THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Real-time three-dimensional echocardiography for left atrial volume assessment in Thoroughbred racehorses: observer variability and comparison with two-dimensional echocardiography

Citation for published version:

Worsman, F, Miller, Z, Shaw, D, Blissitt, K & Keen, J 2020, 'Real-time three-dimensional echocardiography for left atrial volume assessment in Thoroughbred racehorses: observer variability and comparison with two-dimensional echocardiography', *Equine Veterinary Journal*. <https://doi.org/10.1111/evj.13408>

Digital Object Identifier (DOI):

[10.1111/evj.13408](https://doi.org/10.1111/evj.13408)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Equine Veterinary Journal

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

1

2 DR. FRANCESCA WORSMAN (Orcid ID : 0000-0002-7273-9846)

3 DR. JOHN A KEEN (Orcid ID : 0000-0002-7862-5321)

4

5

6 Article type : General Article

7

8

9 **Real-time three-dimensional echocardiography for left atrial volume assessment in**
10 **Thoroughbred racehorses: observer variability and comparison with two-dimensional**
11 **echocardiography**

12 F. C. F. Worsman*, Z. Miller, D. J. Shaw, K. J. Blissitt and J. A. Keen

13 The Royal (Dick) School of Veterinary Studies, University of Edinburgh, Roslin, Midlothian EH25
14 9RG.

15

16 *Corresponding author email: fworsman@exseed.ed.ac.uk

17 **Keywords:** horse, equine left atrial volume, 3DE, intra-class correlation coefficient

18

19 **Summary**

20 **Background:** Left atrial size predicts cardiac morbidity and mortality in humans and dogs. Real-
21 time three-dimensional echocardiography (3DE) may be reliable for assessing left atrial volume
22 (LAV) in horses.

23 **Objectives:** To determine intra and inter-observer variability estimates of 3DE-LAV and compare
24 it to that of 2DE-LAV estimates.

25 **Study design:** Method comparison.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/EVJ.13408](#)

This article is protected by copyright. All rights reserved

Methods: 3DE datasets were obtained from 40 horses, then graded for quality, creating a final study population of 22 horses. The 3DE and 2DE maximum LAV (LAV_{max}) and minimum LAV (LAV_{min}) were measured, and left atrial emptying volume (LA EV) and left atrial ejection fraction (LA EF) calculated, from the same 3D dataset on four occasions using a) a semi-automatic surface recognition algorithm and b) a modified Simpson's method of discs. 3DE LAV measurements were repeated by a second observer.

Results: For 3DE, median LAV_{max} was 596cm^3 for observer one, and 852cm^3 for observer two, LAV_{min} was 373cm^3 for observer one and 533cm^3 for observer two. Low intra-observer measurement variation was observed for LAV_{max} and LAV_{min} , with horse-level intra-class correlation coefficients (ICC_{horse}) for both observers between 76-85% (horse added as random effect). The inter-observer ICC was 58% for LAV_{max} and 50% for LAV_{min} on averaged measurements (with observer added as random effect), indicating consistent differences between observers. While intra-observer variation was similar for 2DE LAV_{max} measurements, it was greater for LAV_{min} ($ICC_{horse} = 67\%$). The inter-method ICC for 3DE vs 2DE was low at 14% for LAV_{max} and $\sim 0\%$ for LAV_{min} , indicating less-consistent differences with method.

Main limitations: Small study population, low observer number, use of different imaging modalities (fundamental frequency and octave harmonics).

Conclusions: 3DE assessment of LAV was reliable, suggesting suitability for longitudinal evaluation of clinical cases. Clinicians should be aware of differences in LAV measurements between observers. More defined measurement guidelines may improve repeatability.

Introduction

Accurate assessment of left atrial size is important and has been shown to predict morbidity and mortality in human and small animal veterinary cardiology [1, 2, 3, 4, 5]. In horses, accurate and repeatable measurement of the left atrium is of particular interest for quantifying volume overload and atrial systolic function in mitral valve regurgitation [6] and atrial fibrillation [7]. Left atrial size is routinely measured as part of the equine echocardiographic examination [6, 8], but measurements have been typically limited to linear dimensions from 2DE datasets, such as the left atrial diameter in a left or right parasternal long-axis view [9, 10] or the right parasternal short axis view. Linear dimensions can be misleading when assessing left atrial size, as the atrium has a complex three-dimensional topography [11] and, when enlarged, linear measurements may not reflect all dimensions of expansion [6]. More recently, 2DE area measurements have been

described and used to calculate functional changes in left atrial size [12]. These images may also be used to estimate volume [12], but such techniques also rely on geometric assumptions. In addition, linear and area measurements by 2DE may be inaccurate owing to foreshortening of the image, which can be difficult to determine. Volume measurement using real-time three-dimensional echocardiography (3DE) makes no such geometric assumptions while mitigating against foreshortening and thus may provide a more accurate and repeatable technique for assessing left atrial volume [13,14,15,16] providing the whole atrium is included in the 3DE dataset. The aim of this study was to assess the repeatability and reproducibility of 3DE to assess left atrial volume estimates in horses, using a software algorithm designed for volume measurement of the left ventricle, and to compare this to estimated volumes calculated from 2DE measurements.

Materials and methods

3DE datasets were obtained from 40 healthy, athletic Thoroughbred horses each examined once. All horses were examined by cardiac auscultation and horses with murmurs higher than grade 2/6 were excluded from the study, as were horses with a heart rate greater than 48bpm. Images were obtained with an ultrasound scanner (Vivid E9¹) using an active matrix 4D volume phased array transducer (4V-D¹) at a frequency of 1.7/3.3 MHz with (*harmonic imaging*) and 2.5MHz without (*fundamental frequency*) octave harmonics. A single operator (*obs*₂) acquired all the images prospectively for the purpose of this study. A single-lead electrocardiogram (ECG) was recorded simultaneously. All images of the left atrium, single cycle multi-beat compilation of 4 cardiac cycles, were obtained from a modified standard right parasternal long-axis view, focusing on the left atrium. Significant time was spent acquiring good quality images ensuring the entire left atrium was incorporated with a low number of stitch artefacts whilst the horse had a low resting heart rate.

Observers analysed the studies offline, retrospectively with analysis software (EchoPac PC version 202²). The images were graded to select only good quality left atrial datasets for analysis (Supplementary Item 1). Images were included if the volumes per second was greater than 15 and all of the atrium was included within the frames. Consequently, by using the grading system, for each horse datasets with the best quality image and highest frame rate were chosen. The same multi-beat compiled cardiac cycle was used for both 2DE and 3DE measurements. Random-generated order, blinded measurements were obtained on 4 occasions by two observers, one internal medicine resident with further cardiology training (*obs*₁) and one

91 European boarded internal medicine Diplomate (*obs₂*) for 3DE; and a single observer (*obs₁*) for
92 2DE images. Maximum and minimum LAV were measured.

93 Maximum 3DE left atrial volume (3DE LAV_{max}) was determined to be at the last frame before the
94 mitral valve opened and minimum 3DE left atrial volume (3DE LAV_{min}) was determined to be at
95 the last frame before the mitral valve closed [17]. The 3DE LAV was measured using a semi-
96 automated algorithm software (4D Auto LVQ in Manual Function; EchoPAC v. 202²), designed for
97 assessing left ventricular volume. This software was used, because at the time of this study,
98 software designed for 3DE atrial volume measurement was not available. The manual option of
99 4D Auto LVQ was chosen, because a pilot study (unpublished data) revealed that, the fully-
100 automated measurement package (4D Auto LVQ in Auto Function; EchoPAC v. 202²) was
101 unreliable at detecting the equine atrial wall and required significant endocardial marker
102 adjustment.

