AUTHOR QUERY FORM

ELSEVIER	**Journal:** JCMG **Article Number:** 4657	**Please e-mail your responses and any corrections to:** **E-mail:** l.loveall@elsevier.com

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. **It is crucial that you NOT make direct edits to the PDF using the editing tools as doing so could lead us to overlook your desired changes.** Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions.

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof.

Location in article	Query / Remark: Click on the Q link to find the query's location in text Please insert your reply or correction at the corresponding line in the proof
	NOTE: The DOI of your paper is located on the proof (in most cases at the bottom of the first page, or near the author correspondence information).
Q1	If there are any drug dosages in your article, please verify them and indicate that you have done so by initialing this query. Also, please review/confirm the spelling of all drug names in figures, tables, and text.
Q2	Please confirm the relationship with industry information for each author is correct, accurate, and complete.
Q3	The Central Illustration may have been altered per the Editors and has been redrawn into journal style using the journal's color palette. Please review it carefully to ensure it is correct as is.
Q4	Please note: figures are redrawn to adhere to journal style; please review each figure carefully to ensure all data representation is correct and that all colors represented in the figure are represented correctly in the figure legend.
Q5	Au: Please provide professional degrees for Ishan Jain, Xiu Tang, and Andrea Ferencz.
Q6	AU: What are the range values in ( )?
Q7	AU: Reference 12: Please provide the date you accessed the website.
Q8	AU: Reference 14: Please provide the date you accessed the website.
Q9	AU: Reference 17: Please provide the date you accessed the website.
Q10	AU: Please provide a title for Figure 1.

(*continued on next page*)

Q11	Please confirm that given names (displayed in pink) and surnames (displayed in blue) have been identified correctly and are presented in the desired order and please carefully verify the spelling of all authors' names. Please check this box or indicate your approval if you have no corrections to make to the PDF file ☐

Thank you for your assistance.

JACC: CARDIOVASCULAR IMAGING VOL. ■, NO. ■, 2024

PUBLISHED BY ELSEVIER

ORIGINAL RESEARCH

Deep Learning-Derived Myocardial Strain

Alan C. Kwan, MD,[a,*] Ernest Chang, MD, PhD,[b,*] Ishan Jain,[a] John Theurer, BS,[a] Xiu Tang,[c] Nadia Francisco, MSc,[c] Francois Haddad, MD,[c] David Liang, MD, PhD,[c] Alexandra Fábián, MD, PhD,[d] Andrea Ferenci,[d] Neal Yuan, MD,[e] Béla Merkely, MD, PhD, DSc,[d] Robert Siegel, MD,[a] Susan Cheng, MD, MMSc, MPH,[a] Attila Kovács, MD, PhD,[d] Márton Tokodi, MD, PhD,[d,†] David Ouyang, MD[a,†]

ABSTRACT

BACKGROUND Echocardiographic strain measurements require extensive operator experience and have significant intervendor variability. Creating an automated, open-source, vendor-agnostic method to retrospectively measure global longitudinal strain (GLS) from standard echocardiography B-mode images would greatly improve post hoc research applications and may streamline patient analyses.

OBJECTIVES This study was seeking to develop an automated deep learning strain (DLS) analysis pipeline and validate its performance across multiple applications and populations.

METHODS Interobserver/-vendor variation of traditional GLS, and simulated effects of variation in contour on speckle-tracking measurements were assessed. The DLS pipeline was designed to take semantic segmentation results from EchoNet-Dynamic and derive longitudinal strain by calculating change in the length of the left ventricular endocardial contour. DLS was evaluated for agreement with GLS on a large external dataset and applied across a range of conditions that result in cardiac hypertrophy.

RESULTS In patients scanned by 2 sonographers using 2 vendors, GLS had an intraclass correlation of 0.29 (−0.01 to 0.53, $P = 0.03$) between vendor measurements and 0.63 (0.48-0.74, $P < 0.001$) between sonographers. With minor changes in initial input contour, step-wise pixel shifts resulted in a mean absolute error of 3.48% and proportional strain difference of 13.52% by a 6-pixel shift. In external validation, DLS maintained moderate agreement with 2-dimensional GLS (intraclass correlation coefficient [ICC]: 0.56, $P = 0.002$) with a bias of −3.31% (limits of agreement: −11.65% to 5.02%). The DLS method showed differences ($P < 0.0001$) between populations with cardiac hypertrophy and had moderate agreement in a patient population of advanced cardiac amyloidosis: ICC was 0.64 (0.53-0.72), $P < 0.001$, with a bias of 0.57%, limits of agreement of −4.87% to 6.01% vs 2-dimensional GLS.

CONCLUSION The open-source DLS provides lower variation than human measurements and similar quantitative results. The method is rapid, consistent, vendor-agnostic, publicly released, and applicable across a wide range of imaging qualities. (J Am Coll Cardiol Img 2024;■:■–■) © 2024 by the American College of Cardiology Foundation.

From the [a]Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA; [b]Leon H. Charney Division of Cardiology, Department of Medicine, New York University Grossman School of Medicine, New York, New York, USA; [c]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA; [d]Heart and Vascular Center, Semmelweis University, Budapest, Hungary; and the [e]Division of Cardiology, Department of Medicine, San Francisco VA, University of California-San Francisco, San Francisco, California, USA. *Drs Kwan and Chang contributed equally to this work. †Drs Tokodi and Ouyang are co-senior authors.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received July 10, 2023; revised manuscript received January 11, 2024, accepted January 17, 2024.

ISSN 1936-878X/$36.00 https://doi.org/10.1016/j.jcmg.2024.01.011

ABBREVIATIONS AND ACRONYMS

2D = 2-dimensional

3D = 3-dimensional

A4C = apical 4-chamber

DLS = deep-learning strain

GLS = global longitudinal strain

ICC = intraclass correlation coefficient

LOA = limits of agreement

LV = left ventricle

MAE = mean absolute error

Myocardial strain can be used to identify subtle findings of cardiac dysfunction.[1] It has specific applications in the longitudinal monitoring of cardiotoxic therapies as well as in assessing cardiomyopathies.[2] Compared with conventional measures of left ventricular (LV) function such as ejection fraction, longitudinal strain has additional prognostic value for predicting cardiovascular outcomes.[3-5] Strain is formally defined as the change in the length (L) of a myocardial segment relative to its initial diastolic length (L_0): strain = $(L - L_0)/L_0$.[6] However, the method of measuring the length of a myocardial segment can significantly vary across vendors, impacting the reproducibility of assessment of strain across studies and vendors.[7]

In echocardiography, speckle tracking, which is based on block-matching algorithms, is used to calculate longitudinal strain[1]; however, the application of this method depends on proprietary software implementations and requires experience for optimal placement of initial contours.[8] Such heterogeneity results in significant intervendor/intravendor variability and limits the ability to compare assessments over time and analyses across ultrasound manufacturers and software vendors.[9,10] For example, within cardio-oncology, differences as small as a relative change of ≥15% (3% absolute change from a global longitudinal strain [GLS] of 20%) can be clinically significant.[2] Given the widespread emergence of artificial intelligence anatomical segmentation models in medicine,[11] adapting strain measurement to leverage these models may improve accessibility and generalizability beyond traditional block-matching-based approaches. Creation of an automated, open-source, vendor-agnostic method to measure GLS from standard B-mode images would greatly improve research applications and may streamline patient analyses.

