ESC European Society of Cardiology

European Heart Journal - Cardiovascular Imaging (2025) **00**, 1–14
https://doi.org/10.1093/ehjci/jeaf155

ORIGINAL PAPER

Vendor differences in 2D-speckle tracking global longitudinal strain: an update on a 10-year standardization effort

Anca Balinisteanu [1,2], Jürgen Duchenne [1,3,4], Alexis Puvrez[1], Laurine Wouters[1], Stéphanie Bézy[1], Ahmed Youssef[1], Lennert Minten[1], Youri Bekhuis[1,5], Leen Van Langenhoven[6], Konstantina Papangelopoulou[1], Aleksandra Cieplucha[1], Irene Cattapan[1], Paulo Tostes[1], Jan Bogaert [7,8], Dragos Vinereanu[2,9], James D. Thomas[10], Luigi P. Badano[11,12], and Jens-Uwe Voigt[1,13]*

[1]Department of Cardiovascular Sciences, KU Leuven, Herestraat 49, 3000 Leuven, Belgium; [2]Department of Cardiology and Cardiovascular Surgery, University of Medicine and Pharmacy "Carol Davila," Bucharest, Romania; [3]Research Group Cardio and Organ Systems, UHasselt, Hasselt, Diepenbeek, Belgium; [4]Department of Cardiology, Jessa Hospital, Hasselt, Belgium; [5]Faculty of Medicine and Life Sciences, UHasselt, Hasselt, Diepenbeek, Belgium; [6]Department of Biostatistics, KU Leuven, Leuven, Belgium; [7]Department of Radiology, University Hospitals Leuven, Leuven, Belgium; [8]Department of Imaging and Pathology, KU Leuven, Leuven, Belgium; [9]Department of Cardiology, University Emergency Hospital, Bucharest, Romania; [10]Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [11]Department of Cardiology, Istituto Auxologico Italiano, IRCCS, Milan, Italy; [12]Department of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy; and [13]Department of Cardiovascular Diseases, University Hospitals Leuven, Leuven, Belgium

Received 28 November 2024; revised 10 March 2025; accepted 12 May 2025; online publish-ahead-of-print 26 May 2025

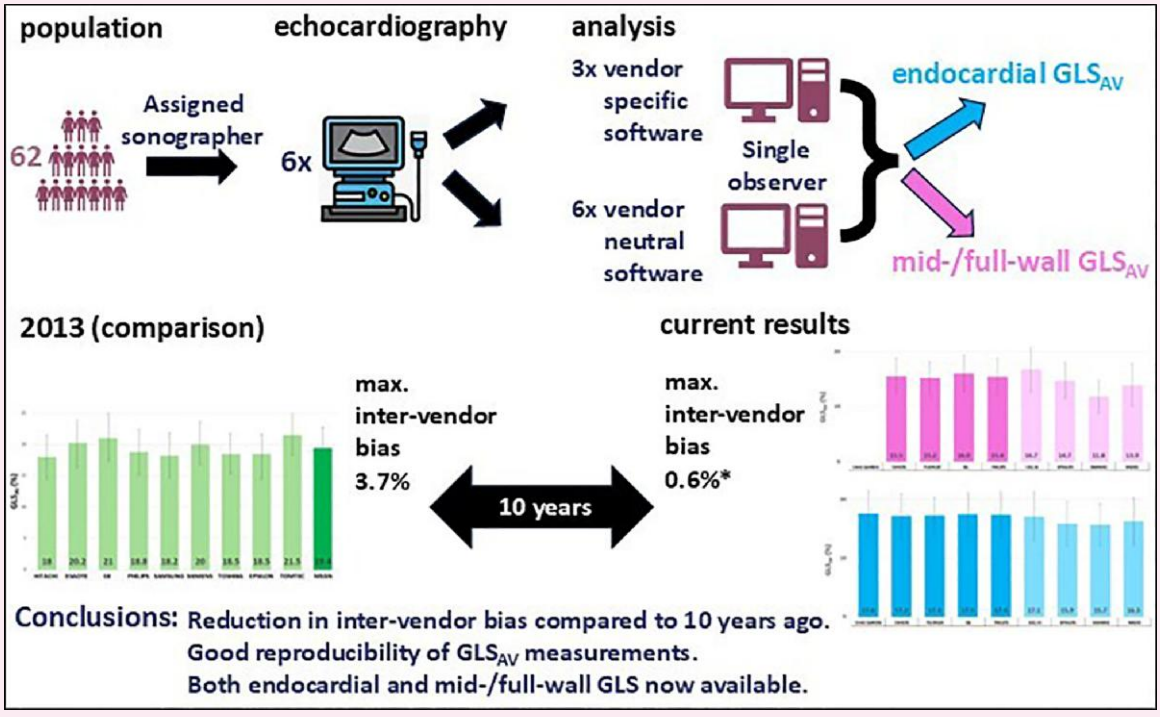
Aims To assess the inter-vendor differences in global longitudinal strain measurements and determine a potential improvement compared with the situation 10 years ago.

Methods and results 372 echocardiographic exams were performed in 62 subjects (50 male, age 56 ± 17) with LV ejection fraction ranging from 30% to 68%, using ultrasound systems from six manufacturers: GE, Philips, Canon, Siemens, Fujifilm and Esaote. Each subject was scanned consecutively on all machines by the same assigned sonographer, with two image sets per subject to assess test–retest setting reproducibility. Average peak systolic global strain from the three apical views (GLS_{AV}) was measured on three vendor-specific (Canon, Siemens, and Fujifilm) and six vendor-agnostic (GE, Philips, US2.AI, Caas Qardia, Medis, and Epsilon) software solutions (SWS). Endocardial and mid-/full-wall GLS were measured and compared with the mean GLS of contemporary semi-automated clinical software: GE, Philips, Canon, Fujifilm, and Caas Qardia. Endocardial and mid-/full-wall GLS measurements from contemporary semi-automated clinical software showed minimal inter-vendor differences, with an average maximum bias of 0.6% strain units. There was a remaining inter-vendor bias with and among some other vendors. The average minimal detectable change with contemporary semi-automated clinical software was 2.5 and 2.4 strain% for endocardial and mid-/full-wall GLS, resp. These values were higher for and among some other vendors. Test–retest variability of GLS measurements was good and similar to that of LV ejection fraction (6.6% vs. 6.5%, $P > 0.05$), indicating consistent results across repeated scans.

Conclusion In this controlled study setting, GLS measurements from companies that provide contemporary semi-automated clinical software have become more consistent, compared with 10 years ago. Mid-/full-wall strain was now available in all but one software.

* Corresponding author. E-mail: jens-uwe.voigt@uzleuven.be

Graphical Abstract



GLS$_{AV}$ = mean peak systolic global strain from the three apical views.

*Maximum average inter-bias for endocardial and mid-/full-wall GLS$_{AV}$ across study population using contemporary semi-automated software solutions for clinical use.

Keywords strain • speckle-tracking echocardiography • inter-vendor • left ventricle • standardization

Highlights

- We present latest findings on GLS inter-vendor differences compared with the baseline study from 10 years ago
- Both endocardial and mid-/full-wall GLS measurements were compared among three vendor-specific and six vendor-agnostic software solutions
- Contemporary semi-automated software solutions for clinical use showed an average inter-vendor bias of just 0.6% strain units which is a significant improvement compared with 10 years ago
- Under the strictly controlled conditions of this study, test–retest reproducibility for both endocardial and mid-/full-wall GLS was in general good and similar to EF, while larger variability remains among some vendors

Introduction

With the introduction of speckle tracking in Echocardiography in the early 2000s, global longitudinal strain (GLS) had been proposed as a new parameter for quantifying myocardial function.[1] Early clinical experience and validation studies indicated that the new method was sensitive to subtle left ventricular (LV) function changes and had diagnostic and prognostic potential.[1] Nevertheless, it also became obvious that there was a need to standardize strain measurements among different vendors.[2]

A Strain Standardization Task Force, initiated by the European Association of Cardiovascular Imaging (EACVI) in cooperation with the American Society of Echocardiography (ASE) and Industry Partners,[3,4] developed a consensus document to set standards and to enhance the comparability of strain measurements. At the same time, an initial Inter-vendor Comparison Study confirmed a good reproducibility of the technique, but also revealed significant differences in strain measurements among vendors, with the highest inter-vendor absolute mean difference of 3.7% strain units.[5]

Although the majority of strain data had been derived from full-wall speckle tracking, the initial Inter-vendor Comparison Study could only compare endocardial strain amongst vendors, as back in 2013, most companies could not provide mid-wall or full-wall strain measurements. In the meantime, most strain analysis packages provide both options, so that a comparison of both, endocardial as well as mid-wall/full-wall strain measurements is now possible.