103 The three left atrial slices (60° to the long axis), derived from the 3DE dataset, were re-aligned in
104 the software to orientate the dorsal atrium uppermost, consistent with the apical orientation used
105 to measure ventricular volume in human medicine. The long axis four chamber view was aligned
106 first to allow the corresponding intersection line of all planes to be placed in the middle of the
107 atrial cavity, crossing the theoretical apex (dorsal atrium) and the centre of the mitral valve
108 opening in each view [18]. Aligning one plane automatically changed the other planes. The
109 alignment of the other views was manually refined. These initial steps were completed for
110 maximum atrial volume and did not need to be repeated to acquire minimum atrial volume which
111 was taken to be at the end of ventricular diastole and is automatically detected by the software,
112 the point at which it expects there to be maximum ventricular volume.

113 For maximum left atrial volume, the *Move ED²* button was therefore manually adjusted to the
114 correct frame before the mitral valve opened (Supplementary Item 2). For minimum atrial volume
115 the *Move ES²* button was similarly adjusted to the correct frame before the mitral valve closed.
116 The subsequent steps were performed to acquire maximum left atrial volume and repeated to
117 acquire minimum left atrial volume. At each time point, two electronic markers were placed on the
118 endocardium at the hinge points either side of the mitral valve annulus and one was placed at the
119 theoretical apex (dorsal aspect of the atrium) on all three 60° long axis slices. Positioning of the
120 markers was simultaneously shown on the other long axis slices as well as one short axis slice
121 for further guidance. The software then automatically generated a volume by endocardial border
122 recognition in six atrial views. The LAV was then minimally adjusted manually to ensure accurate
123 tracking and the final LAV was generated. The pulmonary veins and left atrial appendage were
124 not included. The software then generated a dynamic surface rendered left atrial cast and the

125 resulting volume-time plot (Figure 1) was examined to confirm that it demonstrated normal atrial
126 filling and bi-phasic atrial emptying pattern.

127 For 2DE analysis, LAV was measured by a modified Simpson's method of discs (2DE Volume
128 Biplane function EchoPAC v 202²) on the same single cycle-multi beat image. The atrium was
129 measured in a single plane (corresponding to a right parasternal long axis 4 chamber image) for
130 maximum volume by first positioning two markers on the endocardium either side of the mitral
131 valve annulus then positioning a third marker bisecting the dorsal wall. The endocardial border
132 was then traced by the operator using the software and, following manually marking the length of
133 the chamber, a volume comprising multiple stacked discs was automatically generated.
134 Manual adjustments after automated tracing were not performed in this technique. The method
135 was repeated for measuring the minimum volume. The pulmonary veins and left atrial appendage
136 were not included.

137 For both observers for 3DE and for obs_1 for 2DE, left atrial emptying volumes (LA EV) and left
138 atrial ejection fractions (LA EF) were also calculated. The LA EV was calculated by subtracting
139 the minimum LAV (LAV_{min}) from the maximum LAV (LAV_{max}), the LA EF was calculated by
140 dividing the LA EV by the LAV_{max} .

141

142 Data analysis

143 All statistical analyses were carried out in R (v 4.0.0, © 2020 The R Foundation for Statistical
144 Computing)³. All analyses were carried out on combined and separated fundamental frequency
145 and harmonic imaging data. Repeatability was defined as the ability of the same single observer
146 to obtain a same/similar result on a repeated measurement performed on the same sample.
147 Reproducibility was defined as the ability of different observers to come up with a same/similar
148 measurement. Intra-observer variability in both 3DE (in both observers, obs_1 and obs_2) and 2DE
149 (obs_1) measurements for LAV_{max} and LAV_{min} , LA EV and LA EF were assessed via random effect
150 linear models. Which horse the four repeated measures had come from for each
151 observer/method, was entered as the random effect, and the associated horse-level intra-class
152 correlation coefficients (ICCs) calculated (using the *lmer* and *sjPlot* packages³) (Figure 2). A high
153 horse-associated ICC (ICC_{horse}) was taken to indicate low intra-observer variability i.e. the
154 majority of the variability was attributable to the horse. For horse-associated ICCs intra-observer
155 repeatability and reproducibility were determined to be excellent if >90%, good if >75%, moderate
156 if between 74% and 50% and poor if <50% [19]. Also, as well as determining the standard
157 deviation of the four repeated measures per horse, a repeatability coefficient, the absolute value

below which the difference between the measurements will lie with 95% probability was determined. A Repeatability Coefficient (RC) was calculated as $\sqrt{2} \times 1.96 \times$ within-subject standard deviation [20].

Inter-observer variability in 3DE LAV_{max} and 3DE LAV_{min} were also assessed using ICC ($ICC_{inter-observer}$). The four repeated-measures per observer/method were averaged and random-effect linear models on mean 3DE LAV_{max} and 3DE LAV_{min}, 3DE LA EV and 3DE LA EF data, with observer added as the random effect were carried out (Figure 2). A high $ICC_{inter-observer}$ was taken to indicate high variability between observers i.e. the minority of the variability was attributable to the horse. Inter-method variability (3DE vs. 2DE) in obs_1 were also assessed via ICC ($ICC_{inter-method}$), with method added as the random effect and a high $ICC_{inter-method}$ was taken to indicate high variability in LAV between 3DE/2DE methods. Therefore, the $ICC_{inter-method}$ compares the correlation among 4 mean observations from each of the 22 horses by 2DE and 3DE methods, by observer 1 (Figure 2). For observer and method associated ICCs observer and method repeatability and reproducibility were determined to be excellent if <10%, good if <25%, moderate if between 26% and 50% and poor if >50%. Agreement between observers in 3DE measurements and for 3DE versus 2DE for obs_1 were also assessed using Bland Altman analyses (using the *BlandAltmanLeh* package³).

The relationship between both LAV_{max} and LAV_{min} in the different observers and methods of measurement, were assessed by linear regression. For each horse a single mean 3DE LAV_{max} and 3DE LAV_{min} value for each observer was calculated and then the difference between the observers was calculated, this difference was then regressed against horse bodyweight.

Results

Images selected for further analysis

After grading the image quality, 22 images were used for further analysis, from 21 geldings and 1 mare aged 4-9 years with a bodyweight range from 411-534kg (mean 472kg). The median heart rate was 35bpm (range 27-48bpm). Applying the grading system to all images obtained from all horses resulted in selection of a single image, either obtained by fundamental frequency or harmonic imaging, from each horse to be measured. Twelve selected images were obtained using harmonic imaging, 10 using fundamental frequency. The image volumes per second median for all images was 29.3 (range 15.8-42.1VPS); for harmonic imaging 24.3 (range 15.8-40.5VPS), and for fundamental frequency imaging was 37.8 (range 20.9-42.1VPS). Plots of

190 measurements obtained each day were assessed and evaluation of LAV over the 4 days by both
191 observers revealed no bias from day of measurement (Supplementary Item 3).