In this study, we aimed to develop an accessible strain assessment pipeline that leverages an automated, deep learning semantic segmentation model in order to reduce interobserver and vendor variability. We analyze limitations in current strain approaches, present an open-source deep learning-based solution to quantify LV longitudinal strain, and evaluate its performance against traditional manual strain measurements in both internal and external validation cohorts. We hypothesized that our method of deep learning strain (DLS) assessment could provide consistent and robust strain measurements in an automated fashion, without the need for observer-specific frame selection or segmentation, and vendor-specific proprietary motion estimation.

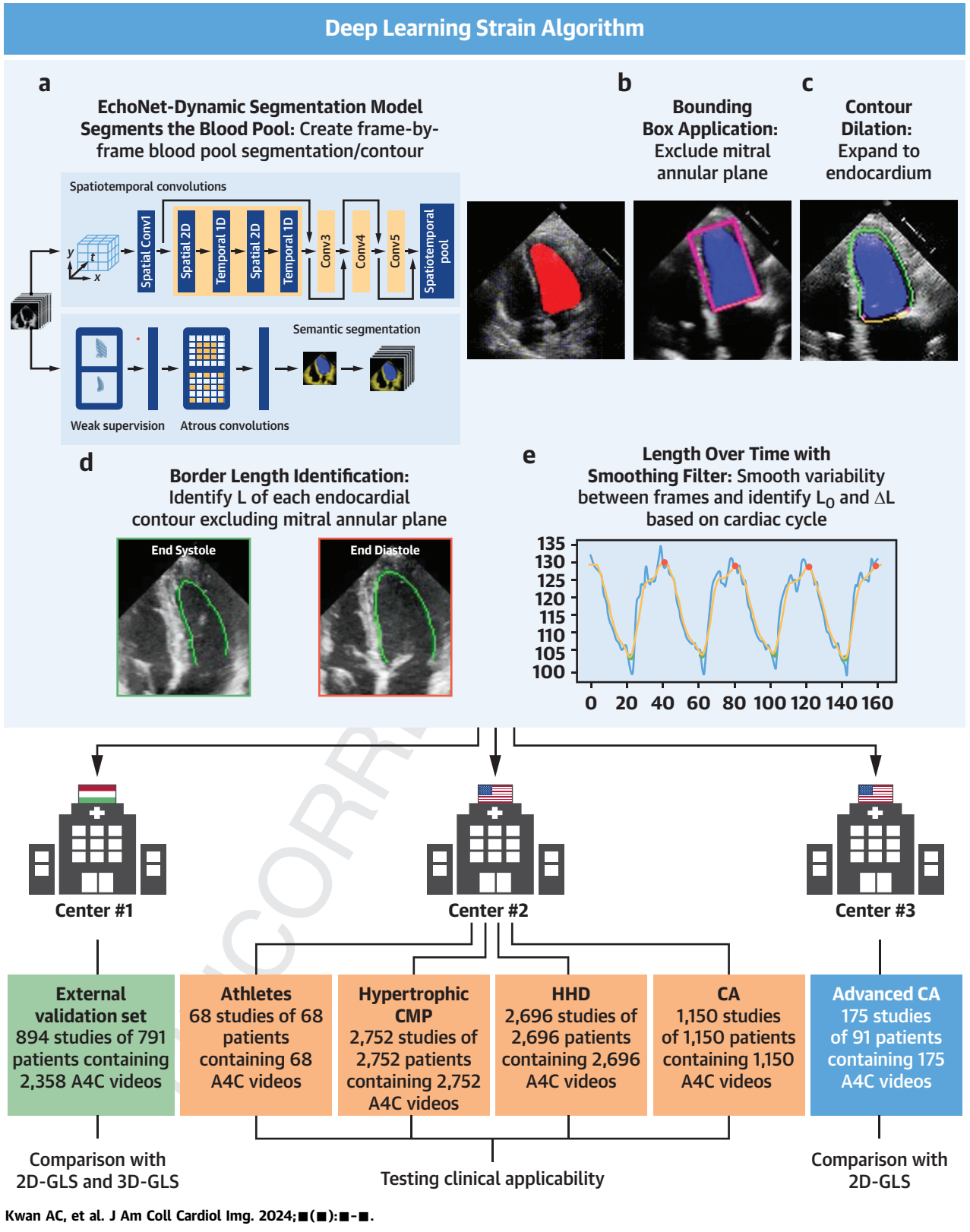
METHODS

QUANTIFYING CLINICAL VARIATION IN TRADITIONAL STRAIN. Intervendor and observer variability is a challenge for standard GLS. The variability in standard clinical measurements of GLS was assessed using consecutive image acquisitions performed by 2 sonographers with advanced cardiac certification and more than 15 years of experience. Both sonographers scanned the same 43 unique patients using a Philips Epiq 7C (GLS in QLAB, version 10.2) and a GE Vivid E95 (GLS in EchoPAC, version 2.01) ultrasound system and assessed 2-dimensional (2D)-GLS on the acquired videos using the corresponding speckle-tracking software packages, resulting in a total of 4 2D-GLS values per patient. The sonographers were blinded to each other's acquisitions and results, and both scanned each patient on the same day to acquire a total of 4 acquisitions per patient in a prospective manner.

Manual contour placement or adjustment of automated contours as necessary for most speckle-tracking approaches also has inherent variability. To assess the impact of minor variations in contouring on 2D-GLS, we measured the quantity of change in 2D-GLS after shifting the input contour by specified pixel distances. Original end-diastolic contours from 262 clinical cases were taken, and pixel-by-pixel translations in the positive x-axis direction were created from 0 to 6 pixels. Each end-diastolic contour was provided as the baseline contour to a validated custom speckle-tracking-based approach (code and validation available at Chang),[12] with measurement of 2D-GLS by 2D normalized cross-correlation algorithm with a sliding frame of reference implemented in MATLAB 2022b. Mean absolute error and proportional difference was measured compared with the initial GLS measurement.

DEEP LEARNING ALGORITHM DESIGN AND IMPLEMENTATION. We adapted EchoNet-Dynamic, a deep learning-based segmentation network tailored to analyzing echocardiograms,[13] to identify the endocardium-blood pool interface across the cardiac cycle using apical 4-chamber (A4C) view echocardiographic videos (**Central Illustration**). EchoNet-Dynamic was trained on 7,465 A4C view echocardiographic videos of 7,465 unique patients from Stanford Healthcare and validated internally on an additional 2,565 videos. The dataset is publicly available in Ouyang[14] and has been previously described in detail.[13] Briefly, the model is a convolutional neural network with atrous convolutions to

CENTRAL ILLUSTRATION



Kwan AC, et al. J Am Coll Cardiol Img. 2024;■(■):■-■.

Continued on the next page

perform semantic segmentation identifying the blood pool of the LV. The model was trained using manual segmentations of the LV at end-systole and end-diastole with weak supervision to create frame-level semantic segmentation (eg, the blood pool is identified in all frames within the provided echocardiography clip) (**Central Illustration A**). Following the segmentation performed on an A4C view video, a bounding box is placed over the segmented LV blood pool, which can be used to identify the mitral annular plane as a flat section of the contour (**Central Illustration B**). By measuring the perimeter of the segment with excluding the mitral annular plane, we can identify the longest and shortest lengths of the endocardial contour in a single cardiac cycle (**Central Illustration D**) to calculate per-beat strain and the average strain over all of the cardiac cycles within a single video. The methodology can be thought of as a deterministic measurement of change in the blood pool contour, applied to an output from a deep learning-based semantic segmentation model and, therefore, could be easily adapted to other segmentation models.

