



Original Article

Intra- and inter-operator reproducibility of automated cloud-based carotid lumen diameter ultrasound measurement



Luca Saba^a, Sumit K. Banchhor^b, Tadashi Araki^c, Klaudija Viskovic^d,
Narendra D. Londhe^b, John R. Laird^e, Harman S. Suri^f, Jasjit S. Suri^{f,g,h,*}

^a Department of Radiology, University of Cagliari, Italy

^b Department of Electrical Engineering, NIT Raipur, Chhattisgarh, India

^c Division of Cardiovascular Medicine, Toho University Ohashi Medical Center, Tokyo, Japan

^d Department of Radiology and Ultrasound, University Hospital for Infectious Disease, Zagreb, Croatia

^e UC Davis Vascular Centre, University of California, Davis, CA, USA

^f Monitoring and Diagnostic Division, AtheroPoint[®], Roseville, CA, USA, USA

^g Point-of-Care Devices, Global Biomedical Technologies, Inc., Roseville, CA, USA

^h Department of Electrical Engineering, University of Idaho (Affil.), ID, USA

ARTICLE INFO

Article history:

Received 27 September 2017

Accepted 14 January 2018

Available online 31 January 2018

Keywords:

Atherosclerosis
Carotid
Cloud-based
Reproducibility
Reliability

ABSTRACT

Background: Common carotid artery lumen diameter (LD) ultrasound measurement systems are either manual or semi-automated and lack reproducibility and variability studies. This pilot study presents an automated and cloud-based LD measurements software system (AtheroCloud) and evaluates its: (i) intra/inter-operator reproducibility and (ii) intra/inter-observer variability.

Methods: 100 patients (83 M, mean age: 68 ± 11 years), IRB approved, consisted of L/R CCA artery (200 ultrasound images), acquired using a 7.5-MHz linear transducer. The intra/inter-operator reproducibility was verified using three operator's readings. Near-wall and far carotid wall borders were manually traced by two observers for intra/inter-observer variability analysis.

Results: The mean coefficient of correlation (CC) for intra- and inter-operator reproducibility between all the three automated reading pairs were: 0.99 ($P < 0.0001$) and 0.97 ($P < 0.0001$), respectively. The mean CC for intra- and inter-observer variability between both the manual reading pairs were 0.98 ($P < 0.0001$) and 0.98 ($P < 0.0001$), respectively. The Figure-of-Merit between the mean of the three automated readings against the four manuals were 98.32%, 99.50%, 98.94% and 98.49%, respectively.

Conclusions: The AtheroCloud LD measurement system showed high intra/inter-operator reproducibility hence can be adapted for vascular screening mode or pharmaceutical clinical trial mode.

© 2018 Published by Elsevier B.V. on behalf of Cardiological Society of India. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Carotid artery disease (CAD) and stroke have been predicted as the foremost cause of death in the United States and worldwide.¹ Progression of atherosclerosis disease² blocks the arteries³ limiting the oxygenated blood flow to the brain (see Fig. 1). Rupture of the arterial cap can eventually block the carotid arteries, leading to stroke.^{4–6} Ultrasound (US) examination is mostly recommended by the American Society of Echocardiography for patients with cardiovascular risk. US-based measurements are preferred over other imaging modalities (CT, MRI, and OCT), as it is safe, less expensive, less time consuming and provides real-time data.^{7,8}

The distance between near and far-wall borders in the carotid artery is called lumen diameter (LD).⁹ Recent studies had correlated LD with the stroke events.^{10,11} There are several reasons for automated LD measurements and its clinical significance. First, it helps in direct estimation of blockage of the artery along the common carotid (CCA) or internal carotid artery (ICA). Second, the LD estimation helps in evaluating the number of bumps along the CCA/ICA, which is indicative of the aggressiveness of the atherosclerosis disease. Third, the minimum LD estimation along the artery shows the growth of the arterial disease at that particular location, which can then be used for monitoring the plaque buildup for that spots. Fourth, due to multifocal¹² nature of calcium in the carotid arteries in both near and far walls (double sided), a decrease in LD can predict the stenosis severity. Fifth, an accurate end-diastolic LD is an essential parameter in measuring arterial elasticity.¹³

* Corresponding author at: Monitoring and Diagnostic Division, AtheroPoint[™], Roseville, CA, USA.

E-mail address: Jasjit.Suri@atheropoint.com (J.S. Suri).

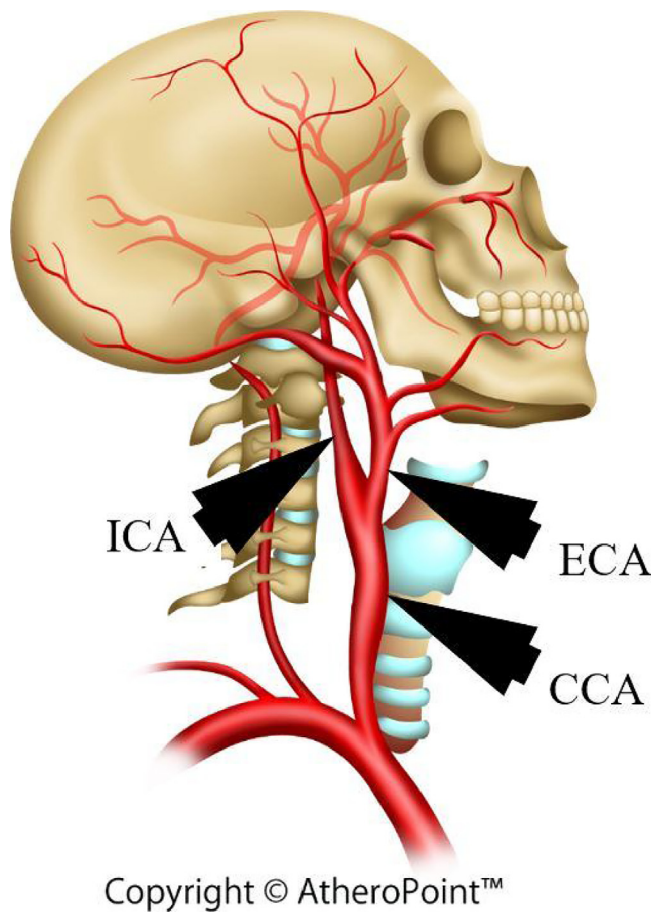


Fig. 1. (a) – Illustration of plaque formation in the carotid artery (Courtesy of AtheroPoint®, Roseville, CA, USA).

Even though carotid intima media thickness (cIMT) is more popular as compared to LD, recent findings show that LD measurement in CCA/ICA can provide equally informative risk assessment¹⁴. Thus, LD measurement can be used as a secondary validated measure which is essential during the planning of different surgical procedures such as stenting or endarterectomies^{15,16} and can be considered as a surrogate marker for risk of cardiovascular disease.

In high-resolution US images, the cardiologist can manually trace the near and far-wall borders of the CCA/ICA using calipers.¹⁷ Manual tracing of the near and far wall borders are tedious, prone to errors and have large intra- and inter-observer variability.¹⁸ Recent innovations^{19–23} have enabled us to automate the LD

computations. Previous automated LD measurement systems^{24–30} lack an “anytime-anywhere” solution and telemedicine-based model.³¹ A detailed literature survey is presented in the Discussion section. Recently, Suri and his team had proposed a completely automated cloud-based system for cIMT⁶ and LD³² computation. The proposed AtheroCloud system,³² can automatically compute LD, stenosis severity index (SSI) and total lumen area (TLA) measurement using B-mode ultrasound. This was the first system in an automatic web-based framework but lacked intra- and inter-operator reproducibility and variability analysis. This puts a restraint on the reliability of these automated LD measurement systems. These issues motivate us to demonstrate the intra- and inter-operator reproducibility for an accurate and reliable tool which can perform automated delineation of the near and far-wall borders of the CCA/ICA to measure LD in the ultrasound images. The current study is the first of its kind and is based on the hypothesis that even a novice operator can yield high intra- and inter-operator reproducibility for automated LD measurement in a cloud-based setting. Workflow of the current study is shown in Fig. 2.

Carotid stenosis is an early indicator for CAD.³³ LD is generally used to predict the stenosis severity. Recently, Polak^{10,11,34} link LD with the stroke events. Previous studies also correlate LD to cardiac events,³⁵ age³⁶ and to other risk factors like hypertension, LV mass, systolic blood pressure, gender, smoking, and hyperlipidemia.^{32,37} This clearly proves LD to be a reliable imaging biomarker in the prediction of CAD and myocardial infarction.

2. Data acquisition and preparation

2.1. Patient selection

Two hundred and four patients from July 2009 to December 2010 underwent B-mode carotid US scans. For both left and right CCA artery, a total of 407 US scans (one patient had one image missing) were obtained and retrospectively analyzed. The study was IRB approved and written informed consent was taken from the patients. Due to lack of funding, the dataset used in this study contains a random set of 100 patients (200 CCA ultrasound scans).

2.2. Ultrasound data acquisition

Carotid ultrasonographic examinations for all the patients were performed using ultrasound scanner (Aplio XV, Aplio XG, Xario) equipped with 7.5 MHz linear array transducer. The average resolution factor for carotid US scans was 0.0529 mm/pixel. Same skilled sonographer, having an experience of 15 years scanned all the patients following American Society of Echocardiography (ASE) guidelines.

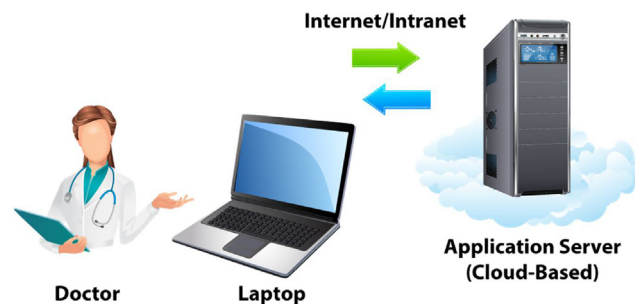


Fig. 2. (a) – Workflow of AtheroCloud and its components. The tower represents server in the cloud. The arrows represent the bi-directional flow of information (Courtesy of AtheroPoint®, Roseville, CA, USA).