192 3DE volumes - intra-observer variability

193 The left ventricular quantification package was a feasible technique for left atrial volume
194 measurement. The overall combined fundamental frequency and harmonic imaging median 3DE
195 LAV_{max} for *obs*₁ was lower compared to *obs*₂ (596cm³ versus 852cm³ respectively, Table 1), and
196 this was also observed in 19 of the 22 horses (Figure 3a). Within observers, there was low intra-
197 observer variability, as assessed by the $ICC_{horse} \geq 76\%$ for both observers (Table 1), and the
198 average standard deviation of the four repeated-measures per horse was 56.6cm³ for *obs*₁ and
199 74.0cm³ for *obs*₂, giving RC's of 156.8cm³ and 205cm³ respectively. There was variation between
200 the observers based on the acquisition method, with the median difference between observers for
201 the data acquired via fundamental frequency lower (*obs*₁ = 635.5cm³, *obs*₂ = 788.5cm³, difference
202 = 153cm³, Table 2a, Figure 4a) compared to the harmonic imaging data (*obs*₁ = 571.0, *obs*₂ =
203 893.5, difference = 322.5cm³, Table 2b, Figure 4b). Intra-observer variability in the acquisition
204 method split data were qualitatively very similar (ICC_{horse} 69 to 85%, Table 2a,b) to the combined
205 data.

206 The overall median 3DE LAV_{min} measurements for the two observers gave a similar bias as the
207 3DE LAV_{max}, with median 3DE LAV_{min} for *obs*₁ lower, than *obs*₂ (Table 1), and this was observed
208 in 18 of the 22 horses (Figure 3b). Within observers, there was even lower intra-observer
209 variability in the 3DE LAV_{min} data, with $ICC_{horse} \geq 84\%$ for both observers (Table 1), and the
210 average standard deviation of the four repeated-measures per horse was 37.7cm³ for *obs*₁ and
211 38.5cm³ for *obs*₂, the RC for *obs*₁ was 104.4cm³ and for *obs*₂ was 106.6cm³. There was less
212 variation between the observers based on the acquisition method, but the median difference
213 between observers for the data acquired via fundamental frequency were still lower (Table 2a,
214 Figure 4c) compared to the harmonic imaging data (Table 2b, Figure 4d). Intra-observer
215 variability in the acquisition method split data were also qualitatively very similar (Table 2a,b) to
216 the combined data.

217 Left atrial emptying volumes (3DE LA EV) and ejection fractions (3DE LA EF) were also
218 calculated to compare between observers given *obs*₂ had a bias to measure larger volumes
219 irrespective of acquisition method (Tables 1, 2a,b). Overall median 3DE LA EV and 3DE LA EF
220 were greater in *obs*₂ compared to *obs*₁ (Table 1), with similar differences observed in the
221 harmonic imaging data (Table 2b). Whilst 3DE LA EV calculated from fundamental frequency
222 images gave a qualitatively similar result, the median 3DE LA EF of *obs*₁ was higher than that of

223 obs_2 (Table 2a). Intra-observer variability was predominately slightly greater for both observers for
224 both 3DE LA EV - with lower ICC_{horse} for combined data (Table 1) and data split into acquired via
225 fundamental frequency (Table 2a) or harmonic imaging (Table 2b)

226 3DE volumes - inter-observer variability

227 The $ICC_{inter-observer}$ for combined 3DE data was 58% for LAV_{max} and 50% for LAV_{min} (Table 1),
228 indicating consistent differences between observers measuring images from the same horse.
229 Bland-Altman analyses revealed that on average there were mean differences of 256cm³ for 3DE
230 LAV_{max} , and 154cm³ for 3DE LAV_{min} (Table 3). There was considerable variation between
231 observers for individual horses in both 3DE LAV_{max} and 3DE LAV_{min} (Table 3, Supplementary
232 Item 4). These differences were not related to bodyweight, as there was no statistically significant
233 relationship with the observed differences between observers in both 3DE LAV_{max} ($P = 0.156$, R^2
234 $= 10.2\%$) and 3DE LAV_{min} ($P = 0.069$, $R^2 = 16.3\%$) (Figure 5).

235 Fundamental frequency acquired data had a lower 3DE LAV_{max} and 3DE LAV_{min} $ICC_{inter-observer}$
236 ($\leq 35\%$, Table 3a), indicating less consistent differences between observers (Figure 4c), but mean
237 differences were 197cm³ for 3DE LAV_{max} and 108 for 3DE LAV_{min} , and maximum differences
238 were 764cm³ and 466cm³, respectively (data not shown). In contrast, data acquired using
239 harmonic images had a higher 3DE LAV_{max} and 3DE LAV_{min} $ICC_{inter-observer}$ (Table 3b), indicating
240 more consistent differences between observers (Figure 4d), with mean differences of 304cm³ and
241 193cm³, respectively, but still considerable variation in differences (maximum = 577cm³ and
242 392cm³, respectively, data not shown).

243 The $ICC_{inter-observer}$ for combined 3DE LA EV data was much lower (27%) than for 3DE LAV_{max} and
244 3DE LAV_{min} (Table 1), indicating less consistent differences between the two observers.
245 However, mean differences were 101 cm³, with wide variation in differences between horses
246 (Table 3, Supplementary Item 4). This relatively lower level of ICC in 3DE LA EV was replicated if
247 the data was split by acquisition method (Table 2a,b), and Bland-Altman analyses showing mean
248 differences of 88-112cm³ for the two methods (data not shown).

249 The $ICC_{inter-observer}$ for 3DE LA EF data whether combined, or divided into fundamental frequency
250 or harmonic images was ~0%, indicating effectively no consistent differences between observers
251 (Tables 1, 2a,b). This is reflected in the mean difference 3DE LA EV of <1% for both combined
252 data (Table 3), and fundamental frequency or harmonic images. However, there was still variation
253 in differences in 3DE LA EF between observers (maximum = 27%, Supplementary Item 4).

254 2DE volumes – intra-observer variability

255 The obs_1 overall median 2DE LAV_{max} was higher at 687.5cm^3 compared to the 3DE
256 measurements (596cm^3 , Table 1), but this was only observed in 11 of the 22 horses (Figure 6a).
257 As with obs_1 3DE measurements, there was low intra-observer variability for the 2DE ($ICC_{horse} =$
258 75%), and the average standard deviation of the four repeated-measures per horse was 63.7cm^3
259 with a RC of 176.5cm^3 . However, the differences calculated by obs_1 for 3DE and 2DE were very
260 dependent on the acquisition method, with the median difference between data acquired via
261 fundamental frequency being 1cm^3 (Table 2a), compared to a 132cm^3 for the harmonic imaging
262 data (Table 2b). Intra-observer variability in the acquisition method split data were qualitatively
263 very similar (ICC_{horse} 66-77%, Table 2a,b) to the combined data.

264 Overall median 2DE LAV_{min} was also higher at 392.5cm^3 compared to the 3DE measurements
265 (373cm^3 , Table 1), but again this was not in all horses, being only observed in 10 (Figure 6b).
266 Intra-observer variability was low ($ICC_{horse} = 67\%$), and the average standard deviation of the four
267 repeated-measures per horse was 59.4cm^3 , giving a RC of 164.5cm^3 . Differences between 3DE
268 and 2DE calculated were less dependent on the acquisition method, with median differences in
269 data acquired via fundamental frequency being 11cm^3 (Table 2a), compared to a 29cm^3 for the
270 harmonic imaging data (Table 2b). Intra-observer variability in the acquisition method split data
271 were qualitatively very similar (ICC_{horse} 60-64%, Table 2a,b) to the combined data.