Given that the exact boundary of the LV may vary due to the presence of trabeculations, we tested varying degrees of dilation of the contour and their effect on reproducibility (**Central Illustration C**). The length of the contour was quantified in each frame. Using Python's SciPy function scipy.signal.find_peaks[15] over a 32-frame sliding window, we identified the maximum and minimum lengths of the contour, corresponding with end-systole and end-diastole. Parameter tuning was performed using the repeated measures cohort by identification of lowest variability between cardiac cycle-to-cycle repeated measurements of strain by DLS within individual video acquisitions. Dilation was used to optimize the tracked region to be closest to the mid-myocardium instead of the endocardium or epicardium, consistently with vendor-specific approaches. For the former, 0-pixel dilation (ie, no dilation) was compared with 3- and 5-pixel dilation. When dilated with 0 pixels, the mean SD of cycle-to-cycle DLS was 2.57%, whereas 3-pixel and 5-pixel dilation resulted in a mean SD of 2.02% and 2.50%, respectively. Given the expected smooth contraction and relaxation of the LV, we evaluated multiple smoothing filters, including Savitzky-Golay (Savgol), convolve average, moving average, and low-pass, to reduce frame-to-frame variance in length across a cardiac cycle (**Central Illustration E**). The mean with the Savgol filter was 1.35%, with the low-pass filter was 1.52%, with the convolve average filter was 1.40%, and with the moving average filter was 1.89%. Therefore, the Savgol filter with 3-pixel dilation was chosen for the final pipeline given the lowest variation between multiple measurements. Algorithms were implemented in Python (version 3.9.5, Python Software Foundation) using open-source libraries, including PyTorch, SciPy, Torchvision, and OpenCV.[16] The code is publicly available in Ouyang.[17]

2D AND 3D EXTERNAL VALIDATION OF THE DLS ALGORITHM. The DLS algorithm was applied to the A4C view videos of patients who underwent clinically indicated 2D and 3D echocardiographic examination at Semmelweis University (Budapest, Hungary) between November 2013 and March 2021. This external cohort consisted of patients with different cardiac diseases (such as heart failure, cardiomyopathies, and valvular heart disease), healthy subjects with no cardiovascular risk factors or established cardiac diseases, and athletes. DLS was compared with 2D echocardiography-derived GLS (2D-GLS) measured on the same A4C view videos using the AutoStrain LV (TomTec Imaging Systems) speckle-tracking software package, and 3-dimensional (3D) echocardiography-derived GLS (3D-GLS) assessed using the 4D LV-Analysis 3 (TomTec Imaging Systems) software package. All videos were reviewed by an experienced sonographer who assessed the image quality using a 5-point Likert scale (1 = nondiagnostic, 2 = poor, 3 = moderate, 4 = good, 5 = excellent).

DLS ANALYSIS OF DISEASE PHENOTYPES. To test potential clinical applicability of the algorithm, we acquired echocardiography videos from patients in

CENTRAL ILLUSTRATION Continued

Schematic illustration of the DLS pipeline, which comprises the following consecutive steps: (a) segmentation of the left ventricular (LV) blood pool, (b) application of a bounding box to exclude the mitral annular plane from the LV blood pool contour, (c) dilation of the contour, (d) calculating the length of the contour in each frame of the video, and (e) calculating strain from the contour length tracing after applying the Savitzky-Golay smoothing filter. The algorithm is then validated in 3 clinical locations including a large external dataset from Semmelweis University, multiple LV hypertrophy-associated conditions from Stanford University cardiology specialty clinics, and advanced cardiac amyloidosis patients from Cedars-Sinai Medical Center. 2D = 2 dimensional; 3D = 3 dimensional; CA = cardiac amyloidosis; CMP = cardiomyopathy; DLS = deep learning-derived strain; GLS = global longitudinal strain; HHD = hypertensive heart disease.

Stanford's sports cardiology clinic, hypertrophic cardiomyopathy clinic, cardiac amyloidosis clinic, and with a diagnosis of hypertension. DLS was applied in each group, and we assessed for the ability to differentiate each population from each other. In parallel, we identified a population of Cedars-Sinai patients with more severe cardiac amyloidosis and performed standard 2D-GLS and DLS in these patients to assess performance in an advanced disease population.

STATISTICAL ANALYSES. Continuous variables are expressed as mean ± SD, whereas categorical variables are reported as frequencies and percentages. For the 43 cases with 4 strain measurements performed by 2 sonographers with 2 machines, we compared the average values between the 2 readers and between measurements made on the 2 machines, reporting intraclass correlation coefficients (ICC). To assess the effect of baseline contour variability, we calculated mean absolute error and proportional differences between each translated vs initial image. In the external validation cohort, we compared DLS with 2D-GLS and 3D-GLS using ICC and Bland-Altman analyses. In addition, linear regression models were applied to assess the impact of subjective image quality on measurement agreement. We compared DLS between athletes and the 3 cohorts with LV hypertrophy of different etiologies using unpaired Student's *t*-tests. Finally, we also compared DLS with 2D-GLS in a cohort of advanced cardiac amyloidosis patients using ICC and Bland-Altman analyses. All ICC estimates were 2-way mixed-effects models (random effects: study, fixed effects: rater/method), assessing absolute agreement with single rater/measurement.[18] Statistical analyses were performed in R (version 4.1.1, R Foundation for Statistical Computing). The study was approved by Stanford University, Semmelweis University (#190/2020), and Cedars-Sinai Institutional Review Boards.

RESULTS

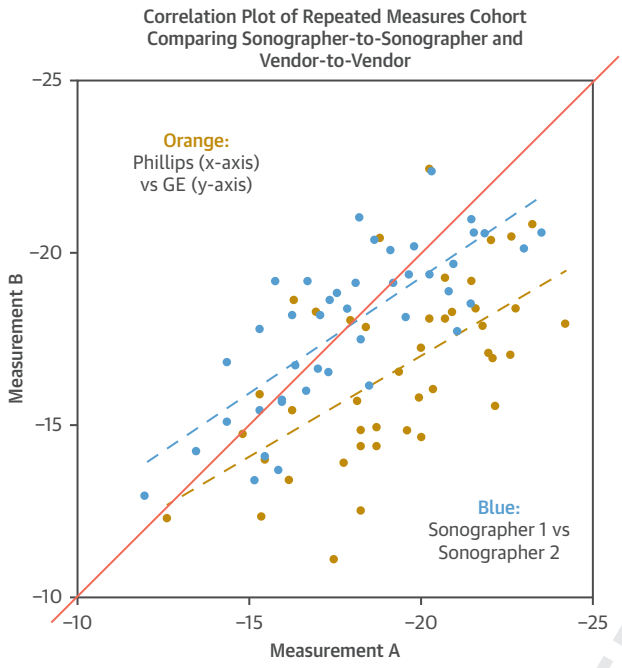
VARIABILITY OF TRADITIONAL STRAIN IN CLINICAL PRACTICE. To assess variation in measurement of traditional strain in clinical practice, we identified a cohort of 43 patients (55 ± 17 years, 38% women) to undergo repeated serial imaging on the same day. Within the cohort, 4 repeated measurements were performed for each patient, consisting of 2 sonographers performing strain analysis using both Philips and GE machines and software. The mean GLS measured by readers 1 and 2 was −18.00% ± 2.64% and -17.90% ± 2.41%, respectively. Measurements on Philips and GE had mean of −19.31% ± 1.74% and −16.58% ± 1.39%, respectively (Figure 1). The ICC between readers 1 and 2 was moderate: 0.63 (0.48-0.74), $P < 0.001$. The ICC between GE and Philips systems was 0.29 (-0.01 to 0.53), $P = 0.03$, showing poor agreement across the different videos of the same patients.