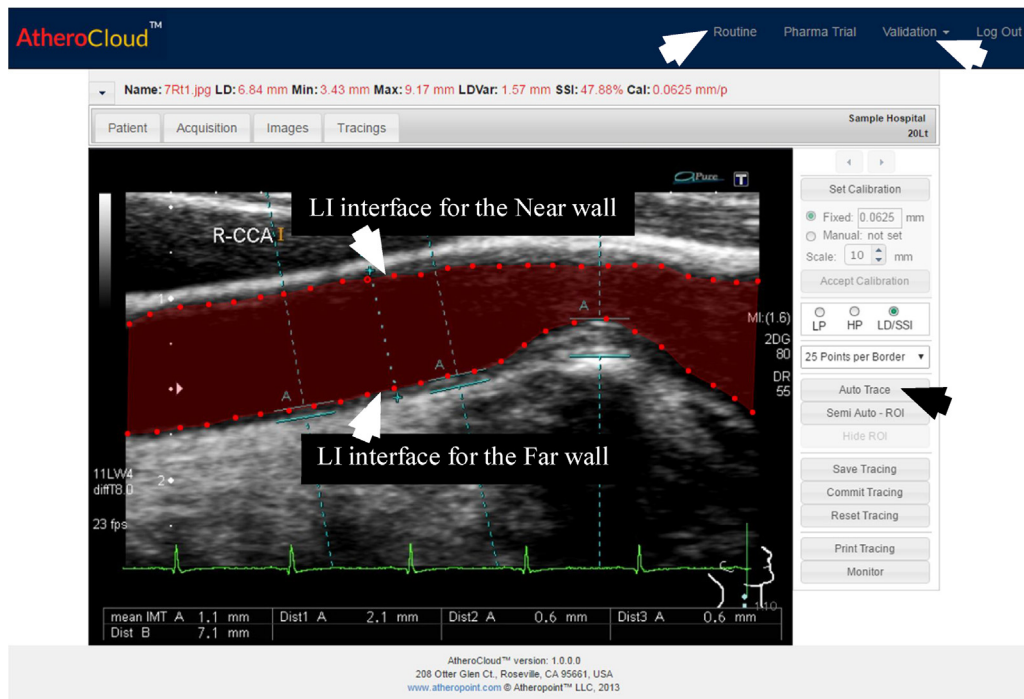


Fig. 3. Routine trial mode automated tracings (red) of the carotid LD region showing lumen-intima (LI) interface for the near wall and far wall using AtheroCloud software (courtesy of AtheroPoint[®], Roseville, CA, USA). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

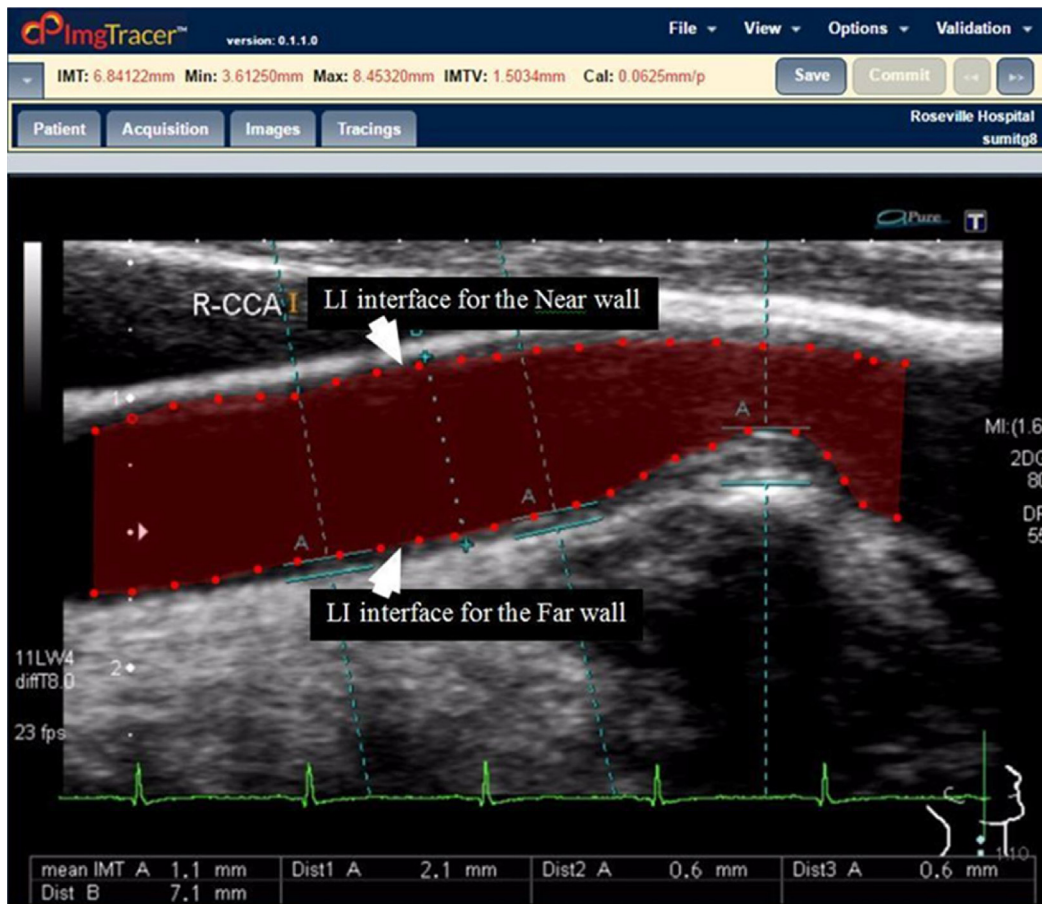


Fig. 4. (color image) – Manual tracings (red) of the carotid LD region showing lumen-intima (LI) interface for the near wall and far wall using PImgTracer software (Courtesy of AtheroPoint[®], Roseville, CA, USA). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. Patient demographics and far wall characteristics

The study sample contains 100 patients (73 M/27F); mean age 68 ± 11 ; ranging from 29 to 88 years, 50 patients had proximal lesion location, 29 at middle and 21 at a distal location. These 100 patients had a mean HbA1c, LDL, HDL and total cholesterol as: 6.40 ± 1.2 (mg/dl), 103.96 ± 31.34 (mg/dl), 51.17 ± 14.04 (mg/dl) and 179.60 ± 38.61 (mg/dl), respectively. Forty-two from the pool of one hundred were smokers.

3. Methodology

3.1. AtheroCloud software for LD computation

A sample view of AtheroCloud software is shown in Fig. 3. Average LD was measured in this study since it is the most effective biomarker for estimating the plaque burden.⁷ The AtheroCloud software computes average LD based on the bi-directional polyline distance method (PDM).^{38,39} PDM is the shortest distance from one point to another and it is better than the Euclidian distance. Before performing LD computation, we first ensure that the carotid near and far wall borders computed by AtheroCloud matches with the near and far wall borders computed by the manual experts. To achieve these requirements, we perform the following series of steps: (i) B-spline smoothing; (ii) assuring that near and far wall borders of both AtheroCloud and manual reading to have a common support. Common support guarantee same start and end coordinated hence ensuring the same length; and (iii) performing interpolation which ensures 100 equidistant interpolated points in both AtheroCloud and manual near and far wall borders. Steps performed to match the borders are similar to a recent study carried out by Suri and his team.^{6,32}

3.2. Intra-and inter-operator reproducibility analysis

Three different operators, one novice (Auto 1) and two experienced (Auto 2 and Auto 3) examined the database and measured the LD by using the AtheroCloud software. Every patient was analyzed three times by each operator. Each operator revisited in a gap of one month and the same analysis was repeated. While performing the analysis, the operators were blinded to each other. Similarly, while reanalyzing (performing the second and third measurements) the operators were blinded to their previous results. Intra-operator reproducibility was analyzed by using the individual analysis by each operator and inter-operator reproducibility was analyzed by comparing the analysis of both the operators. Note that ultrasound acquisition did not undergo any kind of special variations such as gain control or dynamic range.

3.3. Intra-and inter-observer variability analysis

Two different experienced observers (Manual 1, and Manual 2) performed the freehand manual measurements for intra- and inter-observer variability. Both the two observers use a commercially available software system – ImgTracer^{6,40,41} to manually trace the lumen-intima (LI) borders for the near and far wall of carotid arteries. A sample view of ImgTracer is shown in Fig. 4. Both the observers analyzed the database two times. While performing the analysis, the observers were blinded to each other. Similarly, while reanalyzing (performing the second measurement) the observers were blinded to their previous results. Intra-observer reproducibility was analyzed by using the individual analysis by each observer and inter-observer reproducibility was analyzed by comparing the analysis of both the observer.

3.4. Statistical analysis

Three kinds of statistical tests (two-tailed z-test, Mann-Whitney test, and Kolmogorov-Smirnov (KS) test) were performed using MedCalc 16.8.4 statistical software (Osteen Belgium). Two-tailed z-test and Mann-Whitney test compute the statistical significance for both (a) reproducibility and (b) variability analysis. Similarly, KS test evaluates the normality of the distribution of automated and manual readings. All the three tests were applied between three operators (Auto 1, Auto 2 and Auto 3) and two observers (Manual 1, and Manual 2). A level of 5% statistical significance was used throughout the study.