The ultrasound industry has experienced substantial transformations over the past decade. The ultrasound branch of Toshiba has been taken over by Canon, the medical device division of Hitachi has been acquired by FUJIFILM, and Philips has incorporated TOMTEC as—amongst others—provider of strain analysis software. Additionally, there is an

Table 1 Vendors participating in the study with type and version of equipment and software provided

Vendor	Ultrasound machine	Type	Strain analysis software and version
GE	Vivid E95	High end	EchoPac 206
Philips	EPIQ CVx 9.0.3	High end	Tomtec Ultrasound Workspace 2.51.00
Canon	Aplio-I 900 Prims Edition V6.5	High end	UltraExtend NX V1.0
Siemens	Siemens Acuson SC2000 version 6.0	High end	Velocity Vector Imaging 3.0
Fujifilm	Lisendo 880LE version 5.0	High end	2D Tissue Tracking version 8.0a
Esaote	MyLAB X8EXP	Cart base	MyLAB Desk F12
Pie Medical Imaging[a]			Caas Qardia 2.0
Medis[a]			Medis Suite 4.062.4
Epsilon[a]			Echoinsight
US2.AI[a]			US2.AI V1(4.4)

Of note, in comparison to 2013, TOMTEC is now part of Philips. Canon is the successor of Toshiba. Fujifilm is the successor of Hitachi–Aloka. Siemens could not participate with their latest system and software as it was still lacking CE marking at the time of the study. Esaote had to withdraw after data acquisition due to a software issue.
[a]Software-only vendor.

increasing number of vendors providing semi-automated strain analysis software packages, and fully automated, artificial intelligence (AI) based echo analysis software solutions (SWS) have entered the market. Also driven by the efforts of the joint Strain Standardization Task Force, companies have worked on improving their software packages. Most vendors have released several new versions of their software in recent years.

The current study is a follow-up on the Inter-vendor Comparison Study conducted in Leuven, Belgium in 2013.[5] It was designed to: (i) assess the potential systematic bias among GLS measurements obtained by the analysis software of the different vendors, (ii) compare the robustness and test–retest variability of GLS measurements in a clinical setting, and, (iii) compare the results to our findings from 2013 in order to identify potential improvements and convergence, and to provide reliable data for future discussions among users and industry partners.

To the best of our knowledge, no other inter-vendor study in the last decade has compared in a clinical setting such a comprehensive number of both vendor-specific and vendor-agnostic software packages, incorporating both semi-automated and fully automated systems.

Methods

Study population

A list of potential patient candidates was created from hospital records of patients with myocardial infarction (MI) within 3 years before the study, treated at the University Hospitals Leuven. From this list, we selected those who had undergone a cardiac magnetic resonance (CMR) exam including late gadolinium enhancement imaging for scar definition. Care was taken that patients had been in stable condition, had not experienced any other ischaemic events or undergone any cardiac interventions between the CMR and the image acquisitions for the study. Of those, we identified sixty-three patients >18 years of age, with sinus rhythm, proper echocardiographic image quality, ability to walk and to lie in supine position for 2 h that finally consented to participate. Since we aimed at a wide range of LV functional states, we also recruited healthy volunteers as study subjects in stand-by. Those were asked to replace patients in case of no-show or other reasons preventing participation. Out of the 63 invited patients, 13 dropped out (3 due to atrial fibrillation and 10 due to no-show) and were replaced by healthy volunteers. One additional patient had to be excluded without replacement since an arrhythmia occurred late during the scanning session. Thus, the study population included complete data sets of a total of 62 subjects. A detailed overview of the subject inclusion process is provided in Supplementary data online, *Figure S1*. The study was approved by the Ethical Committee of the University Hospitals Leuven. All subjects provided written informed consent prior to inclusion.

Industry partner recruitment

An open letter was sent to potentially interested industry partners. Six ultrasound manufacturers (Philips, Andover, Massachusetts; GE Vingmed Ultrasound, Horten, Norway; Canon, Otawara, Japan; FUJIFILM, Tokio, Japan; Siemens, Mountain View, California; and Esaote, Florence, Italy) consented to participate and provided cardiac ultrasound machines and their proprietary 2D speckle-tracking software platforms. Furthermore, these companies agreed to send application specialists for technical support during data acquisition. In addition, five companies providing vendor-agnostic strain analysis software packages (US2.AI, Singapore; Philips (TOMTEC), Unterschleissheim, Germany; Epsilon, Ann Arbor, Michigan; Pie Medical Imaging, Maastricht, Netherlands; Medis, Leiden, Netherlands) participated in the study. All companies provided dedicated training sessions to make the reader of the study data familiar with their analysis software. Vendors that provided a version of their latest semi-automated clinical strain analysis SWS were GE, Philips (TOMTEC), Canon, Fujifilm, and Pie Medical Imaging with their software Caas Qardia. US2.AI provided a fully automated AI strain analysis solution, operating without any human intervention. Siemens participated with an older ultrasound machine and software, as the latest products were not yet CE-marked for human use at the time of the study. Epsilon participated with an older analysis software version, as the latest version was not yet available. Medis participated with software intended for research use only. Esaote had to withdraw from the study after completing image acquisition due to software issues encountered during the data analysis. However, the company continued to be indirectly involved through Caas Qardia multivendor software, which is integrated into Esaote's cardiac ultrasound machines. The complete list of participants, device and software versions are provided in *Table 1*.

Study protocol

Data acquisition was planned for one week in April 2023, comprising 9 scanning sessions of 2 h each. In each of these sessions, 7 study subjects were scanned on all 6 ultrasound machines in a randomized order. The ultrasound machines and beds were arranged in one single room and separated from each other by screens for patient privacy. In order to minimize variability in image acquisition, each patient was scanned by one dedicated

experienced echocardiographer on all ultrasound machines. A total of 15 echocardiographers participated, each with more than 6 years of experience and over 4000 echocardiographic exams. While the echocardiographers were responsible for providing optimal clinical image quality (e.g. exact scan planes, sample volume positions, etc.), the application specialists from the ultrasound manufacturers provided technical support and were responsible for optimizing machine settings according to the respective vendor recommendations for strain analysis.

After providing written informed consent, weight and height of each study subject was measured. Before and at the end of each scanning session, blood pressure was measured. Subjects were not allowed to drink or smoke during the scanning session in order to avoid changes in hemodynamic status.

Subjects were examined in left lateral decubitus position. Standard echocardiographic views from parasternal and apical windows were acquired with frame rate and image settings recommended by the respective company for speckle-tracking strain analysis. Three consecutive cardiac cycles were captured from each apical view during breath hold. Blood pool Doppler traces of the mitral valve were taken in order to have a time reference within the cardiac cycle. At the end of acquiring the first set of images, the echocardiographer was asked to stand up, leave the scanning bed and walk around for approximately one minute. Thereafter, scanning was resumed with the acquisition of a second set of apical views for test–retest variability analysis. All image data was stored as full frame rate Digital Imaging and Communication in Medicine (DICOM) data and—if available—in the respective proprietary raw data format of the company. Once a scan was completed, echocardiographer and patient moved on to the next ultrasound machine.

Data analysis

Image data from all ultrasound machines were analysed with the dedicated post-processing SWS provided by the respective companies (*Table 1*) using proprietary rate raw data, when possible. Full frame rate DICOM images (on average 56 frames per second (fps)) from GE were used for the analysis on the vendor-agnostic software packages from US2.AI, Caas Qardia, Medis and Epsilon. TOMTEC Ultrasound Workspace software was tested with Philips image data with an average of 64 fps as TOMTEC had been acquired by Philips. All strain measurements were performed by a single experienced observer (AB), a cardiologist with 9 years of experience in transthoracic echocardiography and over 10 000 echocardiographic exams. Prior to analysis, the observer completed dedicated training for each software provided by the respective vendor.

From the two acquired data sets, image loops with the best image quality were chosen as primary data set for the comparisons among vendors. The other data set was used in comparison to the first for determining test–retest variability.

In order to reduce bias from user interaction, we commonly accepted the beginning and the end time points of the cardiac cycle as automatically determined by the software packages. All SWS analysed a single cardiac cycle, except for US2.AI, which measured GLS as an average over multiple cycles.