SIMULATION OF PERFORMANCE PERTURBED BY VARIATION IN INITIAL CONTOUR. 2D-GLS using the initial (ie, the reference) contour was -22.92% ± 3.57%. The mean absolute error increased with each shift, from a 1-pixel shift (mean absolute difference, 1.93%; proportional difference vs. the mean, 8.19%) to a 6-pixel shift (mean absolute difference, 3.48%; proportional difference vs. the mean, 13.52%), suggesting intolerance for initial contour tracing variability (Figure 2).

2D AND 3D EXTERNAL VALIDATION OF THE DLS ALGORITHM. In the external center, we initially identified 3,888 videos from 1,017 examinations of 907 patients that seemed suitable for analysis. Nevertheless, several videos had to be excluded due to the following reasons (Supplemental Figure 1): 1) overt changes in the imaging field settings (eg, depth change or rotation; n = 76); 2) acquired from an right ventricle-focused view inadequately capturing the LV (n = 835); 3) 1 or more LV segments not captured within the imaging sector (n = 202), 4) significant foreshortening (n = 88), 5) failure of contouring due to other issues (n = 182), or 6) unfeasible manual assessment of 2D-GLS or 3D-GLS (n = 147). Accordingly, the final external validation set comprised 2,358 videos from 894 examinations of 791 patients (46 ± 24 years, 37% women) (Table 1).

2D-GLS ranged from −1.30% to −31.12%, with a mean of −16.78% ± 5.38%, 3D-GLS ranged from −1.16% to −31.83%, with a mean of −17.75% ± 4.91%, whereas DLS ranged from −6.10% to −38.94%, with a mean of −20.10% ± 5.07%. The ICC between DLS and 2D-GLS measurements was 0.56 (0.19-0.74) ($P = 0.002$), with a bias of −3.31%, limits of agreement (LOA) −11.65% to 5.02%, and mean absolute error (MAE) of 4.29% (Figure 3A). Between DLS and 3D-GLS, the ICC was 0.58 (0.37-0.71), $P < 0.001$, with a bias of −2.35%, LOA of −10.62% to 5.91%, and MAE of 3.78% (Figure 3B). Interestingly, although the correlation between DLS and both GLS measurements was similar, the absolute difference between DLS and 2D-GLS was significantly associated with subjective image quality (ß, −0.89; SE, 0.10; $P < 0.001$), whereas the absolute difference between DLS and 3D-GLS was not (ß, 0.12; SE, 0.10; $P = 0.23$) (Supplemental Figure 2). No difference in performance was seen in patients with recordings acquired using GE vs. Philips ultrasound systems.

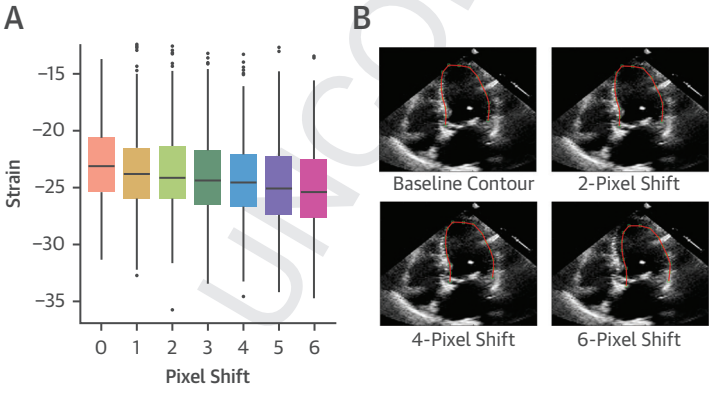
FIGURE 1 ■■■



Interreader (blue) and intervendor (orange) variability of 2-dimensional echocardiography-derived global longitudinal strain measurements. The identity line (ie, the line of equality) is shown in red.

PERFORMANCE ON CLINICAL COHORTS. Strain has been previously shown to vary in different forms of cardiomyopathy.[19-22] We identified echocardiography videos from 64 athletes, 1,590 patients with hypertrophic cardiomyopathy, 2,580 with hypertension, and 1,104 with amyloidosis and performed DLS. The DLS in the athletes was the greatest (−20.80% ± 3.86%), followed by patients with hypertrophic cardiomyopathy (−17.75% ± 4.44%), hypertension (−17.21% ± 3.99%), and amyloidosis (−16.37% ± 4.31%) in decreasing order. Comparing the athletes with the other disease cohorts showed a significant difference for all 3 disease cohorts compared with both the athlete cohort, as well as between populations ($P < 0.0001$ for all) (Figure 4). Noting that we were unable to retrospectively perform GLS on these patients, we also analyzed an additional external validation population of 102 advanced cardiac amyloid disease patients with a total of 186 echocardiography studies. The 2D-GLS mean was −9.40 ± 3.64 and DLS was −9.08 ± 2.89. ICC was 0.64 (0.53-0.72), $P < 0.001$, with a bias of 0.57%, LOA of −4.87 to 6.01, and MAE of 2.27% (Figure 5).

FIGURE 2 **Tolerance Testing of Speckle-Tracking Methods by Shifting End-Diastolic Contour in the Positive X-Axis Direction**



(A) Mean and SD by box plot for translation of baseline contour by pixel quantities. (B) Visual comparison of pixel shifting from baseline to 6-pixel shift.

DISCUSSION

In this article, we present a vendor-independent, open-source, deep learning-based strain algorithm, which can be used to modify semantic segmentation models. In our experiments, we demonstrate traditional strain has significant intervendor variation and even minor variation in initial contour results in significantly different measurements of GLS. In contrast, frame-by-frame contouring of the DLS method may reduce the impact of single-frame errors and avoids inherent manual variation. We compared the DLS values with conventional GLS measurement on cohorts of patients to show good concordance as well as the ability to distinguish between types of cardiomyopathy. Our DLS method performed consistently and within the variation seen between different commercial GLS vendors.

In the cohort of patients undergoing repeated GLS evaluation by different sonographers and using different ultrasound systems, we continued to see significant variation in measuring strain. This variation may be explained by how nearly visually imperceptible changes in initial LV contour can result in large changes in measurement of GLS (eg, a 12.58% difference at a 5-pixel shift in simulation experiments). We subsequently developed our own fully automated open-source deep learning-based algorithm to calculate strain, and show that our DLS algorithm correlated well with expert measurements of 2D and 3D strain on an external validation cohort.

We noted that our algorithm tended to overestimate strain in the large external cohort. Although some of the bias is inherent to the DLS algorithm, as it

TABLE 1 Demographic, Clinical, and Echocardiographic Characteristics of the Training and External Validation Datasets

	Original Training Dataset (n = 7,465, 7,465 Videos)	External Validation Dataset (n = 7991, 2,358 Videos)
Age, y	70 ± 22	46 ± 24
Female	3,662 (49)	290 (37)
Heart failure	2,113 (28)	235 (30)
Diabetes mellitus	1,474 (20)	108 (14)
Hypercholesterolemia	2,463 (33)	221 (28)
Hypertension	2,912 (39)	310 (39)
Renal disease	1,475 (20)	112 (14)
Coronary artery disease	1,674 (22)	102 (13)
End-diastolic volume, mL	91.0 ± 46.0	145.2 ± 58.8[a]
End-systolic volume, mL	43.2 ± 36.1	71.9 ± 49.1[a]
Ejection fraction, %	55.7 ± 12.5	53.4 ± 12.3[a]
Echocardiographic machine		
Philips Epiq 7C	4,832 (65)[b]	328 (14)[b]
Philips Epiq 7G	0 (0)[b]	546 (23)[b]
Philips iE33	2,489 (33)[b]	96 (4)[b]
Philips CX50	62 (1)[b]	0 (0)[b]
Philips Epiq 5G	44 (1)[b]	0 (0)[b]
GE Vivid E95	0 (0)[b]	1,378 (58)[b]
Other	38 (1)[b]	10 (0)[b]