272 Median 2DE LA EV and 2DE LA EF were also higher compared to the 3DE measurements in
273 combined data (Table 1) and data split by acquisition method (Table 2a,b). However, intra-
274 observer variability was much higher in 2DE LA EV and 2DE LA EF compared to 3DE
275 measurements (ICC_{horse} 46 and 54%, respectively, Table 1), with this again reflected in the data
276 split by acquisition method (Table 2a,b).

277

278 3DE vs 2DE volumes – inter-method variability

279 The $ICC_{inter-method}$ from comparing 3DE and 2DE was $\sim 14\%$ for LAV_{max} and $\sim 0\%$ for LAV_{min} (Table
280 1), indicating very little or no consistent differences between 3DE and 2DE as measured by obs_1 .
281 LAV_{max} $ICC_{inter-method}$ dropped to $\sim 0\%$ for fundamental frequency, but rose to 33% for harmonic
282 imaging data (Table 2a,b), whereas for LAV_{min} there was no change in the $ICC_{inter-method}$ for
283 fundamental frequency data, and a smaller increase in harmonic image data. The smaller ICC_{inter-}
284 $method$ obtained are reflected in the relatively small mean differences observed from the Bland-
285 Altman analyses ($<83\text{cm}^3$, Table 3). However, there was still considerable variation between
286 methods for individual horses in both LAV_{max} and LAV_{min} (Table 3, Supplementary Item 4).

287 The $ICC_{inter-method}$ for combined LA EV data was similar to that of LAV_{max} (17%, Table 1), again
288 indicating less consistent differences between the two methods. Mean differences were also
289 lower 64cm^3 (Table 3) but there was still wide variation in differences between horses (Table 3,
290 Supplementary Item 4). These LA EV $ICC_{inter-method}$ values were replicated if the data was split by
291 acquisition method (Table 2a,b). The $ICC_{inter-method}$ obtained for 3DE vs 2DE for LA EF, whether
292 combined, or divided into fundamental frequency or harmonic images was low ($\leq 12\%$), again
293 indicating very low or no consistent differences between methods (Tables 1, 2a,b). This was
294 reflected in the mean difference LA EV of 5% (Table 3), but there was still variation in differences
295 in LA EF between methods (maximum = 38%, Supplementary Item 4).

296

297 Discussion

298 This study demonstrates that assessment of left atrial volume using software designed for
299 measuring the left ventricle is feasible with good intra-observer repeatability, reflected in the high
300 intra observer ICC, and better intra-observer repeatability compared to 2DE estimates. The
301 average standard deviation from repeated measurements and repeatability coefficients per horse
302 were low, supporting the good intra-observer repeatability. This will be of value when assessing
303 progression of disease in horses, where progressive left atrial dilatation secondary to mitral
304 regurgitation reduces the left atrial emptying fraction and increases left atrial pressure and
305 contractile dysfunction [21]. Left atrial enlargement is also a known risk factor for atrial fibrillation
306 [12, 23, 23] and for recurrence of fibrillation following conversion to sinus rhythm [24]. Reliable
307 measurement of left atrial size also has clear diagnostic and prognostic value in humans [25] and
308 dogs [16].

309 In humans, measurement of left atrial volume provides a better assessment of left atrial size [14,
310 15, 21] than linear dimensions which may result in under- or over-estimation of the true diameter
311 of the atrium and may not accurately represent left atrial enlargement in all planes. Cardiac MRI
312 is the gold standard method for assessing left atrial volume but neither this technology, nor
313 cardiac CT, are currently feasible in horses. In this study 3DE LAV was compared to that
314 calculated from a single 2DE image using the single plane method of discs. The AAE/EAE
315 guidelines for estimating LAV from 2DE datasets in humans, recommends acquiring apical views
316 and using either an ellipsoid area-length model or a disc summation algorithm (method of discs)
317 [17] as used in this study. The ellipsoid model is considered less accurate because of reliance
318 and limitations of the linear measurements required to carry out this calculation and the disc
319 summation algorithm is recommended where possible for measuring LAV using 2DE datasets in

humans and dogs [16, 17, 26]. The recommended apical approach however, is not possible in horses and 2DE LAV is rarely calculated due to the perceived implications for repeatability of a single parasternal image. Nevertheless, the single plane method of discs was used in this study, calculated from a single 2DE standard right parasternal image orientated to maximise the left atrium. A single plane technique was chosen for 2DE measurements because we wanted to compare 3DE to 2DE with a conventional, standard equine 2DE view and the right parasternal view is currently the most used. There is evidence in human medicine that single plane echocardiography is comparable to biplane 2DE in the evaluation of LAV [27].

Repeat measurement of 2DE images showed more variation in LAV_{min} values than 3DE in the horses in this study. In human cardiology 3DE LAV measurements performed better in terms of correlation, bias, limits of agreement and inter-observer variability [14, 15] and had superior prognostic ability. While reasonable estimates of LAV can be made using 2DE, the inherent geometric assumptions limit the accuracy of such measures [12], particularly in disease where the assumptions may fail. In contrast, 3DE does not make the same assumptions and by encompassing the entire left atrium in the three-dimensional data set, this technique can potentially provide a more accurate assessment of LAV while also reducing variability in different image acquisition planes acquired between different observers. However, despite this theoretical advantage, we cannot comment on variability in 3DE acquisition between observers as only one observer acquired the dataset in this study. The 2DE images from the 3DE dataset are likely to have had lower spatial and temporal resolution compared to images acquired from a 2DE probe which may have contributed to some variation. Nevertheless, the variability noted in this study is consistent with that observed in a study using 2DE datasets for assessing LAV [12].

The left ventricular quantification software (4D Auto LVQ; EchoPAC v. 2022) used in this study provided a simple, feasible technique to measure LAV. The software and algorithms, designed to measure the left ventricle, were flexible enough to trace the inner side of the left atrial endocardium throughout the cardiac cycle. The main difficulty in using the 4D LVQ package for LAV measurement was that the software expected the maximum volume at end diastole and for this to correspond with the ECG trace. Because atrial volume was being assessed, volumes were largest at end systole and this generated a problem for the algorithms. As a consequence, a technique was adopted to adjust the end diastolic point to the end of atrial diastole, the point at which the largest volume was obtained when the mitral valve was about to open. This adjustment was also required for the minimum LAV. Automated graphs were subjectively assessed for filling and biphasic emptying; normal patterns were observed for all atria included. A pilot study (unpublished data) revealed that if the *Move ED²* and *Move ES²* buttons were not manually

adjusted for each measurement then the automated graph would be inaccurate. During this study, left atrial quantification software for this software package became available, however preliminary evaluation indicated the algorithm was not applicable for measuring the equine left atrium. The reason for this is not clear without more information on the assumptions made by the measurement algorithm.