4. Results

4.1. Intra- and inter-operator reproducibility of cloud-based automated system

Intra- and inter-operator reproducibility of the LD computed between all the three automated reading pairs by three operators is shown in Table 1. Our observation clearly shows the high mean correlation coefficient (CC) (close to 1.0) for intra- and inter-

Table 1
Correlation Coefficient (CC) for reproducibility and variability of the automated system.

| Combinations | Correlation Coefficient |
|----------------------------------|--------------------------|
| Reproducibility (Intra-Operator) | |
| Auto 1a versus Auto 1b | CC = 0.9973 (P < 0.0001) |
| Auto 1b versus Auto 1c | CC = 0.9973 (P < 0.0001) |
| Auto 1a versus Auto 1c | CC = 0.9973 (P < 0.0001) |
| Auto 2a versus Auto 2b | CC = 0.9929 (P < 0.0001) |
| Auto 2b versus Auto 2c | CC = 0.9923 (P < 0.0001) |
| Auto 2a versus Auto 2c | CC = 0.9940 (P < 0.0001) |
| Auto 3a versus Auto 3b | CC = 0.9927 (P < 0.0001) |
| Auto 3b versus Auto 3c | CC = 0.9941 (P < 0.0001) |
| Auto 3a versus Auto 3c | CC = 0.9900 (P < 0.0001) |
| Reproducibility (Inter-Operator) | |
| Auto 1a versus Auto 2a | CC = 0.9522 (P < 0.0001) |
| Auto 1a versus Auto 2b | CC = 0.9572 (P < 0.0001) |
| Auto 1a versus Auto 2c | CC = 0.9512 (P < 0.0001) |
| Auto 1b versus Auto 2a | CC = 0.9525 (P < 0.0001) |
| Auto 1b versus Auto 2b | CC = 0.9580 (P < 0.0001) |
| Auto 1b versus Auto 2c | CC = 0.9522 (P < 0.0001) |
| Auto 1c versus Auto 2a | CC = 0.9525 (P < 0.0001) |
| Auto 1c versus Auto 2b | CC = 0.9580 (P < 0.0001) |
| Auto 1c versus Auto 2c | CC = 0.9522 (P < 0.0001) |
| Auto 1a versus Auto 3a | CC = 0.9640 (P < 0.0001) |
| Auto 1a versus Auto 3b | CC = 0.9687 (P < 0.0001) |
| Auto 1a versus Auto 3c | CC = 0.9684 (P < 0.0001) |
| Auto 1b versus Auto 3a | CC = 0.9647 (P < 0.0001) |
| Auto 1b versus Auto 3b | CC = 0.9694 (P < 0.0001) |
| Auto 1b versus Auto 3c | CC = 0.9691 (P < 0.0001) |
| Auto 1c versus Auto 3a | CC = 0.9647 (P < 0.0001) |
| Auto 1c versus Auto 3b | CC = 0.9694 (P < 0.0001) |
| Auto 1c versus Auto 3c | CC = 0.9691 (P < 0.0001) |
| Auto 2a versus Auto 3a | CC = 0.9834 (P < 0.0001) |
| Auto 2a versus Auto 3b | CC = 0.9844 (P < 0.0001) |
| Auto 2a versus Auto 3c | CC = 0.9845 (P < 0.0001) |
| Auto 2b versus Auto 3a | CC = 0.9830 (P < 0.0001) |
| Auto 2b versus Auto 3b | CC = 0.9873 (P < 0.0001) |
| Auto 2b versus Auto 3c | CC = 0.9835 (P < 0.0001) |
| Auto 2c versus Auto 3a | CC = 0.9797 (P < 0.0001) |
| Auto 2c versus Auto 3b | CC = 0.9835 (P < 0.0001) |
| Auto 2c versus Auto 3c | CC = 0.9847 (P < 0.0001) |
| Variability (Intra-Operator) | |
| Manual 1a versus Manual 1b | CC = 0.9688 (P < 0.0001) |
| Manual 2a versus Manual 2b | CC = 0.9927 (P < 0.0001) |
| Variability (Inter-Operator) | |
| Manual 1a versus Manual 2a | CC = 0.9888 (P < 0.0001) |
| Manual 1a versus Manual 2b | CC = 0.9911 (P < 0.0001) |
| Manual 1b versus Manual 2a | CC = 0.9693 (P < 0.0001) |
| Manual 1b versus Manual 2b | CC = 0.9732 (P < 0.0001) |

Table 2
Results of MAD, MAE, PoM, FoM, and AUC for Auto and Manual LD readings.

| Mean Absolute Difference (MAD) | | | | | |
|--|--------------|--------------|--------------|--------------|--------------------|
| | Manual 1a | Manual 1b | Manual 2a | Manual 2b | Average of Manuals |
| Auto 1a | 0.08 ± 0.06 | 0.07 ± 0.06 | 0.06 ± 0.05 | 0.06 ± 0.05 | 0.06 ± 0.05 |
| Auto 1b | 0.09 ± 0.06 | 0.07 ± 0.06 | 0.06 ± 0.05 | 0.06 ± 0.05 | 0.06 ± 0.05 |
| Auto 1c | 0.09 ± 0.06 | 0.07 ± 0.06 | 0.06 ± 0.05 | 0.06 ± 0.05 | 0.06 ± 0.05 |
| Auto 2a | 0.07 ± 0.06 | 0.05 ± 0.04 | 0.05 ± 0.04 | 0.05 ± 0.04 | 0.04 ± 0.04 |
| Auto 2b | 0.07 ± 0.05 | 0.05 ± 0.04 | 0.04 ± 0.04 | 0.05 ± 0.04 | 0.04 ± 0.03 |
| Auto 2c | 0.07 ± 0.05 | 0.05 ± 0.04 | 0.04 ± 0.04 | 0.05 ± 0.04 | 0.04 ± 0.03 |
| Average of Autos | 0.07 ± 0.05 | 0.05 ± 0.04 | 0.04 ± 0.04 | 0.05 ± 0.04 | 0.04 ± 0.03 |
| Mean Absolute Error (MAE) | | | | | |
| Auto 1a | 9.59 ± 6.86 | 8.01 ± 6.71 | 6.72 ± 6.26 | 6.81 ± 6.30 | 6.56 ± 5.82 |
| Auto 1b | 9.65 ± 6.91 | 8.05 ± 6.80 | 6.74 ± 6.28 | 6.89 ± 6.32 | 6.63 ± 5.88 |
| Auto 1c | 9.65 ± 6.92 | 8.04 ± 6.88 | 6.77 ± 6.35 | 6.96 ± 6.35 | 6.65 ± 5.95 |
| Auto 2a | 8.08 ± 6.68 | 5.79 ± 4.80 | 5.62 ± 5.06 | 5.81 ± 4.75 | 4.71 ± 4.25 |
| Auto 2b | 7.79 ± 6.39 | 5.43 ± 4.05 | 5.44 ± 4.60 | 5.56 ± 4.38 | 4.29 ± 3.71 |
| Auto 2c | 7.79 ± 6.38 | 5.43 ± 4.04 | 5.44 ± 4.59 | 5.56 ± 4.37 | 4.29 ± 3.70 |
| Average of Autos | 8.24 ± 6.16 | 6.22 ± 4.93 | 5.42 ± 4.67 | 5.60 ± 4.56 | 4.96 ± 3.86 |
| Mean Precision-of-Merit (PoM) | | | | | |
| Auto 1a | 90.44 ± 6.86 | 91.99 ± 6.71 | 93.28 ± 6.26 | 93.19 ± 6.30 | 93.44 ± 5.82 |
| Auto 1b | 90.35 ± 6.91 | 91.95 ± 6.80 | 93.26 ± 6.28 | 93.11 ± 6.32 | 93.37 ± 5.88 |
| Auto 1c | 90.35 ± 6.92 | 91.96 ± 6.88 | 93.23 ± 6.35 | 93.04 ± 6.35 | 93.35 ± 5.95 |
| Auto 2a | 91.92 ± 6.68 | 94.21 ± 4.80 | 94.38 ± 5.06 | 94.19 ± 4.75 | 95.29 ± 4.25 |
| Auto 2b | 92.21 ± 6.39 | 94.57 ± 4.05 | 94.56 ± 4.60 | 94.44 ± 4.38 | 95.71 ± 3.71 |
| Auto 2c | 92.20 ± 6.38 | 94.56 ± 4.04 | 94.55 ± 4.59 | 94.43 ± 4.37 | 95.70 ± 3.70 |
| Average of Autos | 91.76 ± 6.16 | 93.78 ± 4.93 | 94.58 ± 4.67 | 94.40 ± 4.56 | 95.04 ± 3.86 |
| Mean Central Difference/Mean Figure-of-Merit (FoM) | | | | | |
| Auto 1a | 95.43 | 95.80 | 99.45 | 99.54 | 97.52 |
| Auto 1b | 95.29 | 95.66 | 99.30 | 99.39 | 97.37 |
| Auto 1c | 95.29 | 95.66 | 99.31 | 99.39 | 97.38 |
| Auto 2a | 98.18 | 98.56 | 97.68 | 97.60 | 99.67 |
| Auto 2b | 98.54 | 98.93 | 97.30 | 97.22 | 99.30 |
| Auto 2c | 98.53 | 98.92 | 97.29 | 97.21 | 99.29 |
| Average of Autos | 96.88 | 97.26 | 99.04 | 98.95 | 99.00 |
| Area Under the Curve (AUC) | | | | | |
| Auto 1a | 0.956 | 0.968 | 0.969 | 0.964 | NA |
| Auto 1b | 0.956 | 0.968 | 0.972 | 0.965 | NA |
| Auto 1c | 0.955 | 0.967 | 0.971 | 0.964 | NA |
| Auto 2a | 0.967 | 0.984 | 0.981 | 0.979 | NA |
| Auto 2b | 0.970 | 0.988 | 0.980 | 0.987 | NA |
| Auto 2c | 0.970 | 0.988 | 0.980 | 0.987 | NA |
| Average | 0.962 | 0.977 | 0.976 | 0.974 | NA |
| SD | 0.01 | 0.01 | 0.01 | 0.01 | NA |

operator reproducibility between all the three automated reading pairs, demonstrating high statistical significance. Fig. A1 shows the regression plots for (a) intra-operator reproducibility and Figs. A2–A4 shows the regression plots for (b) inter-operator reproducibility of LD measurements using AtheroCloud software (see Appendix A1). This proves our hypothesis that even a novice operator can yield high reproducibility when computing automated LD readings.