For the assessment of global strain, the endocardial border was traced or automatically recognized and a region of interest (ROI) covering the myocardium was created according to the software requirements. Only when necessary, minimal manual corrections were made after a visual inspection how tracking indicators follow the myocardial motion. A maximum of five attempts was made to improve tracking in that way. If no appropriate tracking could be achieved, the respective segments were excluded from further analysis. In the case that tracking was deemed inadequate in more than one segment, the entire apical view was excluded.

Both endocardial and mid-/full-wall GLS measurements were extracted and compared. Under the term mid-/full-wall GLS, we summarize strain values derived either from tracking along a mid-wall ROI, or from averaging across the entire myocardial wall, depending on the respective vendor-specific SWS.[4]

Table 2 Subject characteristics ($N = 62$)

Demographics		
Age, years	56 ± 17	(23–78)
Male, *n* (%)	50	(80)
Clinical		
BMI, kg/m^2	26.4 ± 3.9	(18–39)
Baseline SBP, mmHg	138.7 ± 14.8	(102–179)
Baseline DBP, mmHg	81.1 ± 10.5	(57–110)
Final SBP, mmHg	144.8 ± 17.7	(105–183)
Final DBP, mmHg	82.9 ± 10.4	(60–114)
HR, beats/min	60.0 ± 9.1	(41–94)
Echocardiographic		
Biplane EF, %	50.5 ± 7.8	(30–68)
LVEDV, mL	126.9 ± 40.6	(65–306)
IVS, mm	10.2 ± 2.3	(6–21)
LVPW, mm	9.1 ± 1.6	(5–12)
LVEDD, mm	51.0 ± 6.7	(39–76)
E wave velocity, m/s	0.6 ± 0.2	(0.3–1)
E/*A* ratio	1.1 ± 0.4	(0.5–2.2)

Values are expressed as *n* (%) or mean ± SD (range). BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; EF, ejection fraction; LVEDV, left ventricular end-diastolic volume; IVS, interventricular septum; LVPW, left ventricular posterior wall; LVEDD, left ventricular end-diastolic diameter.

The mean peak systolic GLS (GLS_{AV}) was obtained by the average of peak systolic GLS values from the apical views: 4-chamber, 3-chamber, and 2-chamber views. In the current paper, GLS values are reported as absolute numbers.

Conventional echocardiographic parameters were determined by the same single observer on GE images using GE EchoPAC software version 206. We measured interventricular septum thickness, LV end-diastolic diameter, posterior wall thickness, LV end-diastolic volume, E wave velocity, *E*/*A* ratio and Biplane LV EF (EF_{Bi}) using automated ejection fraction function (AutoEF) (*Table 2*).

Since there is no gold standard for GLS measurements *in vivo*, we used as comparison the mean of the individual GLS measurements derived from SWS from all companies that participated with their latest version of a CE-marked conventional, semi-automated clinical tracking software (GE, Philips, Canon, Fujifilm, and Caas Qardia).

Statistical analysis

Statistical calculations were performed using the software SPSS 20.0 (Statistical Package for the Social Sciences) for Windows, ©SPSS Inc. Continuous variables were expressed by mean ± SD. The data followed a normal distribution (Kolmogorov–Smirnov test). Inter-vendor comparison of GLS measurements was determined by a linear mixed-effect model with software as a fixed effect and a random intercept for patient. Bonferroni correction was used to adjust for multiple testing in the post-hoc tests. GLS average values, as previously defined, were compared between vendors and with the mean of companies that participated with the most recent version of their semi-automated clinical SWS. Pearson correlation coefficients and Bland–Altman plots were used to describe the association and agreement between GLS_{AV} measurements from different vendors, and each vendor and mean reference GLS.

Test–retest variability was assessed under strictly controlled conditions in a true test–retest setting with two different sets of image acquisitions.

Absolute mean error, defined as the absolute value of the difference between test and retest measurements, and relative mean error (RME), defined as the ratio between the absolute difference of the measurements and the mean of the measurements, were used to describe test–retest variability. Inter-vendor comparison of test–retest variability was determined by a similar linear mixed-effect model. Comparison of test–retest variability of GLS measurements of each SWS with LV EF was performed using linear fixed-effect model. Paired *t*-tests were used to compare the RMEs between endocardial and mid-/full-wall GLS measurements within the same company. All the statistical tests were two-tailed and *P* value < 0.05 was considered statistically significant.

Test–retest reproducibility parameters within a specific vendor and among vendors were calculated using a single-measurement, absolute-agreement, 2-way random effects model. We used Fisher's transformation to calculate the 95% confidence intervals. We calculated standard error of measurement (SEM) as sqrt(1-intra-class correlation coefficient). The minimal detectable change is then calculated as 1.96*sqrt(2)*SEM.

Results

A total of 372 echocardiographic examinations were performed. Six single-image acquisitions were not suited for analysis. GLS_{AV} measurements were available from all vendors in 62 subjects except for Fujifilm, where three patients were excluded due to a technical acquisition problem and three others because of arrhythmia during the acquisition (56 subjects). All tracking SWS provided endocardial GLS measurements, while mid-/full-wall GLS measurements were available from all but one (Caas Qardia).

Subjects' demographic, clinical, and conventional echocardiographic characteristics are reported in *Table 2*. Mean age of the study cohort was 56 ± 17 years. There was a minor, but significant increase in systolic blood pressure during the scanning session of 4.3% (*Table 2*, $P < 0.05$). In our study population, LVEF ranged from 30% to 68% with a mean value of 50.5 ± 7.8%.

Inter-vendor variability

Mid-/full-wall GLS_{AV}

Mid-/full-wall GLS_{AV} measurements are shown in *Figure 1*. Companies that provided latest semi-automated clinical SWS showed similar results within a narrow range (mean GLS of vendors 15.5 ± 0.3%, with a maximum difference of mean GLS <1 strain%) (*Figure 1* and *Table 3*), while some of the other companies differed significantly from the field ($P < 0.05$).

The highest inter-vendor absolute difference amongst all vendors was 4.9% strain units (*Table 3*).

We found strong correlations between mid-/full-wall GLS_{AV} measurement of most of the vendors and the mean from clinical SWS ($r > 0.8$, $P < 0.0001$, *Figure 2A*). Bland–Altman plots of the same comparison are displayed in *Figure 2B*.

Endocardial GLS_{AV}

Endocardial GLS_{AV} measurements are shown in *Figure 3*. Consistent with the mid-/full-wall GLS_{AV} results, the measurements from companies providing clinical SWS were very close (mean GLS 17.4 ± 0.1%, with a maximum difference of mean GLS <1 strain%) (*Figure 3*).

The highest inter-vendor absolute difference between all vendors was 1.9% strain units (*Table 4*).

GLS_{AV} of each SWS vs. the mean of semi-automated clinical SWS showed significant correlation for most of the companies ($r > 0.8$, $P < 0.0001$, *Figure 4A*). Bland–Altman plots of the same comparison are displayed in *Figure 4B*.

Test–retest variability

Relative test–retest mean errors of mid-/full-wall GLS measurements ranged from 5.4% to 11.8% (*Figure 5*). A post-hoc test revealed a statistically significant difference to other vendors only for Medis (linear fixed-effect model, $P = 0.01$). For the endocardial GLS_{AV}, relative test–retest mean errors ranged from 5.2% to 8.9% (*Figure 6*). Post-hoc comparisons showed significant difference only for Medis ($P = 0.01$).

The absolute test–retest mean errors for mid-/full-wall GLS measurements ranged from 0.8% to 1.6% and from 0.9% to 1.4% for endocardial GLS (see Supplementary data online, *Figures S2* and *S3*). Detailed test–retest reproducibility parameters are presented in *Table 5* and Supplementary data online, *Tables S1* and *S2*.

Among semi-automated clinical software, staying within the same company, allows a minimal detectable strain change (MDC) between 2.0% (GE, Mid-/Full-wall) and 2.9% (Philips, Endocardial and Mid-/Full-wall) (*Table 5*). Switching to a software of another vendor during follow-up can increase this value (MDC on average 3.9% for both mid-/full-wall and endocardial strain) (see Supplementary data online, *Tables S1, S2*).

Test–retest RME of LV EF_{BI} was 6.5%. Several mid-/full-wall GLS_{AV} test–retest RMEs tended to be lower than that for EF, but the differences were neither on a per-company-level, nor on average statistically significant, except for one software provider, Medis (*Figure 5*). The same accounts for endocardial GLS_{AV} (*Figure 6*).

There were no statistically significant differences in test–retest variability between the endocardial and mid-/full-wall GLS_{AV} within each company.