Of the 40 datasets available, 18 were excluded from the study, because of either a volume rate <15 per second, some of the atrial wall not contained within the data set or a combination of both. At the time of collection, the operators were relatively novice with 3DE acquisition techniques and the large rejection rate in this study emphasises the need to acquire as good quality datasets as possible for analysis. Such good quality acquisition is a significant challenge with 3DE and the operator took several datasets to get the best multi-beat acquisition for analysis, yet there was still considerable rejection of images at subsequent image grading. The LVQ software instructs a minimum of 12 volumes per second for analysis however 15 or greater was used as a lower limit in this study; in most cases, this was significantly higher with a median of 29.3 VPS (range 15.8-42.1VPS). Direct comparison between fundamental and harmonic images in this study was not possible as images were not matched; i.e. images were from different horses. Comparisons on the imaging modality and its effect on measurement repeatability are therefore limited as bias may have been introduced. The horses were not assigned into groups, but categorised purely by initial image quality. At initial acquisition, fundamental frequency and harmonic imaging images were recorded in all horses. Then analysis using the grading system resulted in only single images of either type of imaging modality to be acceptable for measurement. Further study should prospectively compare imaging modalities on 3DE LAV measurement repeatability and reproducibility,

The fit athletic horses in this study were of a single breed from which good quality images could be obtained, which is a limitation of this study. The results of this study may therefore not hold true for datasets from breeds where image quality is poorer. Further study with a larger population of mixed breeds would be required to evaluate the effects of different breeds on image quality and observer repeatability and reproducibility. The horses were imaged unsedated on a yard so that the setting would correspond with a practical clinical situation. The LAV range (Figure 3a,b) was wide which is unsurprising given that this population of athletic horses had different weights, ages [6, 28] and were at different stages of training [29]. It is worth noting also that minimal changes in left atrial diameter (and therefore radius) will have a significant impact on volume measurements, creating a wider range than that observed with linear measurements.

387 Given the expected range of left atrial diameters of the horses contained in this study, the results
388 for LAV in this study are considered realistic.

389 In echocardiography, measurements are typically taken over at least three cardiac cycles to
390 provide an average and decrease the variability associated with frame selection. In this study, the
391 multi-beat acquisition, a composite of 4 successive cycles was used. The multi-beat acquisition
392 offers the advantage of enabling the use of harmonic imaging and thus improving image quality
393 for this relatively large volume. Observers measured the same dataset and multi-beat cycle so
394 minimising variation in image acquisition. It may be that greater variation is induced by
395 acquisition between observers although this was not measured in this study. In theory however,
396 3DE should mitigate against the differences in acquisition between observers, since all image
397 planes are sampled simultaneously. Efforts should be made to acquire the best LAV dataset
398 from a modified parasternal four chamber view, which can be challenging. Of further
399 consideration, the time spent measuring is significantly greater than that of single linear or area
400 measurements; to reproduce a realistic clinical situation therefore, the authors felt that
401 establishing variability from one good measurement was optimal. Single beat acquisition is
402 possible but the image quality and frame rate suffer significantly: both these factors may
403 introduce poorer accuracy and greater observer variability. In the case of arrhythmias such as
404 atrial fibrillation, multi-beat acquisition is not possible. Further studies should therefore assess
405 the repeatability of single beat measurements rather than multi-beat ones.

406 In humans, when multi-beat images are acquired, subjects are asked to breath hold, to limit the
407 variability attributable to respiratory motion. In a canine study [16], dogs were anaesthetised
408 allowing operator control of breathing. Neither techniques were feasible in this study, nor
409 clinically applicable, and minor artefacts, due to 'stitching' the multi-beat images, have to be
410 accepted in horses. As long as these were subjectively minor, this was deemed acceptable. The
411 lower resting respiratory rate of horses (approximate maximum 33% of the heart rate) helps in
412 this regard, when the multi-beat setting is four. In previous equine studies, allometric scaling has
413 allowed for correction of difference in atrial size associated with bodyweight [6], this was not
414 assessed in this study where the primary aim was to assess repeatability and reproducibility of
415 the technique. The impact of variables such as heart rate, sex, age, intensity of training, image or
416 variability throughout the breathing cycle were not analysed with respect to LAV. Future studies
417 that aim to establish reference ranges for LAV should provide normal values indexed to another
418 morphometric characteristic or intra-cardiac structures (e.g. the aorta).

419 The intra-class correlation coefficient was elected to be the most desirable way of determining
420 repeatability and reproducibility for intra-, inter-observer and inter-method measurements,
421 reflecting the degree of correlation, repeatability and reproducibility between results [19, 30, 31].
422 The same observers were repeatedly measuring exactly the same cardiac cycle. To reduce
423 observer bias, measurements were obtained in a random generated order to decrease error that
424 could result from sequential measuring, with observers unaware of previous measurements or
425 horse identification. In this study, repeatability of 3DE measures of LAV by a single person was
426 good, as evidenced by the high ICC for repeated measurements in horses in the same observer.
427 Calculated repeatability coefficients suggest that approximately 200cm³ change in volume at 3DE
428 LAV_{max} or a 100cm³ change in volume at 3DE LAV_{min} would suggest a significant enlargement.
429 Lower intra-observer variation was observed for minimum atrial volume compared to maximal
430 atrial volume (Table 1). Reasons for this are unknown but may include better image quality during
431 this stage of the cardiac cycle when the atrium is smaller or a more repeatable frame selection at
432 this cycle stage.

433 Sources of measurement variation using the 4D LVQ algorithm could have arisen from several
434 sources. Firstly, consistent determination of end systole and end diastole is crucial. Determining
435 the last frame before mitral valve opening relies on some degree of subjectivity, especially when
436 the frame rate is higher. The volume rate per second will have a significant impact on whether
437 the frame chosen is truly the end of diastole/systole. Furthermore, during image analysis,
438 accurate placement of markers at the hinge points of the mitral valve and atrial apex is important.
439 Following creation of the volume boundaries, manual border manipulation may also contribute to
440 variability. The sources of variation noted above are also fundamentally affected by the image
441 quality, determined by volume rate and line density, both of which are compromised when large
442 volumes are required. The effects of image quality were further evaluated by analysing variability
443 between images attained with the fundamental frequency from those obtained using harmonic
444 imaging, which is used to improve image resolution. Intra-observer 3DE variability for both
445 observers was still >75% when only the images obtained with fundamental frequency were
446 analysed. The average volume rate for the fundamental frequency images was higher than that
447 for harmonic imaging as expected, since harmonic imaging compromises temporal resolution.
448 When images obtained by fundamental frequency only were analysed, intra-observer variability
449 decreased (ICC_{horse} increased) for 3DE LAV_{max} for observer one and 3DE LAV_{min} for observer
450 two, and inter-observer and inter-method variability decreased ($ICC_{inter-observer}$ and $ICC_{inter-method}$
451 decreased) suggesting that a higher frame rate decreased measurement variability. In contrast
452 however, variability increased slightly with fundamental frequency for 3DE LAV_{min} for observer

one and 3DE LAV_{max} for observer two. Given the design of this study it is difficult to draw firm conclusions about whether fundamental frequency or harmonic imaging is preferred. As a general rule however, obtaining images with the highest resolution and frame rate for a given volume is preferable. Although basic criteria for measurement were set in advance for both observers, it is likely that refined measurement guidelines, being more prescriptive about precise cursor placement, timing of measurement at end systole and diastole, volume border demarcation at the mitral valve and criteria for border manipulation may decrease inter-observer variation. This emphasises the need for different observers to adhere to the same strict measurement guidelines. The study was also limited by the low observer number, further study would be indicated with refined measurement guidelines and more observers of differing expertise to evaluate 3DE and 2DE LAV measurement repeatability, reproducibility and guidelines.