4.2. Intra- and inter-observer variability for validation of cloud-based automated system

Intra- and inter-observer variability of the LD computed between both the manual reading pairs by the two observers is shown in Table 1. Our observation clearly shows the high mean correlation coefficient (CC) (close to 1.0) for intra- and inter-observer variability between both the two manual reading pairs. The corresponding regression plots for intra- and inter-observer variability are shown in Fig. A5 (Appendix A1), respectively. Strong correlations for intra- and inter-observer variability guarantee the reliability of the proposed automated system.

5. Performance evaluation

In the current study, performance evaluation of the AtheroCloud system was evaluated using four different strategies: (1) computing deviation and error between the automated and

manual readings using mean absolute difference (MAD) and mean absolute error (MAE), respectively; (2) computing closeness between the automated and manual readings using: (a) Precision-of-Merit (PoM) and (b) Figure-of-Merit (FoM). PoM demonstrates the closeness between each sample of manual and auto readings, whereas, FoM (also known as Mean central difference) demonstrate the closeness between the mean of manual readings against the mean of auto readings; (3) comparing the performance based on Bland-Altman plots; (4) comparing the diagnostic performance using receiver operating characteristics (ROC) curves.

5.1. Performance evaluation based on computing deviation and error: MAD and MAE analysis

MAD and MAE are mathematically expressed as shown in Eqs. (B3) and (B4) in Appendix B1. Results of MAD and MAE are shown in Table 2. It was observed that average of all the auto combinations for MAD (0.34 ± 0.40 mm) and MAE ($5.73 \pm 7.14\%$) are lowest for Manual 2a.

5.2. Performance evaluation based on computing closeness: PoM and FoM analysis

PoM and FoM are mathematically expressed as shown in Eqs. (B5) and (B8) in Appendix B1. Results of PoM and FoM are shown in

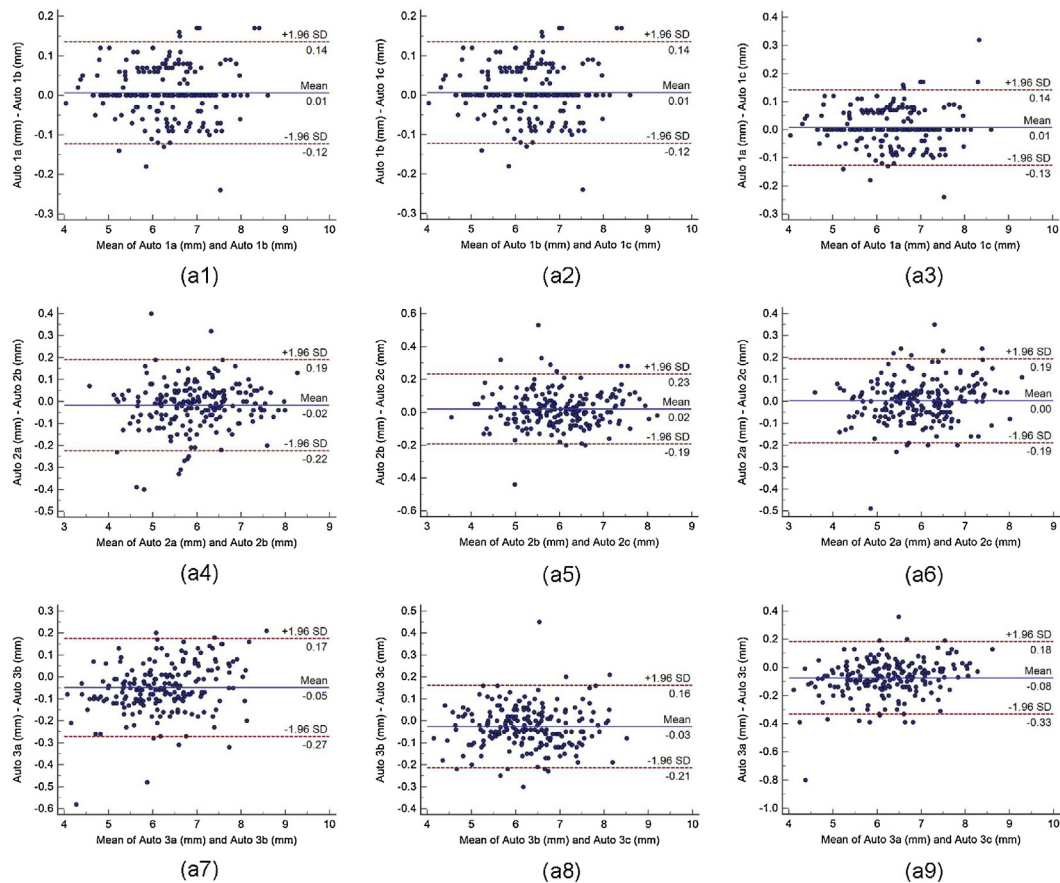


Fig. 5. The Bland-Altman plot for intra-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows Bland-Altman plot between (Auto 1a-Auto 1b), (Auto 1b-Auto 1c), (Auto 1a-Auto 1c), (Auto 2a-Auto 2b), (Auto 2b-Auto 2c), (Auto 2a-Auto 2c), (Auto 3a-Auto 3b), (Auto 3b-Auto 3c), and (Auto 3a-Auto 3c) of LD measurements using AtheroCloud software.

Table 2. It was observed that average of all the auto combinations for PoM ($94.27 \pm 7.14\%$) and FoM (98.94%) are highest for Manual 2a.

5.3. Performance evaluation based on bland-Altmanplots

Bland-Altman plot indicates the bias or average difference between the two readings. Fig. 5 shows the Bland-Altman plots for (a) intra-operator reproducibility and Figs. 6–8 shows the Bland-Altman plots for (b) inter-operator reproducibility of LD measurements using the AtheroCloud software. Similarly, Bland-Altman plots for (c) intra-observer variability and (d) inter-observer variability of LD measurements using ImgTracer software are shown in Fig. 9. Results show a high degree of agreement between the computed LD readings.

5.4. Performance evaluation based on receiver operating characteristics (ROC) curves

The (ROC) analysis was performed on auto against the manual measurements. Since three operators underwent the LD readings three times, there are three sets of the operators' combinations: (Auto 1a, 1b, 1c), (Auto 2a, 2b, 2c), and (Auto 3a, 3b, 3c). There are four manual readings, taken from two manual observers who perform LI borders tracings for the near and far wall of carotid arteries two times (Manual 1a, 1b, 2a, 2b). Comparison of the ROC's of three auto readings: (Auto 1a, 1b, 1c), (Auto 2a, 2b, 2c), and (Auto

3a, 3b, 3c), taking four different manual readings (Manual 1a, 1b, 2a, 2b) as ground truth are shown in Figs. 10–12, respectively.

For this study, 6 mm LD risk threshold was empirically selected. All the borders whose LD is greater than 6 mm are low-risk (non-diseased), while borders whose LD is less than 6 mm are high-risk (diseased). Table 2 shows the area under the curve (AUC) for all the auto combinations. It was observed that the highest mean AUC ($=0.937$) for all the possible combination was obtained while taking Manual 2a as the ground truth. We detected a high AUC for all the possible combinations.

5.5. Statistical tests

For measuring reliability and stability of the system, we have performed three statistical tests (z-test, Mann-Whitney, and Kolmogorov-Smirnov (KS) test). Table 3 shows the result of the three statistical tests. The negative z-value indicates that raw result is below the mean. For all the paired samples, P-values were greater than 5% hence they passed the Mann-Whitney test. The normal distribution for the entire auto and manual readings were greater than 5% thus they passed the KS test.

6. Discussion

6.1. Our system

The objective of this study was to demonstrate the intra- and inter-operator reproducibility for cloud-based automated LD

