebrae—analysis of data from digitised CT images. *Eur Spine J* 2000;9:242–8.
- 26 Johnson AW, Adams L, Kho JB, *et al.* Extended field-of-view ultrasound imaging is reliable for measuring Transversus abdominis muscle size at rest and during contraction. *BMC Musculoskelet Disord* 2021;22:282.
- 27 Tozawa R, Katoh M, Aramaki H, *et al.* Reliability and validity of an ultrasound-based imaging method for measuring interspinous process distance in the lumbar spine using two different index points. *J Phys Ther Sci* 2015;27:2333–6.

Supplemental material figure 1. Bland-Altman plots for significant correlations**A****B**