Human studies have determined that 3DE LAV volumes are larger [13, 14, 15, 18, 32] compared to 2DE. In this study, the mean difference for 2DE volumes were greater than 3DE (Table 3) but 2DE did not give consistently higher values than 3DE. Although the accuracy of both techniques could not be established due to the lack of a gold standard method to compare LAV in horses, similarity with comparative studies in humans and dogs, would suggest that 3DE also offers a more accurate measurement of LAV in horses. Despite the low ICC between 2DE and 3DE measurements of LAV, the wide difference limits at individual horse level supports the suggestion that LAV obtained by different methods should not be used interchangeably, a conclusion also reached in a canine study [16]. Fundamental frequency datasets improved inter-method variability to ~0%, suggesting that image quality may be a contributing factor to the variability.

Conclusion

Assessment of LAV by 3DE is feasible and shows good intra-observer repeatability and moderate inter-observer reproducibility. Results using the 2DE method of discs for LAV assessment should not be used interchangeably with that of 3DE. Variability using both techniques for LAV assessment may improve with higher frame rates and refined measurement guidelines but three-dimensional techniques for assessing volume, if available, are likely to be preferable for longitudinal evaluation of left atrial volume due to their better repeatability.

Authors' declarations of interest

No competing interests have been declared.

Ethical animal research

485 The study protocol was approved by the Veterinary Ethic Research Committee, University of
486 Edinburgh.

487 **Informed consent**

488 Owners gave consent for their animals' inclusion in the study.

489 **Data accessibility statement**

490 The data that support the findings of this study are available from the corresponding author upon
491 reasonable request

492 **Source of funding**

493 The primary author's residency and the study funded by the Horserace Betting Levy Board.

494 **Authorship**

495 F. Worsman had full access to all the data in the study, takes responsibility for integrity of the
496 data and accurate data analysis, contributed to the study execution, data analysis and
497 interpretation, preparation of the manuscript and final approval of the manuscript. Z. Miller
498 contributed to study design, data analysis and interpretation and had final approval of the
499 manuscript. D. Shaw contributed to data analysis and interpretation, preparation and final
500 approval of the manuscript. K. Blissitt contributed to study design, study execution, preparation
501 and final approval of the manuscript. J. Keen contributed to study design, study execution, data
502 interpretation and preparation and final approval of the manuscript.

503

504

505 **Manufacturers' addresses**

506

- 507 1. 4V-D Active Matrix 4D Volume Phased Array Transducer 1.5-4.0MHz, Vivid E9, GE
508 Healthcare, GE Med Systems Ltd, Chalfont St Giles, Bucks, UK.
- 509 2. EchoPac PC version 202, GE Healthcare, Horten, Norway.
- 510 3. *lmer*, *sjPlot* and *BlandAltmanLeh* packages accessed <https://www.R-project.org> (2020) R
511 version 4.0.0, R Core Team, The R Foundation for Statistical Computing, Vienna, Austria.

512

513 **Figure Legends**

514 **Figure 1:** 4 planes of LAV from 3DE dataset (left) showing left atrial volume changes throughout
515 the cardiac cycle. On the right, the time volume curve is enlarged: note the filling to maximal
516 volume at end diastole (*ED*) then the passive and active emptying phases to minimal volume at
517 end systole (*ES*).

518
519 **Figure 2:** Flow chart illustrating study design for Intra-Observer, Inter-Observer and Inter-Method
520 data analysis.

521
522 **Figure 3:** Boxplots with individual values overlaid for *obs*₁ and *obs*₂ for a) 3DE LAV_{max} (cm³) b)
523 3DE LAV_{min} (cm³), n = 22 horses. Horizontal black lines indicate the median for that horse from
524 both observers.

525
526 **Figure 4:** Boxplots with individual values overlaid for *obs*₁ and *obs*₂ for (a,b) 3DE LAV_{max} (cm³)
527 and (c,d) LAV_{min} (cm³) divided into whether fundamental frequency, n = 10 horses (a,c) or
528 harmonic imaging n = 12 horses (b,c) were used to acquire the data.

529
530 **Figure 5:** Relationships between left atrial volumes and bodyweights.

531
532 **Figure 6:** Boxplots with individual values overlaid for *obs*₁ for 3DE and 2DE measured data for a)
533 LAV_{max} (cm³) and b) LAV_{min} (cm³), n = 22 horses.

534
535 **Supporting information**

536 **Supplementary Item 1:** 3DE LAV Image Grading System.

537 **Supplementary Item 2:** Video: Loop of 3DE LAV_{max}.

538 **Supplementary Item 3:** Day-to-day LAV measurement variation.

539 **Supplementary Item 4:** Bland Altman plots.

540

541 **References**

- 542 1. Borgarelli M., Santilli R.A., Chiavegato D., D'Angelo G., Zanatta R., Mannelli A. and
543 Tarducci A. (2006) Prognostic indicators for dogs with dilated cardiomyopathy. *J Vet*
544 *Intern Med* **20** 104-10.
- 545 2. Borgarelli M., Savarino, P., Crosara S., Santilli R.A., Chiavegato D., Poggi, M. Bellino, C.
546 La Rosa G., Zanatta R., Haggstrom J. and Tarducci A. (2008) Survival characteristics and
547 prognostic variables of dogs with mitral regurgitation attributable to myxomatous valve
548 disease. *J Vet intern Med* **22** 20-8.
- 549 3. Rossi A., Cicoira M., Zanolla L., Sandrini R., Golia G., Zardini P. and Enriquez-Sarano M.
550 (2002) Determinants and prognostic value of left atrial volume in patients with dilated
551 cardiomyopathy. *J Am Coll Cardiol* **40** 1425-30.
- 552 4. Sargent J., Muzzi R., Mukherjee R., Somarathne S., Schranz K., Stephenson H., Connolly
553 D., Brodbelt D. and Luis Fuentes V. (2015) Echocardiographic predictors of survival in
554 dogs with myxomatous mitral valve disease. *J Vet Cardiol* 2015 **17** 1-12.
- 555 5. Wu V.C-C., Takeuchi M., Kuwaki H., Iwataki M., Nagata Y., Otani K., Haruki N., Yoshitani
556 H., Tamura N., Abe H., Negishi K., Lin F. C. and Otsuji Y. (2013) Prognostic value of LA
557 volumes assessed by transthoracic 3D echocardiography. *J Am Coll Cardiol Imaging*
558 2013 **6** 1025-35.
- 559 6. Huesler I. M., Mitchell K. J. and Schwarzwald C. C. (2016) Echocardiographic assessment
560 of left atrial size and function in Warmblood horses: reference intervals, allometric scaling,
561 and agreement of different echocardiographic variables. *J Vet Intern Med* **30** 1241-1252.
- 562 7. De Clercq D., van Loon G., Tavernier R., Duchateau L. and Deprez P. (2008) Atrial and
563 ventricular electrical and contractile remodelling and reverse remodelling owing to short-
564 term pacing-induced atrial fibrillation in horses. *J Cardiovascular electrophysiol* **11** 773-
565 784.
- 566 8. Schwarzwald C. C. (2019) Equine Echocardiography. *Vet Clin North Am Equine Pract* **35**
567 (1) 43-64.
- 568 9. Reef V. B. (1998) Cardiovascular ultrasonography. In Reef V. B., ed. *Equine Diagnostic*
569 *Ultrasound*. Philadelphia, PA: WB Saunders 215-272
- 570 10. Marr C. M. and Patteson M. (2010) Echocardiography. In: Marr C. M., Bowen M., eds
571 *Cardiology of the horse*, 2nd ed. Edinburgh: Saunders Elsevier 105-126
- 572 11. Vandecasteele T., Cornillie P., van Steenkiste G., Vandevelde K., Gielen I.,
573 Vanderperren K. and van Loon G. (2018) Echocardiographic identification of atrial-related