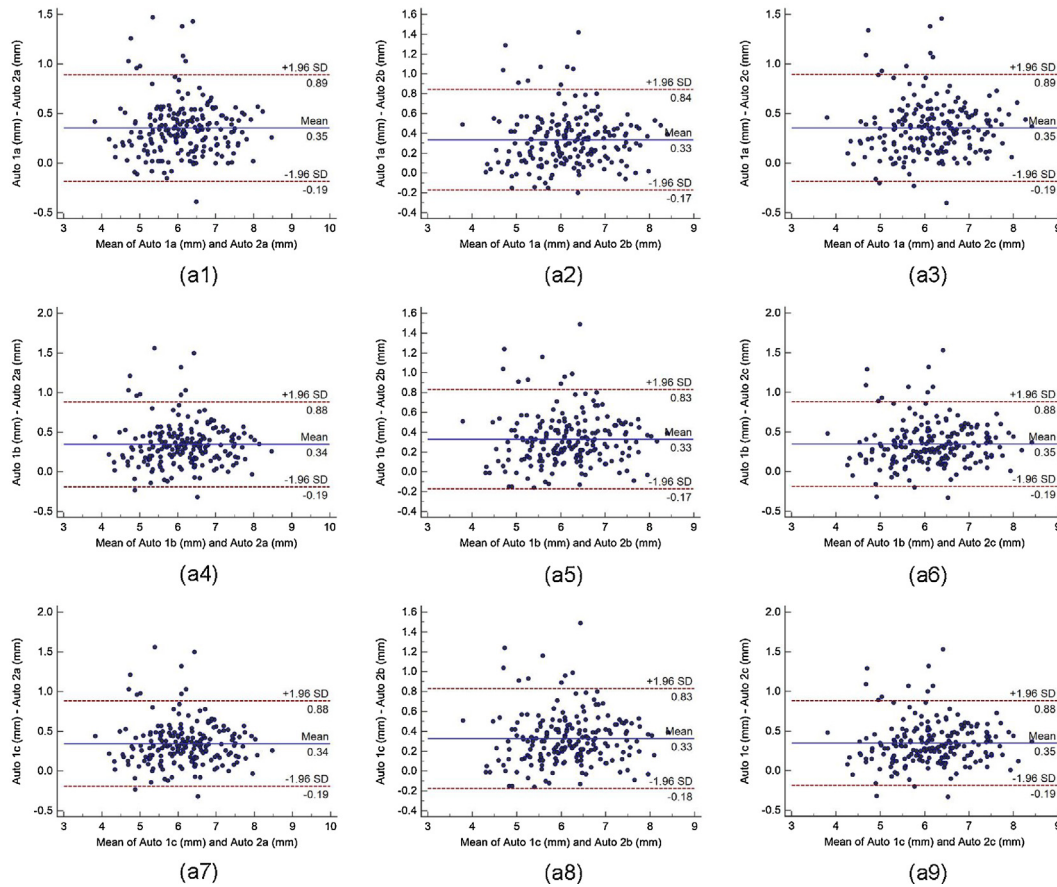


Fig. 6. The Bland-Altman plot for inter-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows Bland-Altman plot between (Auto 1a-Auto 2a), (Auto 1a-Auto 2b), (Auto 1a-Auto 2c), (Auto 1b-Auto 2a), (Auto 1b-Auto 2b), (Auto 1b-Auto 2c), (Auto 1c-Auto 2a), (Auto 1c-Auto 2b), and (Auto 1c-Auto 2c) of LD measurements using AtheroCloud software.

Table 3
Statistical tests for reproducibility of the automated system.

| Combinations | Two-tailed z-Test | | Mann-Whitney Test |
|---|-------------------|---------|-------------------|
| | z | P-value | P-value |
| Reproducibility (Intra-Operator) | | | |
| Auto 1a versus Auto 1b | 0.0604 | <0.9518 | 0.9393 |
| Auto 1b versus Auto 1c | 0.0046 | <0.9963 | 0.9931 |
| Auto 1a versus Auto 1c | 0.0558 | <0.9555 | 0.9465 |
| Auto 2a versus Auto 2b | 0.1493 | <0.8813 | 0.8920 |
| Auto 2b versus Auto 2c | 0.0422 | <0.9664 | 0.9921 |
| Auto 2a versus Auto 2c | 0.1076 | <0.9144 | 0.9105 |
| Reproducibility (Inter-Operator) | | | |
| Auto 1a versus Auto 2a | 1.1220 | <0.2619 | 0.2387 |
| Auto 1a versus Auto 2b | 1.2739 | <0.2027 | 0.1878 |
| Auto 1a versus Auto 2c | 1.2345 | <0.2170 | 0.1801 |
| Auto 1b versus Auto 2a | 1.1808 | <0.2377 | 0.2146 |
| Auto 1b versus Auto 2b | 1.3327 | <0.1826 | 0.1697 |
| Auto 1b versus Auto 2c | 1.2935 | <0.1959 | 0.1659 |
| Auto 1c versus Auto 2a | 1.1765 | <0.2394 | 0.2201 |
| Auto 1c versus Auto 2b | 1.3284 | <0.1841 | 0.1709 |
| Auto 1c versus Auto 2c | 1.2891 | <0.1974 | 0.1691 |
| Variability (Intra-Observer) | | | |
| Manual 1a versus Manual 1b | 0.2600 | <0.7948 | 0.8106 |
| Manual 2a versus Manual 2b | 0.1220 | <0.9731 | 0.9790 |
| Variability (Inter-Observer) | | | |
| Manual 1a versus Manual 2a | 1.6667 | <0.0956 | 0.0871 |
| Manual 1a versus Manual 2b | 1.6995 | <0.0892 | 0.0781 |
| Manual 1b versus Manual 2a | 1.5462 | <0.0488 | 0.1081 |
| Manual 1b versus Manual 2b | 1.5800 | <0.1141 | 0.1006 |

measurement system using novice operators. For validation of reproducibility analysis, we compare the automated readings against the expert reading via manual delineation of LI interfaces tracings for the near and far wall of carotid arteries. Further, we performed intra- and inter-observer variability analysis of the expert readings.

6.2. A special note on reproducibility

There are many important factors which can affect the performance of the reproducibility study. These factors include operator’s background and experience, image quality and format, the number of systems used, mode of performance, time of the day, lighting conditions, operator fatigue, internet speed, far and near wall tweaked along with region and extent of tweaking. All the above factors are summarized in Table 4. For reproducibility tracings, all the operators used standardized software ImgTracer, which had been successfully used in many earlier studies on medical imaging.^{40,41,42–44}

6.3. Review of methods reported in literature

Recent innovations have made it possible to automate the LD computation methods by using better image reconstruction tools. Stadler²⁴ et al in 1996 introduced a skew plane by locating the imaging plane rotated slightly off-axis of the brachial artery. The LD was then estimated from the leading edges of the echo zones. The technique required manual selection of the arterial wall edge on B-

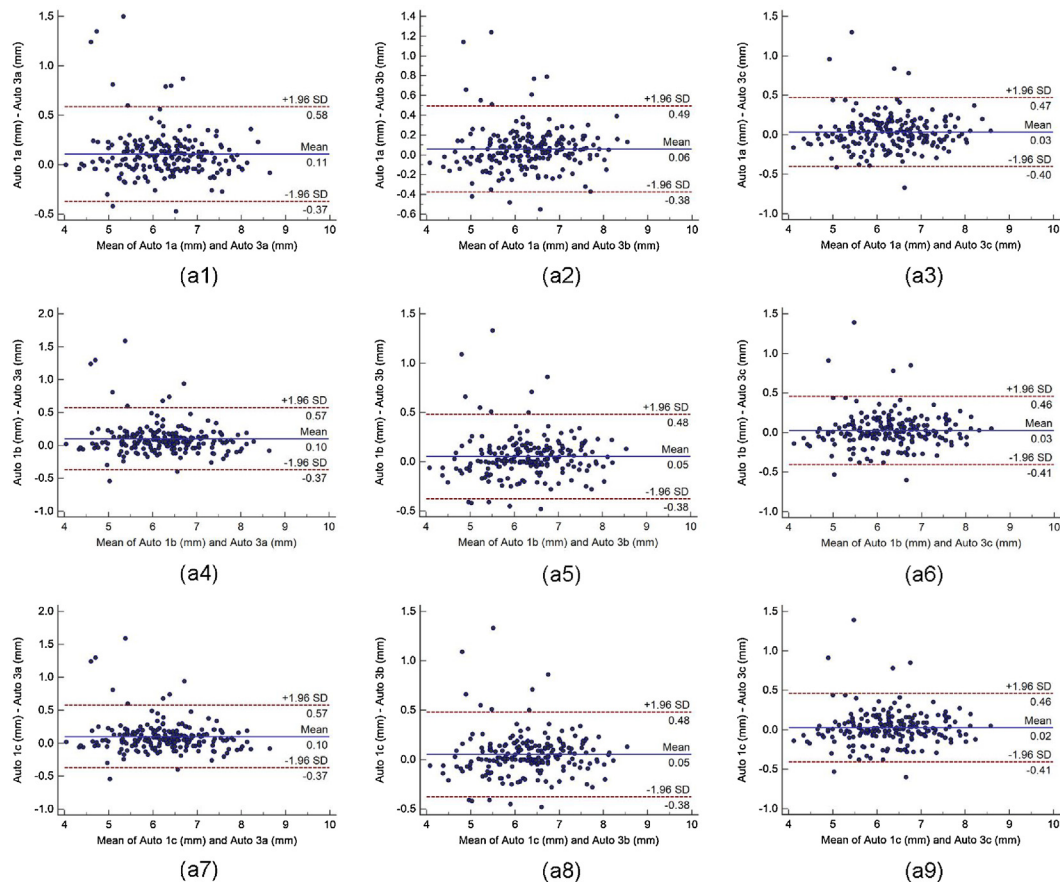


Fig. 7. The Bland-Altman plot for inter-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows Bland-Altman plot between (Auto 1a–Auto 3a), (Auto 1a–Auto 3b), (Auto 1a–Auto 3c), (Auto 1b–Auto 3a), (Auto 1b–Auto 3b), (Auto 1b–Auto 3c), (Auto 1c–Auto 3a), (Auto 1c–Auto 3b), and (Auto 1c–Auto 3c) of LD measurements using AtheroCloud software.

mode images by the operator, as a preliminary step and therefore is time-consuming for analysis of large numbers of studies in a longitudinal trial. Expert readers were also needed to perform the analysis. Beux²⁵ et al. in 2001, removed the limitation of manual selection of the arterial wall by proposing an automated pixel classification based tracking algorithm where the brachial artery diameter was evaluated by parabolic least-square approximation.