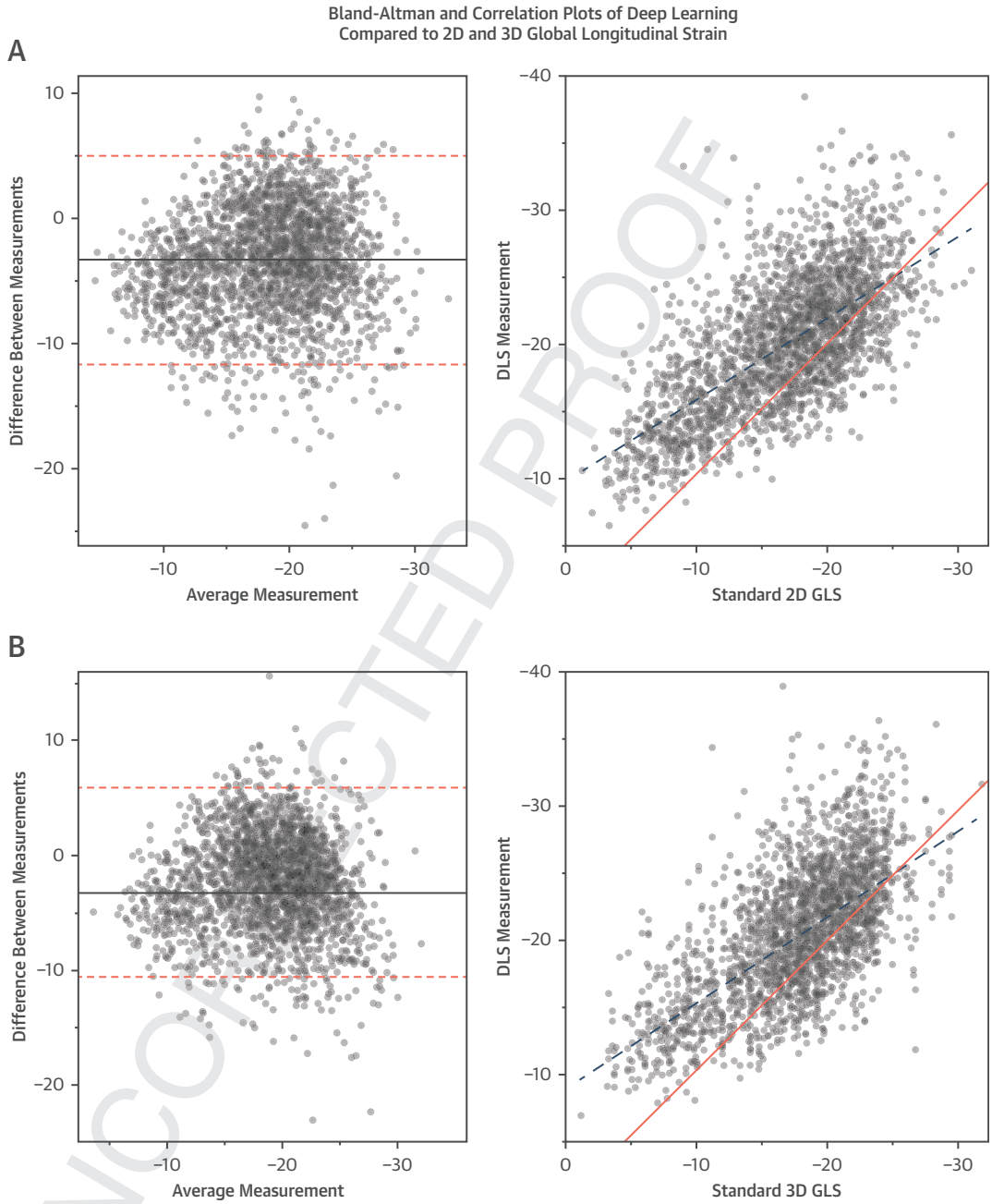
Values are mean ± SD or n (%). [a]Measured with 3-dimensional echocardiography. [b]The number and percentage of videos acquired using the given echocardiographic machine.

has an operating and calibration range, we also observed specific errors may worsen results at extremes of function. In patients with hyperdynamic LV function, segmentation might be challenging for the deep learning algorithm, particularly if mid-chamber obliteration occurs, as the algorithm may only obtain part of the LV cavity because segmentation is typically restricted to a single shape. At lower extremes of function, image quality may be reduced by anatomical deviations due to disease, and the accurate selection of end-systolic and end-diastolic frames on the smoothed endocardial length curve may be more difficult. For example, in patients with minimal change in the contour length, DLS may be overestimated as the algorithm selects the numerically greatest differences instead of the physiological identification of end-systole and -diastole. In addition, in patients with smaller LV, these errors may be magnified. Notably, image quality is an important factor, as worse image quality was associated with weaker ICC and higher absolute error values between DLS and 2D-GLS. Thus, it may be reasonable to visually assess the quality of the segmentation or the selection of end-diastolic and end-systolic frames, especially in patients at the extremes of function, and to exercise greater caution when attempting to directly compare DLS measurements with GLS measurements than when comparing within DLS measurements.

We demonstrated that DLS was significantly different between the cohorts of phenotypic mimics of LV hypertrophy; however, we note the high degree of overlap between the distributions in strain values, particularly between the cardiac amyloidosis, hypertrophic cardiomyopathy, and hypertension cohorts (Figure 4). Several prior studies have applied standard speckle-tracking strain in attempting to differentiate these groups from healthy controls and, in some cases, from each other.[23,24] In general, statistically significant differences were observed, particularly between athletes, hypertensive heart disease, and hypertrophic cardiomyopathy, but frequently with overlapping distributions, suggesting a potential incremental value of 2D-GLS but not its use as a standalone diagnostic approach, similar to our study.[25,26] Other studies leverage layer-specific (eg, either endocardial or epicardial strain in hypertrophic cardiomyopathy) or segmental strains (eg, apical vs. basal strain patterns in amyloidosis) to improve diagnosis,[27-30] which our algorithm is unable to provide at this time. We also noted that other studies reported more severely impaired GLS values in hypertrophic cardiomyopathy or cardiac amyloidosis than those we observed in our cohorts,[24,31,32] which we believe is attributable to the differences in selection criteria. Our cohorts were drawn from subspecialty referral clinics and may include an earlier range of clinical disease status, vs potentially more advanced disease in the previous literature.[27-30] Unfortunately, GLS was not measured in our clinical cohorts; thus, our ability to confirm this is limited. Nonetheless, to assess the performance of our DLS algorithm in more advanced diseases, we also prospectively analyzed a cohort of patients with advanced cardiac amyloidosis who had strain values more in line with prior literature. Agreement results were similar to those in the external cohort, with a mild bias toward lower strain values and a moderate ICC, suggesting that the performance of the DLS algorithm is preserved in this more advanced clinical cohort.