- structures and vessels in horses validated by computed tomography of casted hearts. *Equine Vet J* **51** (1) 90-96
12. Schwarzwald C. C., Schober K. E. and Bonagura J. D. (2007) Methods and reliability of echocardiographic assessment of left atrial size and mechanical function in horses. *Am J Vet Res* **68** (7) 735-747
13. Perez de Isla L., Feltes G., Moreno J., Martinez W., Saltijeral A., de Agustin J. A., Gomez de Diego J. J., Maros-Alberca P., Luaces M., Ferreiros J., Garcia Fernandez M. A. and Macaya C. (2014) Quantification of left atrial volumes using three-dimensional wall motion tracking echocardiographic technology: comparison with cardiac magnetic resonance. *Eur Heart J – CV Imaging* **15** 793-799
14. Keller A. M., Gopal A. S. and Kind D. L. (2000) Left and right atrial volume by free hand three-dimensional echocardiography: in vivo validation using magnetic resonance imaging. *Eur J Echocardiogr* **1** (1) 55-65
15. Artang R., Migrino R. Q., Harmann L., Bowers M. and Woods T. D. (2009) Left atrial volume measurement with automated border detection by 3-dimensional echocardiography: comparison with magnetic resonance imaging. *Cardiovasc Ultrasound* **31** 7-16
16. Bouvard J., Thierry F., Culshaw G. J., Schwarz T., Handel I. and Martinez Pereira Y. (2019) Assessment of left atrial volume in dogs: comparisons of two-dimensional and real-time three-dimensional echocardiography with ECG-gated multidetector computed tomography angiography. *J Vet Cardiol* **24** 64-77
17. Lang R. M., Badano L. P., Mor-Avi V., Afilalo J., Armstrong A., Ernande L., Flachskampf F. A., Foster E., Goldstein S. A., Kuznetsova T., Lancellotti P., Muraru D., Picard M. H., Rietzschel E. R., Rudski L., Spencer K. T., Tsang W. and Voigt J. U. (2015) Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr* **28** (1) 1-39
18. Muraru D., Badano L. P., Piccoli G., Gianfagna P., Del Mestre L., Ermacora D. and Proclemer A. (2010) Validation of a novel automated border-detection algorithm for rapid and accurate quantitation of left ventricular volumes based on three-dimensional echocardiography. *Eur J Echocardiogr* **11** 359-368
19. Koo K. T. and Li M. Y. (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropractic Med* **15** 155-163

20. Bland M. An introduction to medical statistics. 3rd Ed. New York: Oxford University Press; 2000 p. 268-93
21. Pagel P. S., Kehl F., Gare M., Hettrick D. A., Kersten J. R. and Warltier D. C. (2003) Mechanical function of the left atrium: new insights based on analysis of pressure-volume relations and Doppler echocardiography. *Anaes* **98** (4) 975-94
22. Schwarzwald, C. C., Hamlin, R. L., Bonagura J. D., Nishijima Y., Meadows C. and Carnes C. A. (2007) Atrial, SA nodal electrophysiology in standing horses: normal findings and electrophysiologic effects of quinidine and diltiazem. *J Vet Intern Med* **21** 166-175
23. Reef V. B., Bonagura J. D., Buhl R., McGurrian M. K. J., Schwarzwald C. C., van Loon G. and Young L. E. (2014) Recommendations for management of equine athletes with cardiovascular abnormalities. *J Vet Intern Med* **28** 749-761
24. De Clercq D., Decloedt A., Sys S. U., Verheyen T., Van Der Vekens N. and Van Loon G. (2014) Atrial fibrillation cycle length and atrial size in horses with and without recurrence of atrial fibrillation after electrical cardioversion. *J Vet Intern Med* **28** (2) 624-629
25. Wu V. C-C. and Takeuchi M. (2017) Three-dimensional echocardiography: current status and real-life applications. *Acta Cardiol Sin* **33** 107-118
26. Wesselowski S., Borgarelli N., Bello N. M. and Abbott J. (2014) Discrepancies in identification of left atrial enlargement using left atrial volume versus left atrial-to-aortic root ratio in dogs. *J Vet Intern Med* **28** 1527-1533
27. Vieira-Filho, N.G., Mancuso, F.J.N., Oliveira, W.A.A., Gil, M. A., Fischer C.H., Moises, V.A. and Campos O. (2014) Simplified single plane echocardiography is comparable to conventional biplane two-dimensional echocardiology in the evaluation of left atrial volume: a study validated by three-dimensional echocardiography in 143 adults. *Echocardiogr* **31** (3) 265-272
28. Al-Haidar A, Farnir F., Deleuze S., Sandersen C. F., Leroux A. A., Borde L., Cerri S. and Amory H. (2013) Effect of breed, sex, age and body weight on echocardiographic measurements in the equine species. *Res Vet Sci* **95** 255-260
29. Buhl R, Ersbøll A. K., Eriksen L., Koch J. (2005) Changes over time in echocardiographic measurements in young Standardbred racehorses undergoing training and racing and association with racing performance. *J Am Vet Med Assoc* **226** 1881-1887
30. Popović Z. B. and Thomas J. D. (2017) Assessing observer variability: a user's guide. *Cardiovasc Diagn Ther* **3** 317-324
31. Young L. E. and Scott G. R. (1998) Measurement of cardiac function by transthoracic echocardiography: day to day variability and repeatability in normal Thoroughbred horses. *Equine Vet J* **30** (2) 117-122

32. Badano L. P., Miglioranza M. H., Mihăilă S., Peluso D., Xhaxho J., Marra M. P., Cucchini U., Soriani N., Iliceto S. and Muraru D. (2016) Left Atrial Volumes and Function by Three-Dimensional Echocardiography: Reference Values, Accuracy, Reproducibility, and Comparison With Two-Dimensional Echocardiographic Measurements. *Circ Cardiovasc Imaging* **9** (7) <https://doi.org/10.1161/CIRCIMAGING.115.004229>

Table 1: Summary measures (median and range) and Intra-class correlation coefficients (ICC_{horse} , $ICC_{inter-observer}$, $ICC_{inter-method}$) of LAV_{max} and LAV_{min} volume, maximum left atrial emptying volume and left atrial ejection fraction for raw data with 4 repeated measurements per horse per observer per method, and average data where the mean value for each set of 4 values is calculated (n = 22 horses).