Discussion

Main findings

Ten years after the first head-to-head comparison of two-dimensional speckle-tracking echocardiography derived GLS measurements obtained with SWS from different vendors, we compared again GLS results from three vendor-specific and six vendor-agnostic SWSs in a cohort of patients and volunteers with a wide range of LV function. Our main findings were: (i) Contemporary semi-automated clinical SWS show very similar GLS measurements. (ii) Test–retest reproducibility of GLS measurements was good for contemporary, semi-automated clinical SWS, though variability remains non-negligible for certain combinations of different vendors, (iii) in contrast to 10 years ago, almost all companies now provide a mid-/full-wall tracking option, and (iv) both endocardial and mid-/full-wall GLS_{AV} measured with conventional semi-automated SWS for clinical use show low and similar inter-vendor bias and reproducibility.

Reduction in inter-vendor bias

Our study suggests that the observed inter-vendor bias has significantly decreased over the past decade. This is in-line with other publications.[6] A recent study by Chamberlain *et al*. reported excellent agreement for GLS measurements between two vendor-specific and one vendor-independent SWS.[7]

Companies providing contemporary, semi-automated, clinically approved SWS now produce comparable measurement results with a maximum difference of less than 1 strain%. Our data suggest that the work of the strain standardization task force and efforts of the companies to implement its recommendations have been fruitful.

Nevertheless, when considering all software, inter-vendor bias persists. Some older SWS and research-use-only software packages still fell outside this narrow range of consistency. To which extent this will improve with updated versions remains to be tested. A previous inter-vendor study found that updates to vendor-specific speckle-

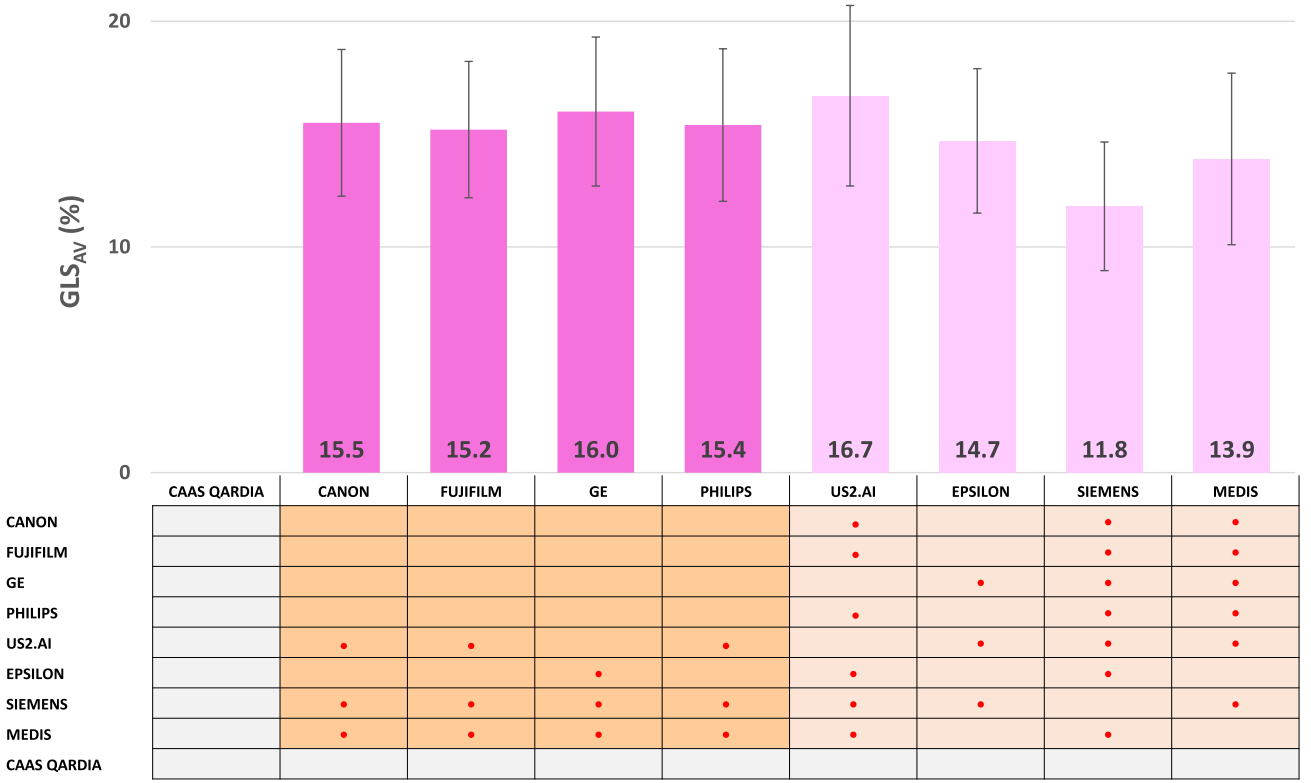


Figure 1 Mean mid-/full-wall GLS$_{AV}$ of our cohort per vendor. Errors bars represent standard deviation. Darker colours indicate companies that provided their latest version of semi-automated clinical software. The table provides the linear fixed-effect model pairwise comparison results. A red dot indicates a significance level of $P < 0.05$. Caas Qardia did not allow mid-full-wall GLS measurements.

Table 3 Bias and limits of agreement for mid-/full-wall GLS_{AV} measurements among all vendors

	CAAS QARDIA	CANON	FUJIFILM	GE	PHILIPS	US2.AI	EPSILON	MEDIS
CANON								
FUJIFILM		−0.5 (−4.1 to 3.2)						
GE		0.5 (−3.1 to 4.1)	1.0 (−2.1 to 4.1)					
PHILIPS		−0.0 (−4.4 to 4.4)	0.3 (−3.6 to 4.2)	−0.5 (−4.4 to 3.3)				
US2.AI		1.2 (−3.6 to 6.0)	1.6 (−2.8 to 6.1)	0.7 (−3.1 to 4.5)	1.2 (−3.5 to 6.0)			
EPSILON		−0.7 (−4.0 to 2.5)	−0.3 (−3.4 to 2.8)	−1.2 (−3.8 to 1.3)	−0.7 (−4.1 to 2.7)	−1.9 (−6.0 to 2.1)		
MEDIS		−1.5 (−6.2 to 3.1)	−1.1 (−5.2 to 3.0)	−2.0 (−5.7 to 1.6)	−1.5 (−6.5 to 3.5)	−2.7 (−7.0 to 1.5)	−0.8 (−4.3 to 2.7)	
SIEMENS		−3.7 (−7.4 to 0.1)	−3.1 (−6.9 to 0.7)	−4.2 (−7.6 to −0.7)	−3.6 (−8.2 to 0.9)	−4.9 (−10.2 to 0.5)	−2.9 (−6.5 to 0.6)	−2.1 (−7.0 to 2.8)

($n = 62$, for Fuji $n = 56$).