Table 4

List of factors affecting the performance of the reproducibility.

| Factors | Auto 1 | Auto 2 | Auto 3 |
|---|---------------------------|---------------------------|---------------------------|
| Operators Background | Novice | Neuroradiologic | Radiologist |
| Experience in Ultrasound | NA | 5–7 years | 2 years |
| Image Quality | Good | Good | Good |
| Image Format | JPEG | JPEG | JPEG |
| Internet Speed | 65.42/37.60 | 26.53/16.68 | 70.53/40.68 |
| (Uploading/ Downloading) | Mbps | Mbps | Mbps |
| Number of Systems Used | One | One | One |
| Lightning Condition | Bright | Not bright | Bright |
| Mode of Performance | In one stroke | In parts | In two parts |
| Tracing on Weekend | No | Yes | Sometimes |
| Time of the Day | All day | Evening | Morning |
| Operator Fatigue | Yes | Yes | Yes |
| Far/Near Wall | Far wall | Far wall | Near wall |
| Region Tweaked | More at peripheral end | More at peripheral end | More at peripheral end |
| Extent of Tweaking | 50% | 50% | 20–30% |
| Number of Points on the Carotid Artery | 25 points are enough | 25 points are enough | 25 points are enough |

The study was performed on 38 healthy patients and two different measurements were obtained on the same day to assess reproducibility of the procedure. Endothelium-dependent vasodilation using flow mediated dilation (FMD) method was evaluated on three parameters: (a) as percent increase in brachial artery diameter at 60 s after cuff release (FMD_{60}), (b) as maximal increase (FMD_{MAX}) and (c) as the integral of percent increase in brachial artery diameter during the 3 min after cuff release (FMD_{AUC}). FMD_{MAX} was observed to be more reproducible (coefficient of variation 14%) than FMD_{60} and FMD_{AUC} . In the same year, Newey and Nassiri²⁶ used Artificial Neural Network to identify and track the brachial artery walls. Back-propagation networks were trained to identify the artery anterior and posterior walls, respectively, by recognizing the edge of the brightest echo of the vessel wall. Two networks were trained to identify the artery walls using over 3200 examples from carotid artery images. The network correctly classified approximately 97% of the randomly selected samples. Above studies uses the data set from brachial arms.

A year later, Gutierrez²⁷ et al proposed an automatic approach to measure cIMT and LD using an active contour technique. A total of 180 images from 30 patients (3 images in diastole and 3 images in systole for each patient) were analyzed. Using the active contour technique, the vessel's border was determined automatically. Potter²⁸ et al in 2007 proposed a new DICOM-based software which uses an automated edge-detection algorithm to locate the carotid arterial wall interfaces within the user selected ROI. From the pool of 30 volunteers, eight subjects were normal healthy volunteers, seven had disordered lipid metabolism and fifteen

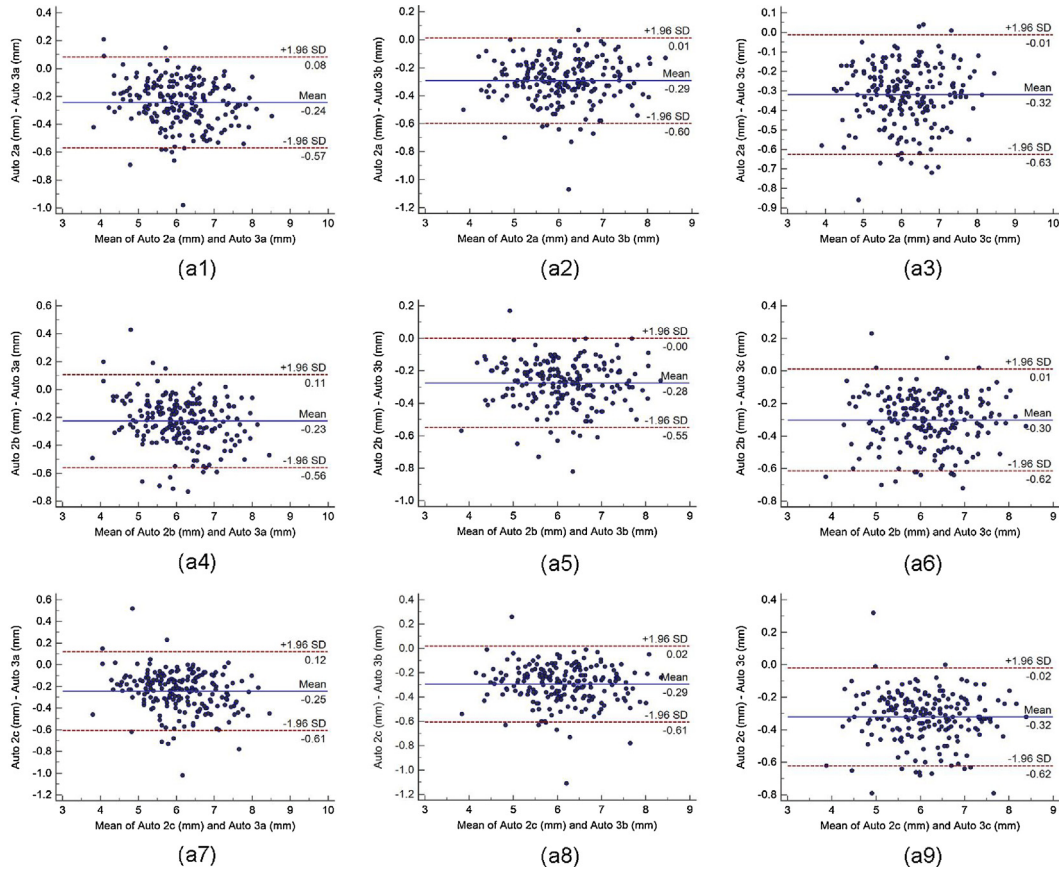


Fig. 8. The Bland-Altman plot for inter-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows Bland-Altman plot between (Auto 2a-Auto 3a), (Auto 2a-Auto 3b), (Auto 2a-Auto 3c), (Auto 2b-Auto 3a), (Auto 2b-Auto 3b), (Auto 2b-Auto 3c), (Auto 2c-Auto 3a), (Auto 2c-Auto 3b), and (Auto 2c-Auto 3c) of LD measurements using AtheroCloud software.

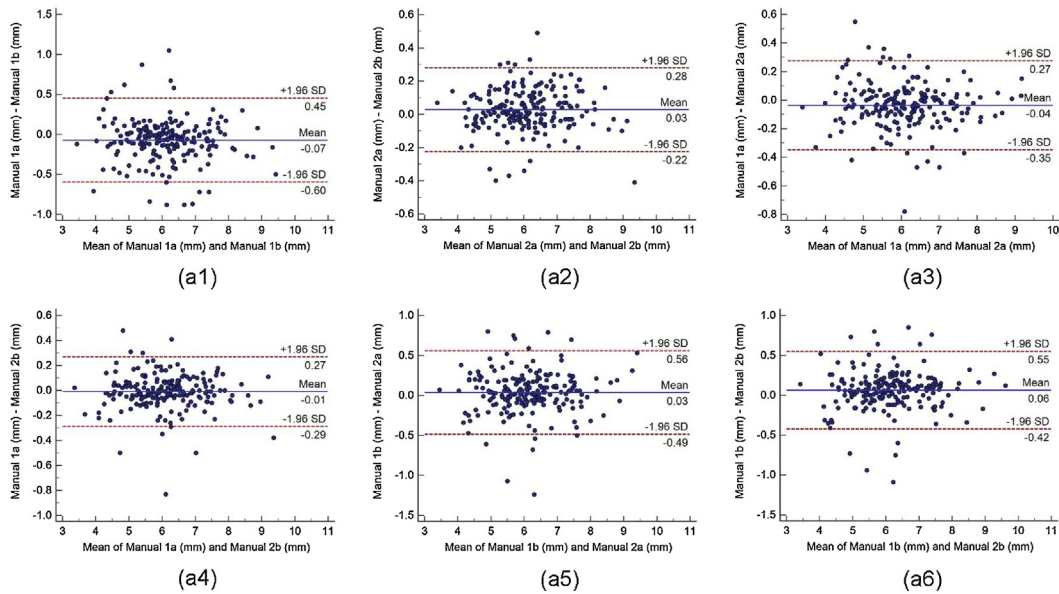


Fig. 9. The Bland-Altman plot for intra- and inter-observer variability. Figure a1, a2 shows Bland-Altman plot for intra-observer variability between (Manual 1a-Manual 1b), (Manual 2a- Manual 2b), and figure a3, a4, a5, and a6 for inter-observer variability between (Manual 1a- Manual 2a), (Manual 1a- Manual 2b), (Manual 1b- Manual 2a), and (Manual 1b- Manual 2b) of LD measurements using AtheroCloud software.

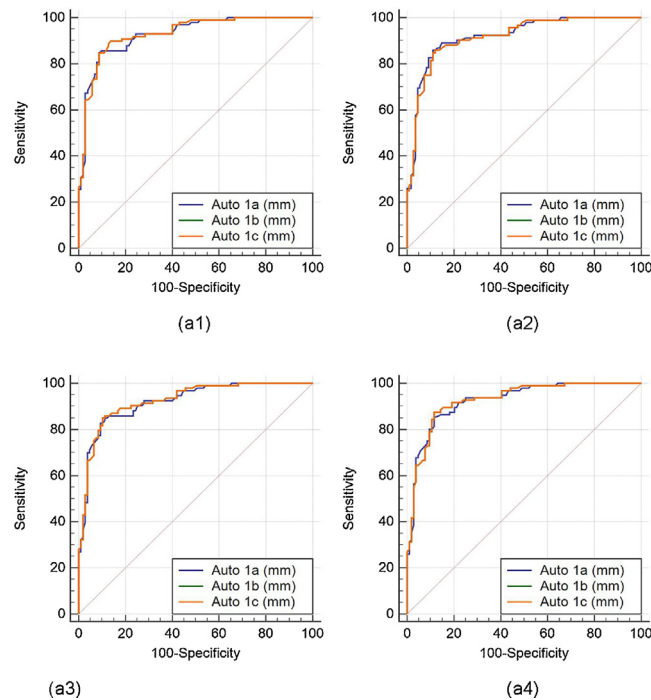


Fig. 10. Comparative receiver operating characteristic curves for three auto readings (Auto 1a, Auto 1b, and Auto 1c) by operator 1, taking Manual 1a, Manual 1b, Manual 2a, and Manual 2b as ground truth are shown in figures a1, a2, a3, and a4, respectively.