Deep learning approaches to semantic segmentation are robust to even poor video quality,[13] and could potentially measure strain in low-quality clinical videos where speckle-tracking or optical flow fails. In addition, this pipeline could be adapted to other semantic segmentation models, allowing generalization to calculate right ventricular or atrial strain. There have been previous publications regarding deep

FIGURE 3 Comparison of 2D and 3D Echocardiography-based GLS vs DLS on External Validation Cohort by Bland-Altman and Correlation Plots



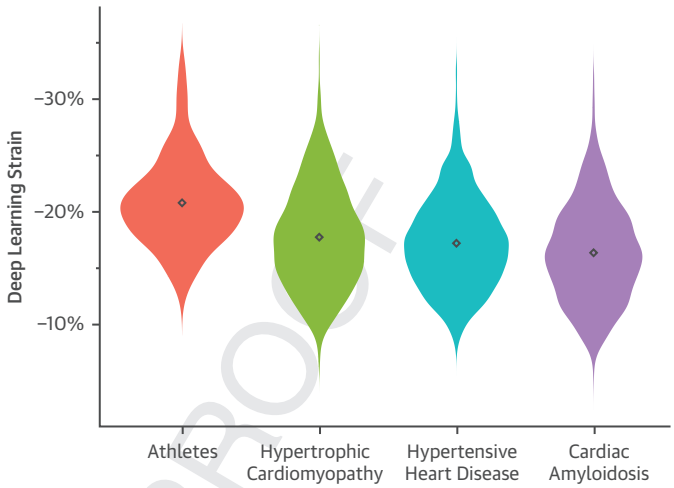
(A) 2D-GLS vs DLS. (B) 3D-GLS vs DLS. 2D = 2 dimensional; 3D = 3 dimensional; DLS = deep learning–derived strain; GLS = global longitudinal strain.

learning-based strain, both in echocardiography and cardiac magnetic resonance imaging. Focusing on the echocardiography approaches, a series of manuscripts by Østvik and Salte and colleagues[33-36] has presented the development of a deep learning-based approach, culminating clinical validation with excellent performance. Their approach leverages a series of neural networks to perform view classification, identify the

timings in the cardiac cycle, perform 4-class segmentation (LV blood pool, LV myocardium, left atrial myocardium, and background), and apply an optical flow field to calculate strain. This approach has higher complexity than ours, which uses a single neural network performing 2-class segmentation (LV blood pool and background) in a deterministic pipeline to assess GLS. The simplicity of our current approach has its advantages (eg, potentially broader/simpler implementation and lower computation requirements) but comes at the expense of slightly lower accuracy compared with the performance of Østvik and Salte's method, although they performed validation only in a smaller internal cohort.[33] Similar publications leveraging optical-flow-focused neural networks have been trialed in smaller populations, with good agreement.[37-39] Our algorithm enables rapid retrospective batch analysis of echocardiographic videos, which may have applications in both research and clinical workflow while eliminating variance related to human-related measurement. Reference ranges of strain are specific to each vendor and its strain analysis package, which can be limiting given the black-box nature of those software solutions. Our DLS algorithm produces comparable strain values within the reference range suggested by professional societies for healthy individuals and is open-source and fully automated, allowing the adaptation, iterative improvement, and easy establishment of normal ranges using local data. Given the lack of algorithmic transparency of the commercially available speckle-tracking packages, we showed our algorithm provided similar results to standard approaches in a large and heterogeneous cohort. The direct measurement and identification of the endocardial contour are both easily understandable and can be visually inspected for sources of error. In addition, the open-source methods of our technique and development using publicly available datasets may improve understandability over proprietary methods.

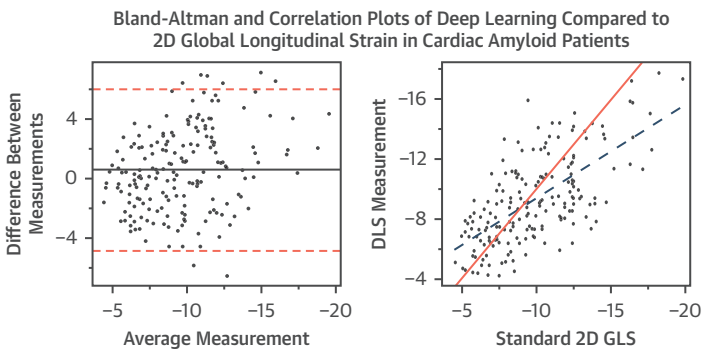
There are a few potential limitations to our methodology. Given the dependence on appropriate deep learning segmentation, we excluded videos/segmentations for which deep learning appeared grossly inaccurate. Noting that the videos used were not acquired specifically with strain in mind, not all cases were able to be accurately contoured, particularly if the video in question did not show all segments of the LV, foreshortening occurred during the clip, or the lateral wall moved outside of the ultrasound sector during a cardiac cycle, which was the case in many segmentation failures resulting in multiple cases being excluded. The tuning of the pipeline parameters (ie, selection of a Savitzky-Golay filter and 3-pixel dilation) was empiric, and although they attempt to approximate anatomical aspects such as beat-to-beat variability and mid-myocardial contours, they are imperfect measures for reflecting patient-level anatomy and physiology. Inherent differences in methodology are present: our proof-of-concept study relies on the A4C view alone, whereas standard 2D-GLS uses all 3 apical views, and 3D-GLS uses 3D acquisitions. However, the measurements are comparable, and the methodology and the codebase are

FIGURE 4 Violin Plots of Clinical Groups' Deep Learning Strain Values



Athletes: −20.80% ± 3.86%, hypertrophic cardiomyopathy: −17.75% ± 4.44%, hypertension: −17.21% ± 3.99%, amyloidosis: −16.37% ± 4.31%). $P < 0.001$ for difference between all groups.

FIGURE 5 Comparison of 2D Echocardiography-based GLS vs DLS on Internal Advanced Amyloidosis Cohort by Bland-Altman and Correlation Plots



Abbreviations as in Figure 3.

easily generalizable to other views. In addition, our study focused on a single ventricular measurement of GLS, without sub-segmentation by region, which may perform differently from block-matching/optical flow-based approaches.

In conclusion, we present a deep learning-derived strain measurement based on a deep learning-derived endocardial contour. The results show that this measurement can be performed reliably with low variance and within the range of standard measurements. In parallel, we show even subtle changes to input images can significantly change strain calculated by speckle tracking, and highlight the importance of high image quality for conventional strain approaches. The DLS method is rapid, consistent, vendor-agnostic, publicly available, and applicable across a wide range of image qualities.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

Dr Kwan has received funding support from the Doris Duke Charitable Foundation Grant 2020059, American Heart Association Career Development Award 23CDA1053659 and National Institutes of Health (NIH) R01HL131532-06. Dr Chang has received funding from National Institutional Health T32HL097129-13. This study was partially supported by the National Research, Development, and Innovation Fund of Hungary within the framework of the Artificial Intelligence National Laboratory Program and by the Thematic Excellence Program (2020-4.1.1.-TKP2020) of the Ministry for Innovation and Technology in Hungary, within the framework of the Bioimaging thematic program of the Semmelweis University. This work was supported in part by NIH K99 HL157421-01. All other authors have reported that they have no relationships relevant to the contents of this paper to report.

ADDRESS FOR CORRESPONDENCE: Dr Alan C. Kwan, Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, AHSP Pavilion Suite A3100, 127 S. San Vicente Boulevard, Los Angeles, California 90048, USA. E-mail: David.Ouyang@cshs.org; alan.kwan@cshs.org. @alanckwan, @David_Ouyang.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: This study quantifies strain agreement by repeated measures with multiple vendors and sonographers, showing that intervendor agreement is especially poor. This study highlights the potential role that an open-source, non-vendor-associated retrospective strain measurement could play in improving patient management. This method is applied within clinical datasets demonstrating acceptable agreement with standard methods and potential clinical applications.

TRANSLATIONAL OUTLOOK: The code and training data for the study are made publicly available. Further development of the DLS method to assess multiple views for GLS, improved tolerance of suboptimal image quality, and clinical integration are essential future steps.