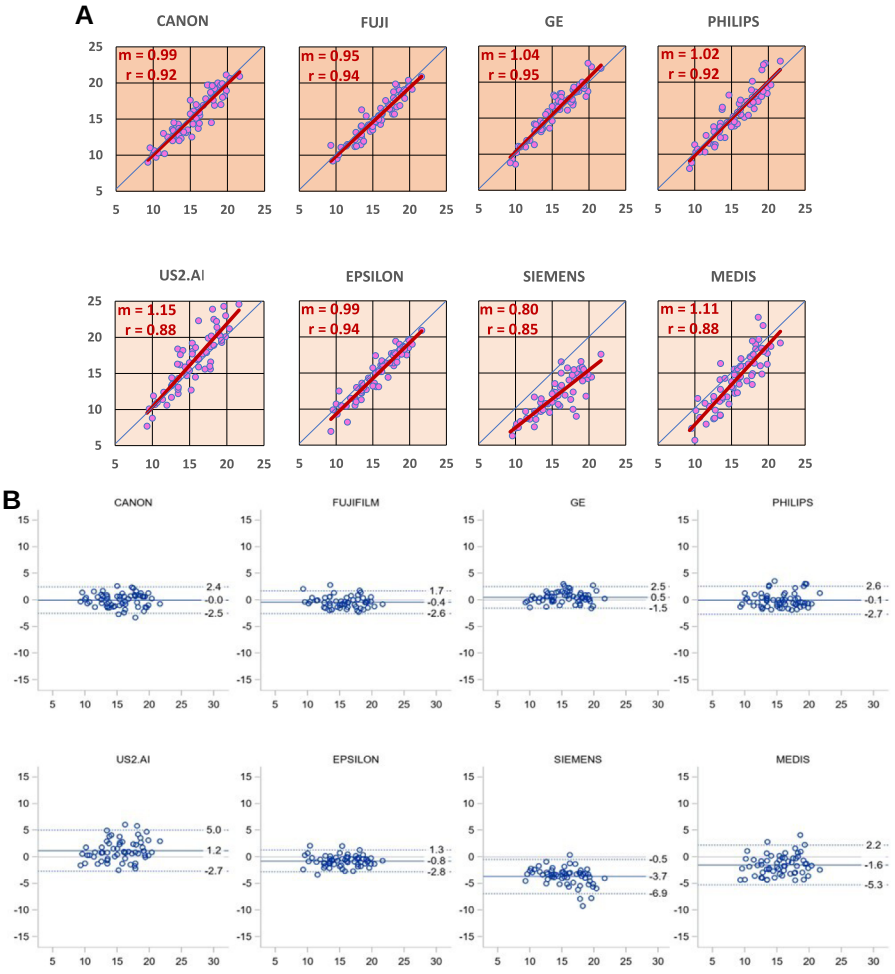


Figure 2 (*A*) Mid-/full-wall individual GLS_{AV} measurement of each vendor (y axis) vs. the mean GLS_{AV} of vendors (x axis). The plot areas shaded in darker colour indicate companies that provided their latest version of semi-automated clinical SWS and that have been included in this mean GLS (i.e. Canon, Fujifilm, GE, and Philips). Regression lines are shown together with their slope (*m*) and correlation coefficients (*r*) (red). (*B*) Bland–Altman plots for the same data showing the difference of the individual GLS_{AV} measurements of each vendor and the mean of vendors as described above (x axis). Blue line indicates bias and red lines show limits of agreement (1.96×SD).

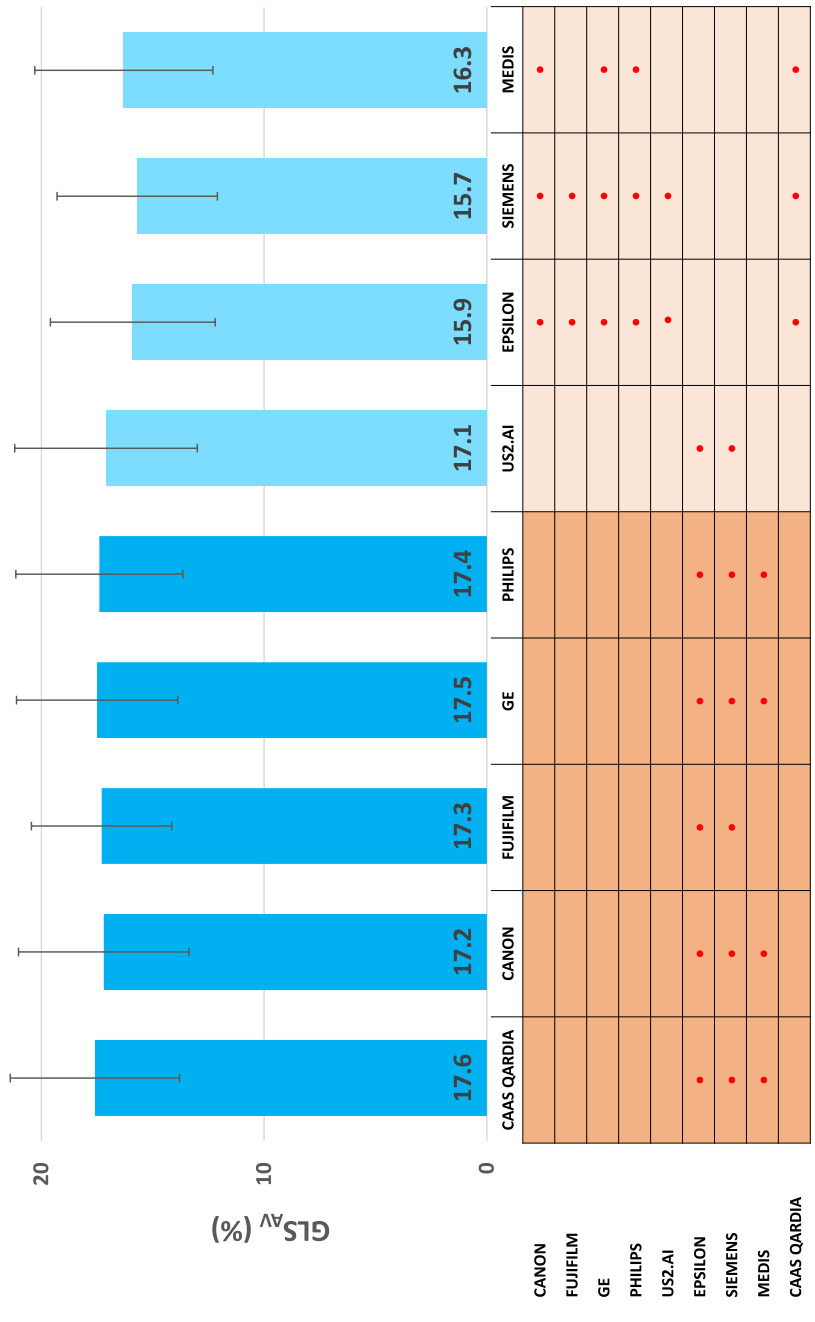


Figure 3 Same display as in *Figure 1*, but for endocardial GLS_{AV}.

Table 4 Bias and limits of agreement for endocardial GLS_{AV} measurements among all vendors

	CAAS QARDIA	CANON	FUJIFILM	GE	PHILIPS	US2.AI	EPSILON	MEDIS
CANON	−0.3 (−4.2 to 3.5)							
FUJIFILM	−0.4 (−4.2 to 3.2)	−0.1 (−4.3 to 4.0)						
GE	−0.1 (−2.7 to 2.5)	0.2 (−3.7 to 4.2)	0.4 (−2.8 to 3.6)					
PHILIPS	−0.2 (−4.1 to 3.7)	0.2 (−4.7 to 5.1)	0.1 (−4.3 to 4.5)	−0.1 (−4 to 3.9)				
US2.AI	−0.4 (−5.0 to 4.1)	−0.1 (−5.1 to 4.8)	−0.0 (−4.7 to 4.7)	−0.3 (−4.2 to 3.5)	−0.3 (−5.2 to 4.6)			
EPSILON	−1.7 (−5.2 to 1.9)	−1.3 (−5.3 to 2.7)	−1.3 (−4.9 to 2.4)	−1.6 (−4.5 to 1.4)	−1.5 (−5.2 to 2.2)	−1.2 (−5.1 to 2.7)		
MEDIS	−1.3 (−5.3 to 2.7)	−0.9 (−5.4 to 3.5)	−0.8 (−5.0 to 3.4)	−1.2 (−4.6 to 2.3)	−1.1 (−6.0 to 3.7)	−0.8 (−5.0 to 3.3)	0.4 (−3.2 to 4.0)	
SIEMENS	−1.9 (−6.2 to 2.5)	−1.5 (−5.9 to 3.0)	−1.2 (−6.0 to 3.5)	−1.7 (−5.7 to 2.3)	−1.7 (−7.1 to 3.7)	−1.4 (−7.0 to 4.2)	−0.2 (−4.9 to 4.6)	−0.6 (−5.6 to 4.5)

($n = 62$, for Fuji $n = 56$).

tracking software resulted in significant and variable changes in GLS measurements.[8]

Sources of residual bias

The remaining inter-vendor bias in both endocardial and mid-/full-wall GLS measurements can partially be explained by noise. Further, vendor-specific, proprietary post-processing software algorithms, with properties that are not disclosed, as well as typical, vendor-specific echo image characteristics may influence measurement results. Our findings strongly suggest that software algorithms play a paramount role in GLS variability, as indicated by the significantly different GLS values obtained from identical single-vendor source images when tested with different vendor-agnostic SWS.

Test–retest variability

This study confirmed that the test–retest variability of strain measurements is in general acceptable. Differences in measurement variability among contemporary, semi-automated clinical SWS were not significant for both mid-/full-wall and endocardial GLS_{AV}. Higher variability and larger MDC values were found with other vendors. Our results are in-line with a comparison study of feasibility, accuracy, and reproducibility of layer-specific GLS among five different vendors (GE, Hitachi, Siemens, Toshiba and Philips).[9]

Overall, our study revealed no relevant changes in GLS_{AV} variability compared with the situation 10 years ago, which confirms the notion that GLS_{AV} remains a reliable parameter for systolic myocardial function assessment, in particular for longitudinal monitoring of patients.