Data	Volume	Method	Observer				$ICC_{inter-observer}$	obs ₁ $ICC_{inter-method}$
			obs_1 Median (range)	ICC_{horse}	obs_2 Median (range)	ICC_{horse}		
Raw (4 repeat measurements / horse)	LAV_{max} (cm ³)	3DE	596.0 (349-1029)	77%	852.0 (516-1471)	80%	-	-
		2DE	687.5 (440-1144)	75%	-	-	-	-
	LAV_{min} (cm ³)	3DE	373.0 (200-721)	84%	533 (314-815)	85%	-	-
		2DE	392.5 (212-776)	67%	-	-	-	-
	Left Atrial Emptying Volume _{max} (cm ³)	3DE	210.5 (36-542)	70%	311.5 (78-941)	75%	-	-
		2DE	289.0 (4-747)	46%	-	-	-	-
	Left Atrial Ejection Fraction (%)	3DE	33.2 (5.2-61.0)	76%	36.1 (13.6-64.0)	71%	-	-
		2DE	43.2 (0.5-67.2)	54%	-	-	-	-
Mean of the 4 repeat measurements / horse	LAV_{max} (cm ³)	3DE	590.4 (405.0-846.0)	-	832.7 (639.8-1214.5)	-	58%	14%
		2DE	689.5 (493.2-1018.8)	-	-	-	-	-
	LAV_{min} (cm ³)	3DE	383.4 (219.2-640.5)	-	528.1 (332.8-777.8)	-	50%	~0%
		2DE	406.0 (239.0-594.2)	-	-	-	-	-
	Left Atrial Emptying Volume _{max}	3DE	216.4 (90.8-430.0)	-	318.5 (108.5-742.8)	-	27%	17%
		2DE	285.3 (99.0-447.2)	-	-	-	-	-

	(cm ³)						
	Left Atrial Ejection Fraction (%)	3DE 2DE	35.3 (14.7-58.3) 41.5 (14.6-58.8)	- -	36.5 (14.9-61.0) -	- -	~0% 6%

Table 2: Summary measures (median and range) and Intra-class correlation coefficients (ICC_{horse} , $ICC_{inter-observer}$, $ICC_{inter-method}$) of LAV_{max} and min LAV volume, maximum left atrial emptying volume and left atrial ejection fraction for raw data with 4 repeated measurements per horse per observer per method, and average data where the mean value for each set of 4 values is calculated in two different methods of imaging a) fundamental frequency data, n = 10 horses and b) harmonic imaging data, n = 12 horses.

2a. Fundamental Frequency								
Data	Volume	Method	Observer				$ICC_{inter-observer}$	obs_1 $ICC_{inter-method}$
			obs_1 Median (range)	ICC_{horse}	obs_2 Median (range)	ICC_{horse}		
Raw (4 repeat measurements / horse)	LAV_{max} (cm ³)	3DE	635.5 (349-1029)	82%	788.5 (516-1471)	76%	-	-
		2DE	634.5 (440-839)	66%	-	-	-	-
	LAV_{min} (cm ³)	3DE	350.5 (210-721)	82%	481.5 (314-782)	88%	-	-
		2DE	339.5 (212-524)	64%	-	-	-	-
	Left Atrial Emptying Volume _{max} (cm ³)	3DE	247.5 (36-433)	61%	297.0 (78-941)	77%	-	-
		2DE	294.5 (58-522)	58%	-	-	-	-
	Left Atrial Ejection Fraction (%)	3DE	39.4 (5.2-58.0)	60%	36.1 (13.9-64.0)	74%	-	-
		2DE	46.2 (10-67.2)	57%	-	-	-	-
Mean of the 4 repeat measurements	LAV_{max} (cm ³)	3DE	617.9 (405.0-846.0)	-	758.5 (639.8-1212.5)	-	35%	~0%
		2DE	612.9 (493.2-768.0)	-	-	-	-	-
	LAV_{min}	3DE	345.8 (246.0-640.5)	-	473.4 (332.8-764.0)	-	24%	~0%

/ horse	(cm ³)	2DE	343.5 (239.0-459.2)	-	-	-	
	Left Atrial Emptying	3DE	236.5 (110.8-315.2)	-	300.3 (155.5-742.8)	-	12%
	Volume _{max}	2DE	278.3 (152.8-437.8)	-	-	-	12%
	(cm ³)						
	Left Atrial Ejection	3DE	40.8 (14.7-46.8)	-	35.2 (22.6-61.0)	-	~0%
	Fraction	2DE	45.6 (26.5-58.8)	-	-	-	12%
	(%)						

2b. Harmonic imaging								
Data	Volume	Method	Observer				$ICC_{inter-observer}$	obs_1 $ICC_{inter-method}$
			obs_1		obs_2			
			Median (range)	ICC_{horse}	Median (range)	ICC_{horse}		
Raw (4 repeat measurements / horse)	LAV _{max} (cm ³)	3DE	571 (404-888)	69%	893.5 (630-1431)	85%	-	-
		2DE	703 (468-1144)	77%	-	-	-	-
	LAV _{min} (cm ³)	3DE	386 (200-569)	90%	569.5 (406-815)	80%	-	-
		2DE	415 (220-776)	60%	-	-	-	-
	Left Atrial Emptying Volume _{max} (cm ³)	3DE	199.5 (64-542)	77%	317.5 (99-648)	75%	-	-
		2DE	279.0 (4-747)	42%	-	-	-	-
	Left Atrial Ejection Fraction (%)	3DE	31.2 (12.4-61.0)	86%	36.4 (13.6-50.5)	68%	-	-
		2DE	39.8 (0.5-65.3)	49%	-	-	-	-
Mean of the 4 repeat measurements / horse	LAV _{max} (cm ³)	3DE	590.4 (456.8-754.2)	-	888.7 (651-1214.5)	-	73%	33%
		2DE	702.9 (514.2-1018.8)	-	-	-	-	-
	LAV _{min} (cm ³)	3DE	406 (219.2-540.0)	-	567.9 (420.8-777.8)	-	68%	11%
		2DE	445.1 (263.2-594.2)	-	-	-	-	-
	Left Atrial Emptying Volume _{max} (cm ³)	3DE	173.5 (90.8-430.0)	-	330.8 (108.5-526.8)	-	36%	14%
		2DE	292.4 (99.0-447.2)	-	-	-	-	-
	Left Atrial Ejection	3DE	28.9 (18.5-58.3)	-	37.2 (14.9-43.2)	-	~0%	~0%

	Fraction (%)	2DE	41.0 (14.6-55.5)	-	-	-	
--	--------------	-----	------------------	---	---	---	--

Table 3: Summary of Bland-Altman analysis of obs_1 3DE measurements in comparison to both obs_2 3DE measurements and obs_1 2DE measurements for LAV_{max} and LAV_{min} volume, maximum Left Atrial Emptying Volume (LA EV_{max}) and Left Atrial Ejection Fraction (LA EF) from the mean value for each set of 4 repeated measurements per horse per observer per method (n = 22 horses).

LAV	LAV_{max} (cm ³)	LAV_{min} (cm ³)	LA EV_{max} (cm ³)	LA EF (%)	LAV_{max} (cm ³)	LAV_{min} (cm ³)	LA EV_{max} (cm ³)	LA EF (%)
Method	3DE	3DE	3DE	3DE	3DE vs 2DE	3DE vs 2DE	3DE vs 2DE	3DE vs 2DE
Observer	obs_1 vs obs_2	obs_1 vs obs_2	obs_1 vs obs_2	obs_1 vs obs_2	obs_1	obs_1	obs_1	obs_1
Mean Difference (cm ³)	-255.6	-154.4	-101.1	-0.7	-82.3	-18.3	-64.0	-5.2
Lower Limit (cm ³)	-689.8	-441.7	-393.3	-26.8	-420.2	-299.5	-316.4	-37.7
Upper limit (cm ³)	178.7	132.8	191.0	25.5	255.6	262.9	188.4	27.3

evj_13408_f1.tiff

evj_13408_f3.tif

evj_13408_f4.tif

evj_13408_f5.tiff

evj_13408_f6.tif