subjects had a history of stroke or transient ischemic attack. The software counts the number of pixels in each vertical pixel column between the marked lines. The threshold for detecting the edges

was calculated using a statistical clustering algorithm and LD and cIMT was computed using the ratio set during calibration. Cinthio²⁹ et al in 2010 proposed a method to measure the relative

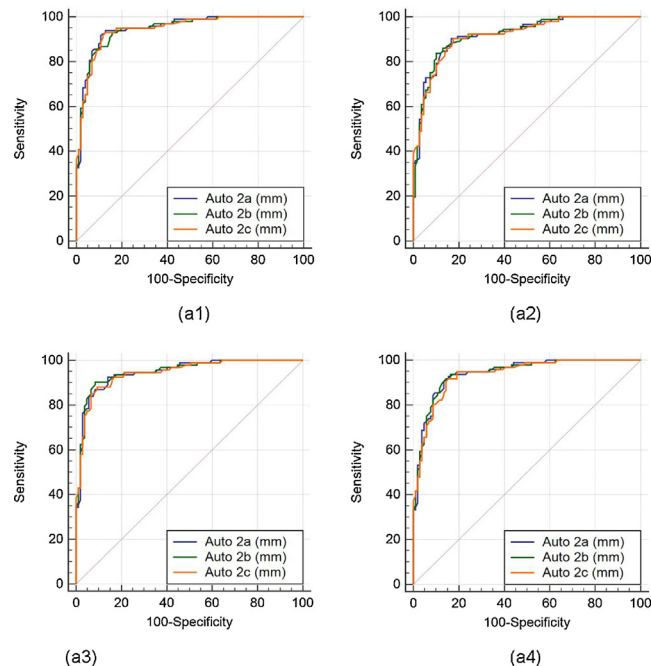


Fig. 11. Comparative receiver operating characteristic curves for three auto readings (Auto 2a, Auto 2b, and Auto 2c) by operator 1, taking Manual 1a, Manual 1b, Manual 2a, and Manual 2b as ground truth are shown in figures a1, a2, a3, and a4, respectively.

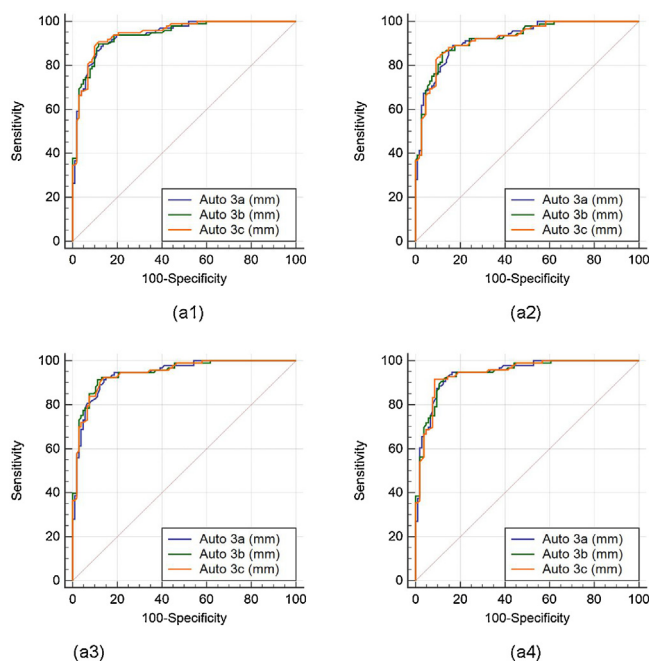


Fig. 12. Comparative receiver operating characteristic curves for three auto readings (Auto 3a, Auto 3b, and Auto 3c) by operator 1, taking Manual 1a, Manual 1b, Manual 2a, and Manual 2b as ground truth are shown in figures a1, a2, a3, and a4, respectively.

diameter changes of the common carotid artery by combining threshold detection and the equation of a straight line. The technique uses the maximum peak of the luminal-intimal echoes during measurement of the diameter change. The measurements were performed on 20 healthy normotensive volunteers and none of the subjects reported previous cardiopulmonary disease, diabetes or smoking. The study indicated that the reproducibility was sufficient for in vivo studies if the width of the ROI is 1.0 mm or wider. Sahani³⁰ et al in 2015 propose a novel method to fuse the information from (a) curve fitting error and (b) distension curve to arrive at an accurate measure of LD computation from a low axial-resolution ultrasound system such as ARTSENS in 85 human volunteers. The study was based on the fusion of information from more accurate distension waveforms with less reliable waveforms to obtain higher accuracy for LD estimation. To bring down the curve fitting time, the method performs autocorrelation of the echoes from opposite walls of the carotid artery

Recently, Suri and his team proposed a completely automated, novel, smart, cloud-based, point-of-care system (AtheroCloud) for cIMT⁶ and LD³² computation in the cloud-based setting. Both the studies^{6,32} uses the same database acquired from 100 diabetic patients (200 images) with mild stenosis overall. The system uses a combination of pixel classification and signal processing approach to determine the LD borders. Finally, LD borders are fed into the PDM system for LD computation. This was the first system of its kind that computes LD/SSI/TLA using carotid B-mode ultrasound in a web-based framework but lacked intra- and inter-operator reproducibility and variability testing. This study presents an automated and cloud-based LD measurements software system (AtheroCloud) and evaluates its: (i) intra/inter-operator reproducibility and (ii) intra/inter-observer variability.

6.4. Strength, weakness, and extension

The proposed system is automated, accurate, stable, reliable and fast (≈ 4 s per frame) with excellent intra- and inter-operator

reproducibility. Comprehensive statistical analyses demonstrate consistency and reliability of the system. Despite above features, there are few challenges and limitations. The current study focuses on carotid US images with mild stenosis consisting predominantly of male subjects. The processing speed of the proposed software completely relies on the internet speed. In future, we anticipate adding more patients and validating the software on both controlled/uncontrolled datasets. Even though we compute the CC of the entire population (100P), a reverse strategy can be adapted to estimate the population size based on the intra-class correlation. The comparison of evaluation can also be performed by varying the gap between the manual tracings. Further, the results can be validated against the histological results.

7. Conclusions

The LD computed by the AtheroCloud software demonstrates an outstanding intra- and inter-operator reproducibility. Statistical analysis exhibits accuracy, reliability, and stability of the system. The results of the pilot study clearly demonstrate that the system can be adapted to the clinical setting for clinical routine or multicenter pharmaceutical trial modes.

Conflict of interest

The authors have none to declare.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgement

We acknowledge Toho Hospital, Japan for the B-mode carotid US scans.

Appendix A1. Regression Plots for intra-operator and inter-observer analysis

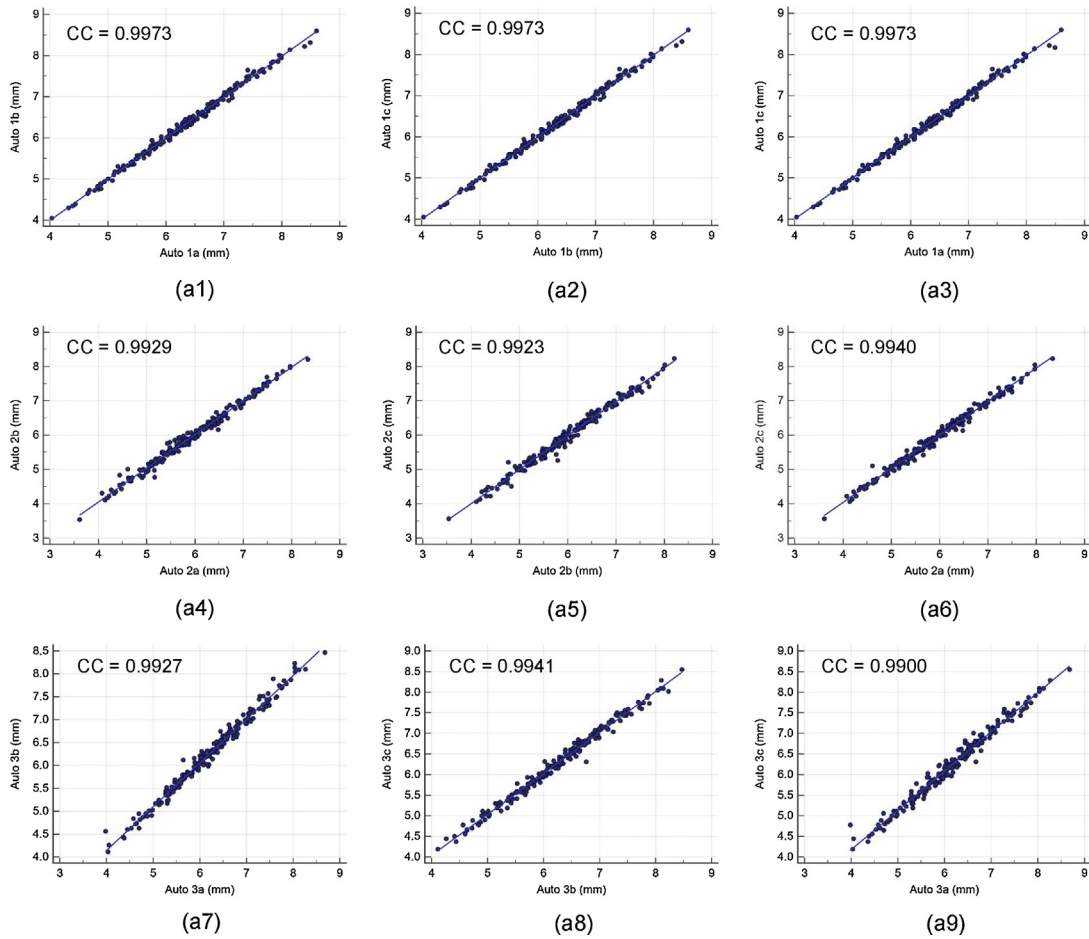


Fig. A1. The regression plot for intra-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows regression plot between (Auto 1a-Auto 1b), (Auto 1b-Auto 1c), (Auto 1a-Auto 1c), (Auto 2a-Auto 2b), (Auto 2b-Auto 2c), (Auto 2a-Auto 2c), (Auto 3a-Auto 3b), (Auto 3b-Auto 3c), and (Auto 3a-Auto 3c) of LD measurements using AtheroCloud software.