Comparison to the first inter-vendor comparison study

This study was conducted in a setting very similar to the first Inter-vendor Comparison Study in 2013.[5] Both studies aimed at assessing inter-vendor bias, as well as determining the reproducibility of GLS measurements in a test–retest scenario. While several vendors or their successors participated in both studies, we could this time also test for the first time a fully automated, AI-based software (US2.AI), which analyses echocardiographic studies without any user interaction.

Both studies compared the variability of GLS measurements to that of LV EF. While the study from 2013 reported a significantly better reproducibility of strain measurements, this advantage could not be reproduced in the current study due to a better performance of EF. We hypothesize, that the automated EF assessment of the current study contributed to the better reproducibility of EF measurements, which was now in the same range as that of GLS. These findings indicate, that part of the reproducibility advantage of GLS lies in the fact that it is an automated measurement with limited user interaction and that EF measurement can profit from this approach likewise.

In contrast to 2013, most companies provide now mid-/full-wall GLS measurements. We could therefore investigate both, endocardial as well as mid-/full-wall GLS_{AV}. We could not detect statistically significant differences between the two approaches, neither in bias nor in reproducibility. Nevertheless, given the significantly different values that both approaches produce, an expert agreement should be found for which parameter should be used by default in clinical reporting. An upcoming update of the 2015 EACVI/ASE consensus document on strain imaging might provide guidance in this respect.

The 2013 first Inter-vendor Comparison Study served as a benchmark for ongoing standardization initiatives. Farsalinos *et al.* revealed in the paper published in 2015 a significant bias in endocardial GLS_{AV} measurements, with the maximum absolute difference between vendors being 3.7% strain units.[5] Our recent findings indicate that, a decade later, there are still vendors that show discrepancies to the main field, both in endocardial and mid-/full-wall GLS measurements, with

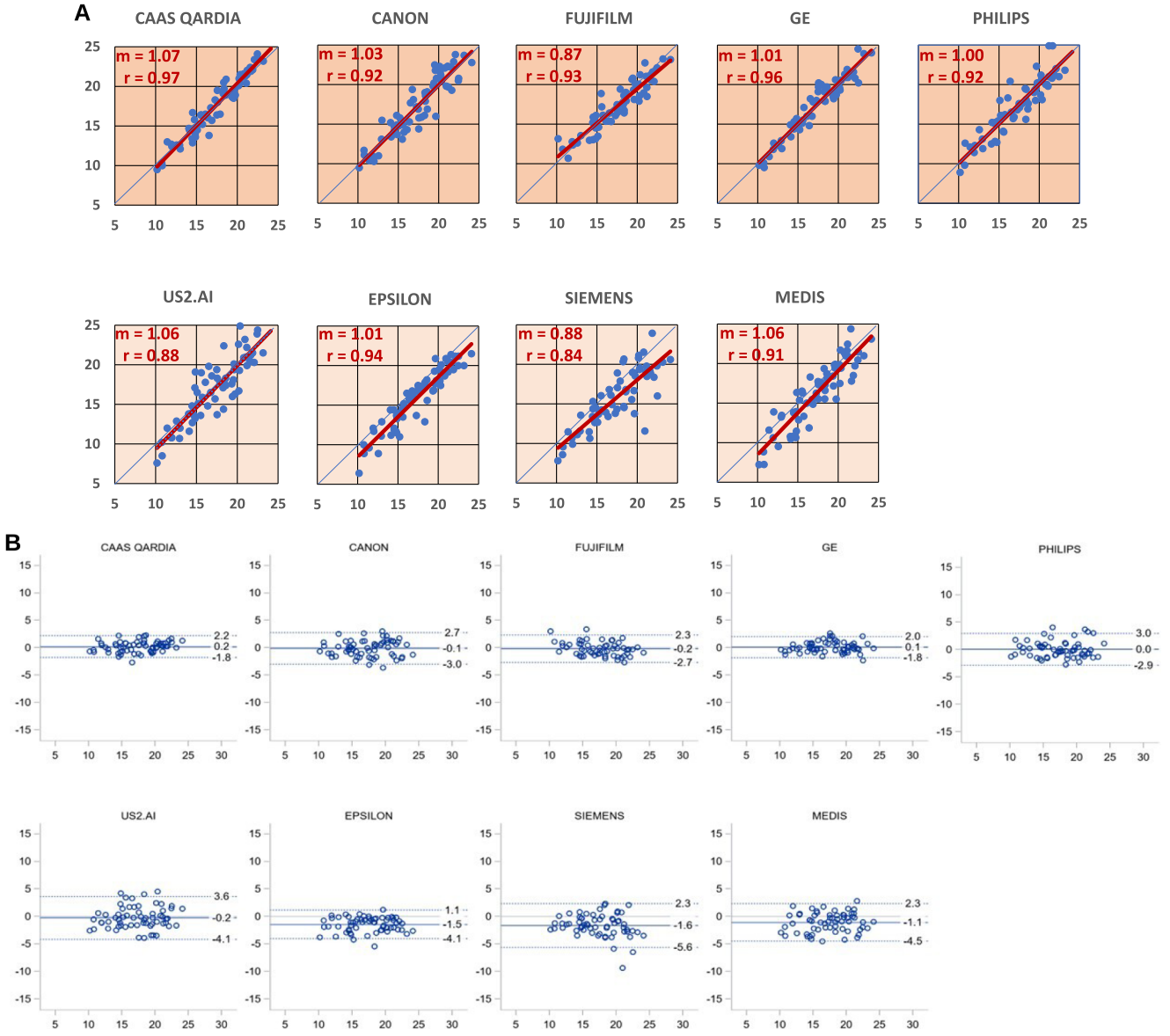


Figure 4 (*A* and *B*) Same display as in *Figure* 2, but for endocardial GLS_{AV}.

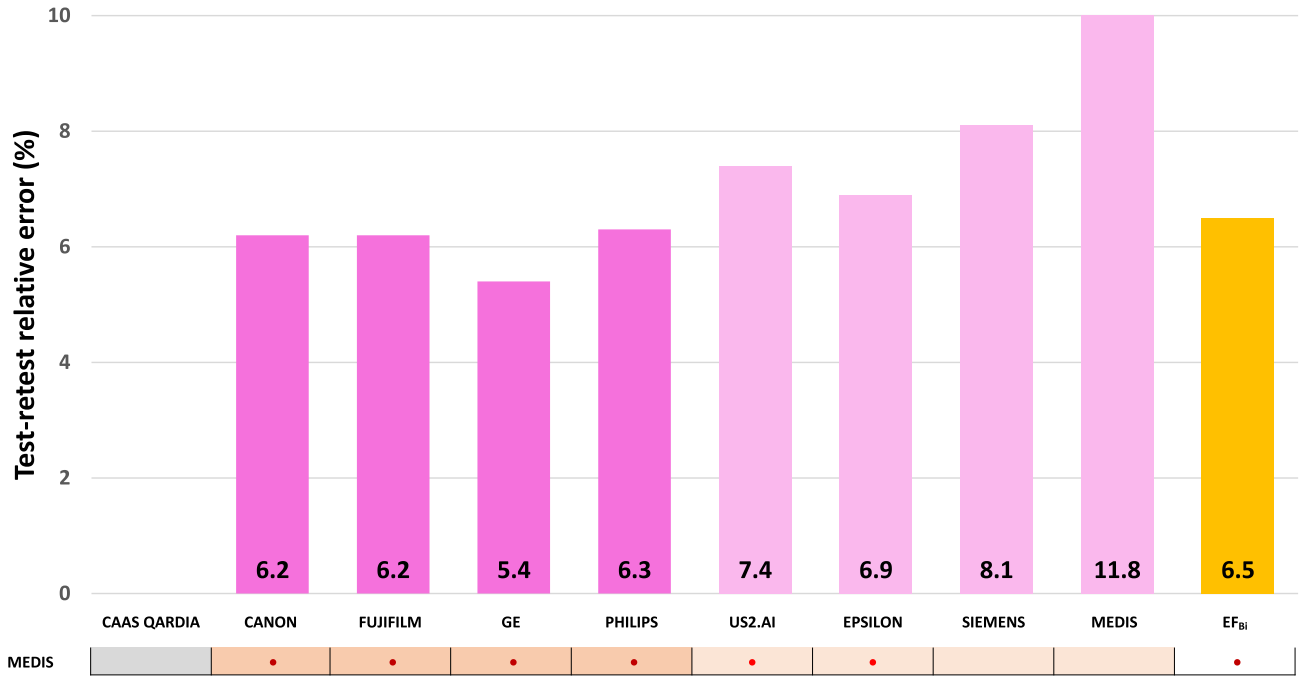


Figure 5 Test–retest RME of mid-/full-wall GLS_{AV} measurements from all vendors (pink) and EF_{BI} (orange). Significant differences were observed for Medis compared with some other vendors (post-test results, linear mixed-effect model), and Medis and EF_{BI} (post-Bonferroni correction, linear fixed-effect model). A dot represents a $P < 0.05$.