REFERENCES

1. Amzulescu MS, De Craene M, Langet H, et al. Myocardial strain imaging: review of general principles, validation, and sources of discrepancies. *Eur Heart J Cardiovasc Imaging*. 2019;20:605–619.

2. Plana JC, Galderisi M, Barac A, et al. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*. 2014;15:1063–1093.

3. Krishnasamy R, Isbel NM, Hawley CM, et al. Left ventricular global longitudinal strain (GLS) is a superior predictor of all-cause and cardiovascular mortality when compared to ejection fraction in advanced chronic kidney disease. *PLoS One*. 2015;10:e0127044.

4. Sengeløv M, Jørgensen PG, Jensen JS, et al. Global longitudinal strain is a superior predictor of all-cause mortality in heart failure with reduced ejection fraction. *J Am Coll Cardiol Img*. 2015;8:1351–1359.

5. Ng AC, Prihadi EA, Antoni ML, et al. Left ventricular global longitudinal strain is predictive of all-cause mortality independent of aortic stenosis severity and ejection fraction. *Eur Heart J Cardiovasc Imaging*. 2018;19:859–867.

6. D'hooge J, Heimdal A, Jamal F, et al. Regional strain and strain rate measurements by cardiac ultrasound: principles, implementation and limitations. *Eur J Echocardiogr*. 2000;1:154–170.

7. Claus P, Omar AMS, Pedrizzetti G, Sengupta PP, Nagel E. Tissue tracking technology for assessing cardiac mechanics: principles, normal values, and clinical applications. *J Am Coll Cardiol Img*. 2015;8:1444–1460.

8. Chan J, Shiino K, Obonyo NG, et al. Left ventricular global strain analysis by two-dimensional speckle-tracking echocardiography: the learning curve. *J Am Soc Echocardiogr*. 2017;30:1081–1090.

9. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*. 2015;16:233–271.

10. Voigt J-U, Pedrizzetti G, Lysyansky P, 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.

11. Kwan AC, Salto G, Cheng S, Ouyang D. Artificial intelligence in computer vision: cardiac MRI and multimodality imaging segmentation. *Curr Cardiovasc Risk Rep*. 2021;15:1–8.

12. Chang E. Agnostic strain. 2023. https://github.com/Ernest-Chang/vendorAgnosticStrain

13. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*. 2020;580:252–256.

14. Ouyang D. EchoNet-Dynamic. 2020. https://echonet.github.io/dynamic/

15. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17:261–272.

16. Van Rossum G, Drake FL Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam; 1995.

17. Ouyang D. Strain analysis. 2022. https://github.com/echonet/strain

18. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163.

19. Tower-Rader A, Mohananey D, To A, Lever HM, Popovic ZB, Desai MY. Prognostic value of global longitudinal strain in hypertrophic cardiomyopathy: a systematic review of existing literature. *J Am Coll Cardiol Img*. 2019;12:1930–1942.

20. Migrino RQ, Aggarwal D, Konorev E, Brahmbhatt T, Bright M, Kalyanaraman B. Early detection of doxorubicin cardiomyopathy using two-dimensional strain echocardiography. *Ultrasound Med Biol*. 2008;34:208–214.

21. Koyama J, Ray-Sequin PA, Falk RH. Longitudinal myocardial function assessed by tissue velocity, strain, and strain rate tissue Doppler echocardiography in patients with AL (primary) cardiac amyloidosis. *Circulation*. 2003;107:2446–2452.

22. Haugaa KH, Goebel B, Dahlslett T, et al. Risk assessment of ventricular arrhythmias in patients with nonischemic dilated cardiomyopathy by strain echocardiography. *J Am Soc Echocardiogr*. 2012;25:667–673.

23. Tanaka H. Efficacy of echocardiography for differential diagnosis of left ventricular hypertrophy: special focus on speckle-tracking longitudinal strain. *J Echocardiogr*. 2021;19:71–79.

24. D'Andrea A, Radmilovic J, Ballo P, et al. Left ventricular hypertrophy or storage disease? The incremental value of speckle-tracking strain bull's-eye. *Echocardiography*. 2017;34:746–759.

25. Neisius U, Myerson L, Fahmy AS, et al. Cardiovascular magnetic resonance feature tracking strain analysis for discrimination between hypertensive heart disease and hypertrophic cardiomyopathy. *PLoS One*. 2019;14:e0221061.

26. Saghir M, Areces M, Makan M. Strain rate imaging differentiates hypertensive cardiac hypertrophy from physiologic cardiac hypertrophy (athlete's heart). *J Am Soc Echocardiogr*. 2007;20:151–157.

27. Phelan D, Collier P, Thavendiranathan P, et al. Relative apical sparing of longitudinal strain using two-dimensional speckle-tracking echocardiography is both sensitive and specific for the diagnosis of cardiac amyloidosis. *Heart*. 2012;98:1442–1448.

28. Baccouche H, Maunz M, Beck T, et al. Differentiating cardiac amyloidosis and hypertrophic cardiomyopathy by use of three-dimensional speckle-tracking echocardiography. *Echocardiography*. 2012;29:668–677.

29. Sun JP, Stewart WJ, Yang XS, et al. Differentiation of hypertrophic cardiomyopathy and cardiac amyloidosis from other causes of ventricular wall thickening by two-dimensional strain imaging echocardiography. *Am J Cardiol*. 2009;103:411–415.

30. Sun JP, Xu TY, Ni XD, et al. Echocardiographic strain in hypertrophic cardiomyopathy and hypertensive left ventricular hypertrophy. *Echocardiography*. 2019;36:257–265.

31. Kato TS, Noda A, Izawa H, et al. Discrimination of nonobstructive hypertrophic cardiomyopathy from hypertensive left ventricular hypertrophy on the basis of strain rate imaging by tissue Doppler ultrasonography. *Circulation*. 2004;110:3808–3814.

32. Butz T, van Buuren F, Mellwig KP, et al. Two-dimensional strain analysis of the global and regional myocardial function for the differentiation of pathologic and physiologic left ventricular hypertrophy: a study in athletes and in patients with hypertrophic cardiomyopathy. *Int J Cardiovasc Imaging*. 2011;27:91–100.

33. Salte Ivar M, Østvik A, Smistad E, et al. Artificial intelligence for automatic measurement of left ventricular strain in echocardiography. *J Am Coll Cardiol Img*. 2021;14:1918–1928.

34. Østvik A, Salte IM, Smistad E, et al. Myocardial function imaging in echocardiography using deep learning. *IEEE Trans Med Imaging*. 2021;40:1340–1351.

35. Østvik A, Smistad E, Espeland T, Berg EAR, Lovstakken L. Automatic myocardial strain imaging in echocardiography using deep learning. In: Stoyanov D, Taylor Z, Carneiro G, et al., eds. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer International Publishing; 2018:309–316.

36. Salte IM, Østvik A, Olaisen SH, et al. Deep learning for improved precision and reproducibility of left ventricular strain in echocardiography: a test-retest study. *J Am Soc Echocardiogr*. 2023;36:788–799.

37. Deng Y, Cai P, Zhang L, et al. Myocardial strain analysis of echocardiography based on deep learning. *Front Cardiovasc Med*. 2022;9:1067760.

38. Evain E, Sun Y, Faraz K, et al. Motion estimation by deep learning in 2D echocardiography: synthetic dataset and validation. *IEEE Trans Med Imaging*. 2022;41:1911–1924.

39. Romero-Pacheco A, Perez-Gonzalez J, Hevia-Montiel N. Estimating echocardiographic myocardial strain of left ventricle with deep learning. *Annu Int Conf IEEE Eng Med Biol Soc*. 2022;2022:3891–3894.