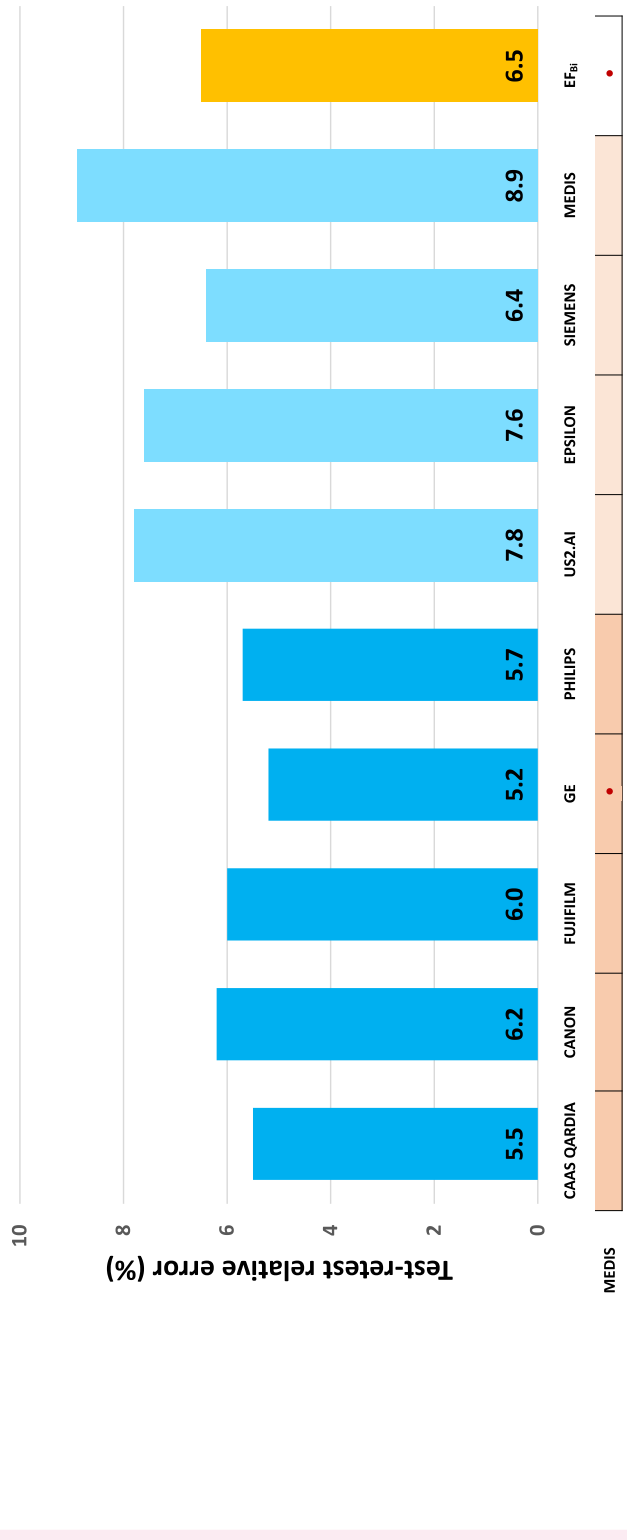


Figure 6 Same display as in *Figure 5*, but for endocardial GLS_{AV}.

Table 5 Test–retest reproducibility parameters by vendor

Layer	Vendor	Test–retest reproducibility		
		ICC (95% CI)	SEM	MDC
ENDOCARDIAL	CAAS QARDIA*	0.937 (0.834;0.977)	0.952	2.636
	CANON*	0.940 (0.895;0.966)	0.955	2.645
	FUJIFILM*	0.922 (0.872;0.953)	0.896	2.481
	GE*	0.951 (0.914;0.972)	0.778	2.156
	PHILIPS*	0.928 (0.884;0.956)	1.034	2.863
	US2.AI	0.916 (0.866;0.949)	1.195	3.311
	EPSILON	0.929 (0.880;0.958)	0.988	2.738
	SIEMENS	0.886 (0.818;0.929)	1.260	3.490
	MEDIS	0.892 (0.828;0.933)	1.320	3.657
MID-/ FULL-WALL	CAAS QARDIA*			
	CANON*	0.934 (0.891;0.961)	0.853	2.362
	FUJIFILM*	0.919 (0.867;0.952)	0.841	2.331
	GE*	0.948 (0.911;0.970)	0.726	2.011
	PHILIPS*	0.912 (0.859;0.946)	1.033	2.863
	US2.AI	0.922 (0.874;0.952)	1.106	3.064
	EPSILON	0.921 (0.839;0.962)	0.882	2.444
	SIEMENS	0.859 (0.769;0.916)	1.088	3.013
	MEDIS	0.831 (0.735;0.894)	1.493	4.135

*Companies that provided their latest version of semi-automated clinical software.
ICC, intra-class correlation; SEM, standard error of measurement; CI, 95% confidence interval based on Fisher's transformation of the ICC. MDC, Minimal detectable change.

maximum absolute differences of 1.9% for endocardial GLS_{AV} and 4.9% for mid-/full-wall GLS_{AV}. Nevertheless, a key finding of the current study is the evident trend towards a better consistency of GLS measurements among vendors offering contemporary semi-automated software that is approved for clinical use. Among those companies (GE, Philips, Canon, Fujifilm, and Pie Medical Imaging with their software Caas Qardia) the bias was particularly low (<1 strain% for both, endocardial and mid-/full-wall strain). This finding indicates that the collaboration between clinicians and industry partners in the strain standardization task force under the umbrella of EACVI and ASE in the past years was fruitful and lead to an improvement for the daily clinical use of speckle-tracking strain.

While the message of the first Inter-vendor study from 2013 was still to be cautious when comparing measurement results from SWS from different vendors, this warning can—to a certain extent—be revised in the light of our current findings. Our data indicate that on average current conventional semi-automated clinical SWS produce comparable measurement results. Nevertheless, the measurement reproducibility showed—although statistically not significant among companies with approved clinical software—a certain range. It can therefore be assumed, that upper and lower limits of normality and thresholds for detecting true changes may still slightly differ, depending on the analysis software used.

Limitations

This study assessed inter-vendor bias and its change over the past 10 years. We used a very comparable set-up and scanning protocol. However, patients were different and the pathology was this time uniformly coronary artery disease. Nevertheless, the range of LV function was comparable in both study cohorts, so that we believe that a comparison of the current results with those from 2013 is valid.

High-quality images are essential for reliable speckle-tracking-derived GLS analysis.[2] We evaluated the effect of software algorithms on measurement variability by comparing GLS results from subjects with suitable image quality. Therefore, this study did not investigate how image quality affects strain measurements on different SWS. Additional research is needed to explore this issue in more depth.

The vendor-agnostic SWS were tested using only images acquired with the machine of one vendor (GE). Although this allowed assessment of measurement variability, the analysis did not offer a comprehensive understanding of how image characteristics may affect tracking results. This aspect will be addressed in future analyses.

The selection of companies included in the mean GLS reference may lead to a stronger correlation of those companies with the mean. Nevertheless, given the lack of an objective reference, we believe that excluding software that is older or not approved for clinical use provided the most meaningful reference for comparison, also with respect the first inter-vendor study.

Some degree of caution is warranted when extrapolating the reproducibility data of this study to an everyday clinical setting. This study aimed primarily at comparing the performance of strain software from different vendors. Therefore, conditions were strictly controlled to minimize physiologic variation. Nevertheless, two datasets of apical images were acquired in each subject, which mimics a clinical follow-up scenario. However, the same examiner scanned the patient, and a company representative helped with optimizing the technical image quality. While this set up provided optimal conditions for the comparison among vendors, it does not fully reflect a true test–retest-scenario in routine clinical practice, where follow-up examinations are often performed by multiple echocardiographers and over extended time intervals.

The level of user interaction and semi-automatic user guidance may significantly affect measurement variability.[10] We therefore aimed at reducing bias from user interaction to a minimum. The data set was excluded, if more than five manual interventions were needed to optimize tracking.

AI for automatic measurements

Innovative AI-based methods for fully automated GLS measurement can provide reliable results without any operator input and with a high level of agreement to conventional semi-automatic speckle-tracking software. Such fully automated approaches will facilitate an even broader clinical adoption of speckle-tracking echocardiography.[11,12]

Outlook

Standardizing strain measurements requires a collaboration among researchers, clinicians, and industry partners. Transparency in inter-machine differences is essential, with manufacturers disclosing performance data and regulatory bodies ensuring standardization. Benchmarking against publicly available databases can enhance accuracy, build trust, and drive technological advancements. Fully automated software now matches traditional methods, reducing variability and improving consistency, leading to greater accuracy and reliability of strain measurement in the clinic.

Conclusions

Our study showed good reproducibility of GLS measurements within a strictly controlled test–retest scenario, which may to some extent reflect real-world clinical conditions. Along with the improved

reproducibility of EF observed in this study, these findings suggest a potential positive impact of (semi-) automated processing of data. GLS measurements can therefore be considered a reliable and robust parameter for LV function assessment, in particular in longitudinal studies.

Most notably, however, is a relevant reduction of inter-vendor-bias for GLS measurements with conventional clinical SWS when compared with the first inter-vendor-study in 2013. It is the understanding of both clinicians and industry partners within the strain standardization task force that its work under the umbrella of EACVI and ASE in the past 10 years has been productive and, despite some remaining variability, contributed to an improvement for the daily clinical use of speckle-tracking strain.

Acknowledgements

The Strain Standardization Task Force was initiated by the EACVI in collaboration with the ASE and industry partners under the lead of L.P.B. and J.D.T. The current work has been conducted under the auspices of the EACVI and the lead of J.-U.V. The authors thank all industry partners for their active support of this project and their continuing interest in and contribution to the work of the standardization task force. We appreciate the active support in study set-up and patient scanning from our colleagues, research fellows and students Kelly De Groef, Leticia Barrios, Christophe Vandebriele, and Guy Odent, as well as our echo lab co-workers Tine Vanden Acker, Dorien Eraerts, Joris Vermunicht, and Ibn Tielens.

Supplementary data

Supplementary data are available at *European Heart Journal - Cardiovascular Imaging* online.

Funding

Y.B. is supported by the Research Foundation Flanders (FWO) (T004420N). J.D., L.W., and J.-U.V. received personal research mandates of the FWO (12ZZN22N, 11PP524N, and 1832922N, respectively). A.Y. is supported by a scholarship from the Egyptian Ministry of Research and Higher Education (2021 recipient) and by the Belgian Fund for Cardiac Surgery (489708).

Conflict of interest: A report of the EACVI Task Force on Strain Standardization (chair: J.-U.V.), a collaboration of health professionals and Industry Partners: Nina Ajmone (University Leiden, NL); Gerard Aurigemma (University of Massachusetts, Worcester, USA); Luigi P. Badano (University Milano-Bicocca, IT); Anca Balinisteanu (KU Leuven, BE and University of Medicine and Pharmacy Carol Davila, Bucharest, RO); Victoria Delgado (Germans Trias i Pujol Hospital, Barcelona, ES); Erwan Donal (University Rennes, FR); Jürgen Duchenne (KU Leuven, BE); Thor Edvardsen (University Oslo, NO); Alan Fraser (University Cardiff, UK); Marc Friedberg (University of Toronto, CA); Ruxandra Jurcut (Institute of Cardiovascular Diseases Prof. C.C. Iliescu, Bucharest, RO); Jaroslaw D. Kasprzak (Medical University of Lodz, PL); Theodore Kolias (University of Michigan, Ann Arbor, USA); Shelby Kutty (Johns Hopkins School of Medicine, Baltimore, USA); Tom Marwick (Baker Institute, Melbourne, AU); Luc Mertens (University of Toronto, CA); Denisa Muraru (University Milano-Bicocca, IT); James D. Thomas (Northwestern University, Chicago, US); Jens-Uwe Voigt (KU Leuven, BE); Wim Gorissen (Canon); Carla Witkam (Canon); Megan McAuliffe (Epsilon Imaging); James Graham (Epsilon Imaging); Marco Maglione (Esaote); Stefano Pedri (Esaote); Tomoaki Chono (FUJIFILM); Saki Ochiai (FUJIFILM); Gunnar Hanssen (GE Vingmed Ultrasound); Hani Nozari Mirar (GE Vingmed Ultrasound); Sten Roar Snare (GE Vingmed Ultrasound); Marieke Vermeulen (Medis Imaging); Verena Roediger (Philips); Marcus Schreckenberg (Philips); Jane Vogel (Philips); Jean-Paul Aben (Pie Medical Imaging); Maarten Heusinkveld (Pie Medical Imaging); Myra Cocker (Siemens Healthineers); Gunilla Fritzon (Siemens Healthineers); Marti McCulloch (Siemens Healthineers); Michael Snyder (Siemens Healthineers); Jose Rivero (US2.AI); Yoran Hummel (US2.AI). Aleksandra Cieplucha is currently affiliated with the University of Medicine and Pharmacy, Poznan, Poland. Irene Catapan is currently affiliated with the University of Medicine and Pharmacy, Padua, Italy.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

1. Reisner SA, Lysyansky P, Agmon Y, Mutlak D, Lessick J, Friedman Z. Global longitudinal strain: a novel index of left ventricular systolic function. *J Am Soc Echocardiogr* 2004;**17**:630–3.
2. Negishi K, Lucas S, Negishi T, Hamilton J, Marwick TH. What is the primary source of discordance in strain measurement between vendors: imaging or analysis? *Ultrasound Med Biol* 2013;**39**:714–20.
3. Thomas JD, Badano LP. EACVI-ASE-industry initiative to standardize deformation imaging: a brief update from the co-chairs. *Eur Heart J Cardiovasc Imaging* 2013;**14**:1039–40.
4. Voigt JU, Pedrizzetti G, Lysyansky P, Marwick TH, Houle H, Baumann R *et al.* Definitions for a common standard for 2D speckle tracking echocardiography: consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. *Eur Heart J Cardiovasc Imaging* 2015;**16**:1–11.
5. Farsalinos KE, Daraban AM, Ünlü S, Thomas JD, Badano LP, Voigt JU. Head-to-head comparison of global longitudinal strain measurements among nine different vendors: the EACVI/ASE Inter-vendor comparison study. *J Am Soc Echocardiogr* 2015;**28**:1171–81.e2.
6. Shiino K, Yamada A, Ischenko M, Khandheria BK, Hudaverdi M, Speranza V *et al.* Intervendor consistency and reproducibility of left ventricular 2D global and regional strain with two different high-end ultrasound systems. *Eur Heart J Cardiovasc Imaging* 2017;**18**:707–16.
7. Chamberlain R, Shiino K, Scalia GM, Sabapathy S, Chan J. Advantage and validation of vendor-independent software for myocardial strain analysis compared to vendor-specific software. *Australas J Ultrasound Med* 2020;**24**:48–57.
8. Nagata Y, Takeuchi M, Mizukoshi K, Wu VC, Lin FC, Negishi K *et al.* Intervendor variability of two-dimensional strain using vendor-specific and vendor-independent software. *J Am Soc Echocardiogr* 2015;**28**:630–41.
9. Ünlü S, Mirea O, Duchenne J, Pagourelias E, Bézy S, Thomas JD *et al.* Comparison of feasibility, accuracy, and reproducibility of layer-specific global longitudinal strain measurements among five vendors: a report from the EACVI-ASE-Strain Standardization Task Force. *J Am Soc Echocardiogr* 2018;**31**:374–80.e1.
10. Mada RO, Lysyansky P, Daraban AM, Duchenne J, Voigt JU. How to define end-diastole and end-systole? Impact of timing on strain measurements. *JACC Cardiovasc Imaging* 2015;**8**:148–57.
11. Salte IM, Østvik A, Smistad E, Melichova D, Nguyen TM, Karlsen S *et al.* Artficial intelligence for automatic measurement of left ventricular strain in echocardiography. *JACC Cardiovasc Imaging* 2021;**14**:1918–28.
12. Sveric KM, Botan R, Winkler A, Dindane Z, Alothman G, Cansiz B *et al.* The role of artificial intelligence in standardizing global longitudinal strain measurements in echocardiography. *Eur Heart J Imaging Methods Pract* 2024;**2**:qyae130.