KEY WORDS deep learning, echocardiography, longitudinal strain

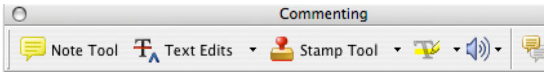
APPENDIX For supplemental figures, please see the online version of this paper.

ELSEVIER

INSTRUCTIONS ON THE ANNOTATION OF PDF FILES

To view, print and annotate your content you will need Adobe Reader version 9 (or higher). This program is freely available for a whole series of platforms that include PC, Mac, and UNIX and can be downloaded from http://get.adobe.com/reader/. The exact system requirements are given at the Adobe site: http://www.adobe.com/products/reader/tech-specs.html.

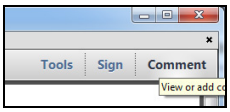
Note: ***Please do NOT make direct edits to the PDF using the editing tools as doing so could lead us to overlook your desired changes.*** *Rather, please request corrections by using the tools in the Comment pane to annotate the PDF and call out the changes you are requesting. If you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file.*

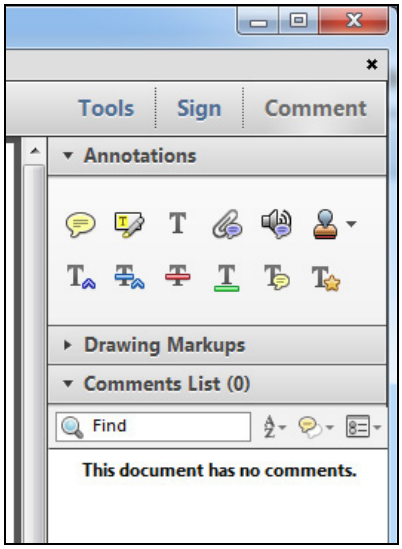
<table>
<tr><th colspan="2">PDF ANNOTATIONS</th></tr>
<tr><th>Adobe Reader version 9</th><th>Adobe Reader version X and XI</th></tr>
<tr><td>When you open the PDF file using Adobe Reader, the Commenting tool bar should be displayed automatically; if not, click on ‘Tools’, select ‘Comment & Markup’, then click on ‘Show Comment & Markup tool bar’ (or ‘Show Commenting bar’ on the Mac). If these options are not available in your Adobe Reader menus then it is possible that your Adobe Acrobat version is lower than 9 or the PDF has not been prepared properly.
 

(Mac)
PDF ANNOTATIONS (Adobe Reader version 9)

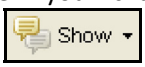
The default for the Commenting tool bar is set to ‘off’ in version 9. To change this setting select ‘Edit | Preferences’, then ‘Documents’ (at left under ‘Categories’), then select the option ‘Never’ for ‘PDF/A View Mode’.
 

(Changing the default setting, Adobe version 9)</td><td>To make annotations in the PDF file, open the PDF file using Adobe Reader XI, click on ‘Comment’.

If this option is not available in your Adobe Reader menus then it is possible that your Adobe Acrobat version is lower than XI or the PDF has not been prepared properly.
 

This opens a task pane and, below that, a list of all Comments in the text. These comments initially show all the changes made by our copyeditor to your file.
 
</td></tr>
</table>

HOW TO...		
Action	**Adobe Reader version 9**	**Adobe Reader version X and XI**
Insert text	Click the 'Text Edits' button  on the Commenting tool bar. Click to set the cursor location in the text and simply start typing. The text will appear in a commenting box. You may also cut-and-paste text from another file into the commenting box. Close the box by clicking on 'x' in the top right-hand corner.	Click the 'Insert Text' icon  on the Comment tool bar. Click to set the cursor location in the text and simply start typing. The text will appear in a commenting box. You may also cut-and-paste text from another file into the commenting box. Close the box by clicking on '_'  in the top right-hand corner.
Replace text	Click the 'Text Edits' button  on the Commenting tool bar. To highlight the text to be replaced, click and drag the cursor over the text. Then simply type in the replacement text. The replacement text will appear in a commenting box. You may also cut-and-paste text from another file into this box. To replace formatted text (an equation for example) please Attach a file (see below).	Click the 'Replace (Ins)' icon  on the Comment tool bar. To highlight the text to be replaced, click and drag the cursor over the text. Then simply type in the replacement text. The replacement text will appear in a commenting box. You may also cut-and-paste text from another file into this box. To replace formatted text (an equation for example) please Attach a file (see below).
Remove text	Click the 'Text Edits' button  on the Commenting tool bar. Click and drag over the text to be deleted. Then press the delete button on your keyboard. The text to be deleted will then be struck through.	Click the 'Strikethrough (Del)' icon  on the Comment tool bar. Click and drag over the text to be deleted. Then press the delete button on your keyboard. The text to be deleted will then be struck through.
Highlight text/ make a comment	Click on the 'Highlight' button  on the Commenting tool bar. Click and drag over the text. To make a comment, double click on the highlighted text and simply start typing.	Click on the 'Highlight Text' icon  on the Comment tool bar. Click and drag over the text. To make a comment, double click on the highlighted text and simply start typing.
Attach a file	Click on the 'Attach a File' button  on the Commenting tool bar. Click on the figure, table or formatted text to be replaced. A window will automatically open allowing you to attach the file. To make a comment, go to 'General' in the 'Properties' window, and then 'Description'. A graphic will appear in the PDF file indicating the insertion of a file.	Click on the 'Attach File' icon  on the Comment tool bar. Click on the figure, table or formatted text to be replaced. A window will automatically open allowing you to attach the file. A graphic will appear indicating the insertion of a file.
Leave a note/ comment	Click on the 'Note Tool' button  on the Commenting tool bar. Click to set the location of the note on the document and simply start typing. Do not use this feature to make text edits.	Click on the 'Add Sticky Note' icon  on the Comment tool bar. Click to set the location of the note on the document and simply start typing. Do not use this feature to make text edits.

HOW TO...		
Action	**Adobe Reader version 9**	**Adobe Reader version X and XI**
Review	To review your changes, click on the 'Show' button  on the Commenting tool bar. Choose 'Show Comments List'. Navigate by clicking on a correction in the list. Alternatively, double click on any mark-up to open the commenting box.	Your changes will appear automatically in a list below the Comment tool bar. Navigate by clicking on a correction in the list. Alternatively, double click on any mark-up to open the commenting box.
Undo/delete change	To undo any changes made, use the right click button on your mouse (for PCs, Ctrl-Click for the Mac). Alternatively click on 'Edit' in the main Adobe menu and then 'Undo'. You can also delete edits using the right click (Ctrl-click on the Mac) and selecting 'Delete'.	To undo any changes made, use the right click button on your mouse (for PCs, Ctrl-Click for the Mac). Alternatively click on 'Edit' in the main Adobe menu and then 'Undo'. You can also delete edits using the right click (Ctrl-click on the Mac) and selecting 'Delete'.

SEND YOUR ANNOTATED PDF FILE BACK TO ELSEVIER

Save the annotations to your file and return as instructed by Elsevier. Before returning, please ensure you have answered any questions raised on the Query Form and that you have inserted all corrections: later inclusion of any subsequent corrections cannot be guaranteed.

FURTHER POINTS

- Any (grey) halftones (photographs, micrographs, etc.) are best viewed on screen, for which they are optimized, and your local printer may not be able to output the greys correctly.
- If the PDF files contain colour images, and if you do have a local colour printer available, then it will be likely that you will not be able to correctly reproduce the colours on it, as local variations can occur.
- If you print the PDF file attached, and notice some 'non-standard' output, please check if the problem is also present on screen. If the correct printer driver for your printer is not installed on your PC, the printed output will be distorted.