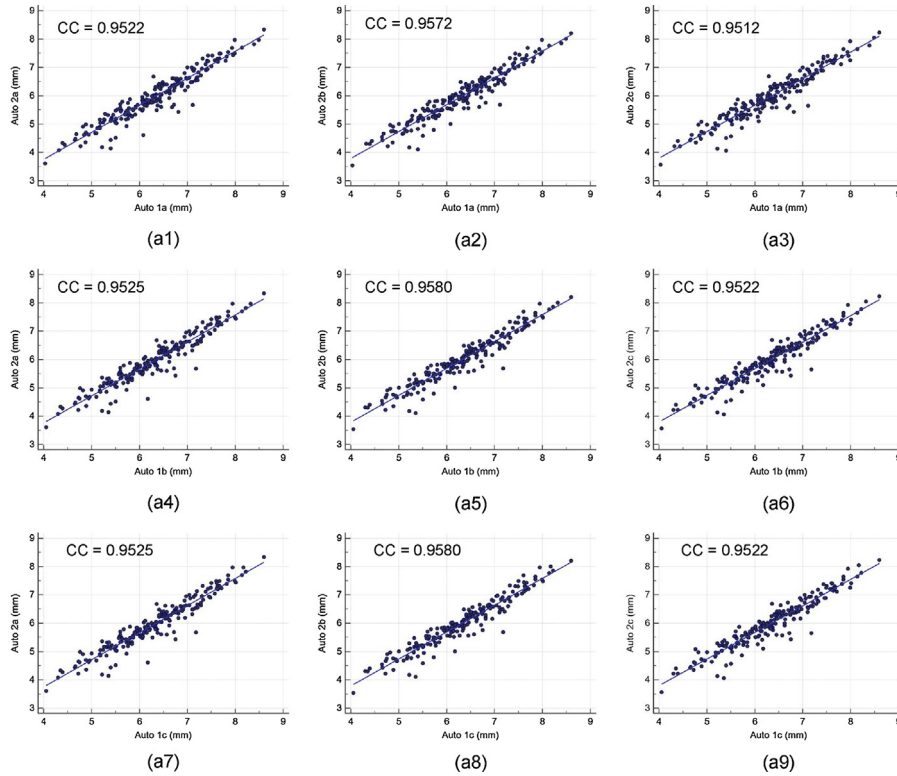


Fig. A2. The regression plot for inter-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows regression plot between (Auto 1a-Auto 2a), (Auto 1a-Auto 2b), (Auto 1a-Auto 2c), (Auto 1b-Auto 2a), (Auto 1b-Auto 2b), (Auto 1b-Auto 2c), (Auto 1c-Auto 2a), (Auto 1c-Auto 2b), and (Auto 1c-Auto 2c) of LD measurements using AtheroCloud software.

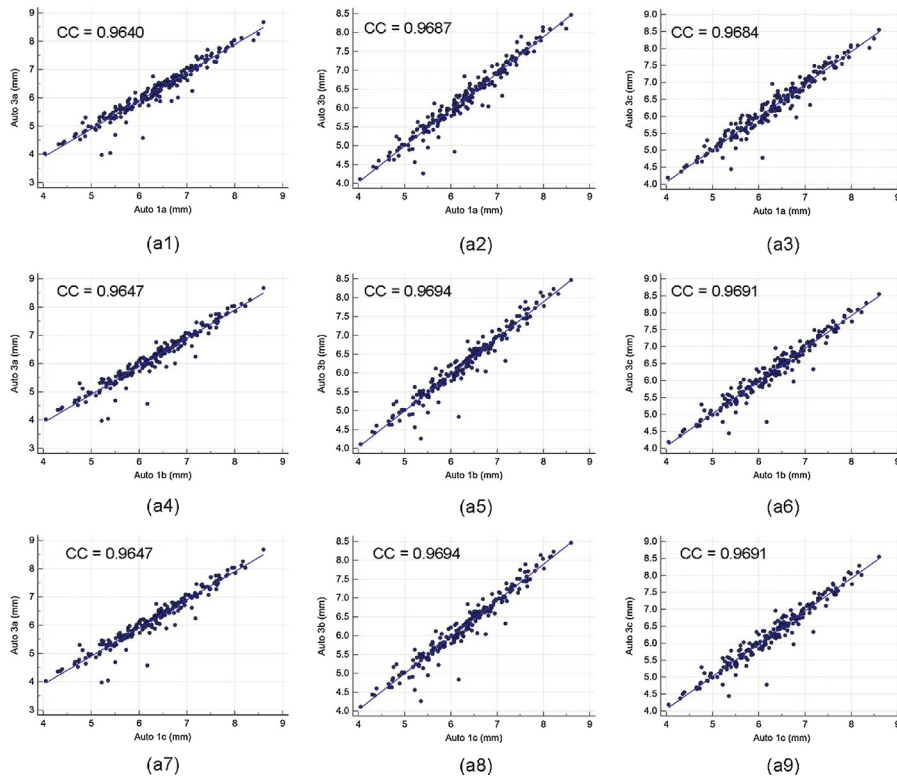


Fig. A3. The regression plot for inter-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows regression plot between (Auto 1a-Auto 3a), (Auto 1a-Auto 3b), (Auto 1a-Auto 3c), (Auto 1b-Auto 3a), (Auto 1b-Auto 3b), (Auto 1b-Auto 3c), (Auto 1c-Auto 3a), (Auto 1c-Auto 3b), and (Auto 1c-Auto 3c) of LD measurements using AtheroCloud software.

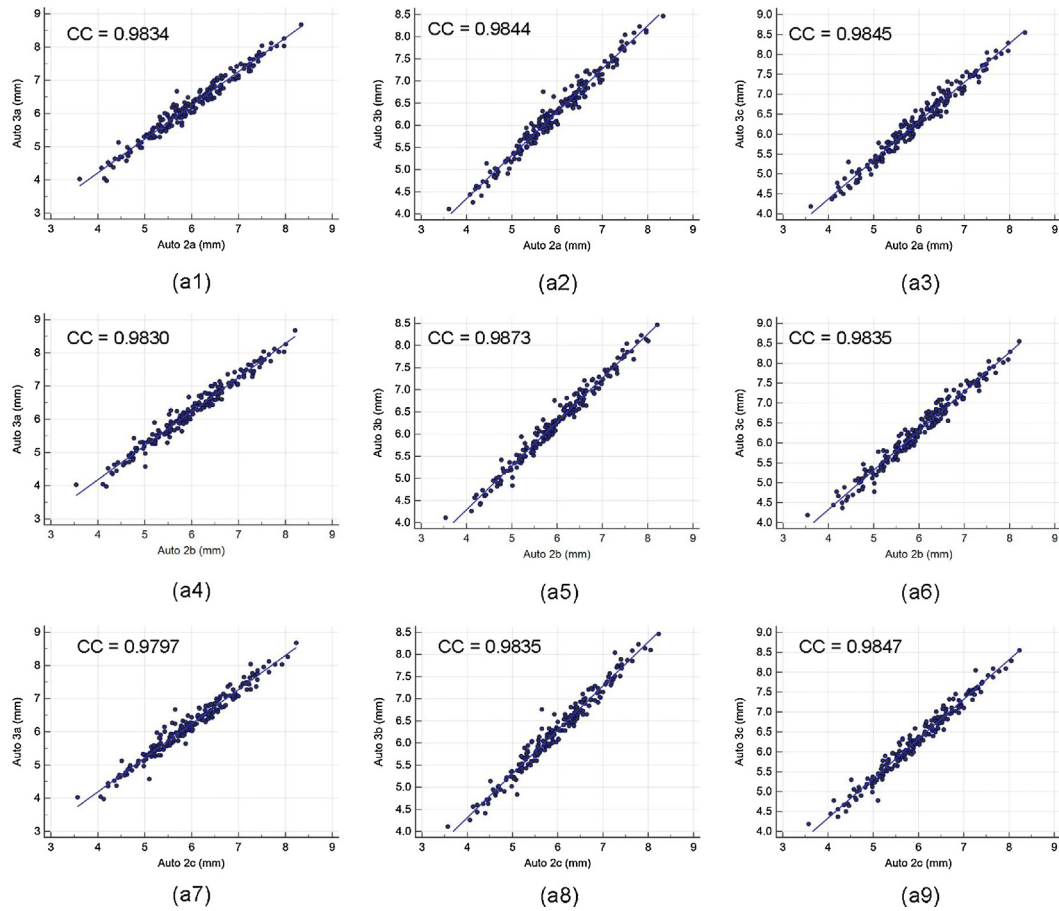


Fig. A4. The regression plot for inter-operator reproducibility. Figure a1, a2, a3, a4, a5, a6, a7, a8 and a9 shows regression plot between (Auto 2a-Auto 3a), (Auto 2a-Auto 3b), (Auto 2a-Auto 3c), (Auto 2b-Auto 3a), (Auto 2b-Auto 3b), (Auto 2b-Auto 3c), (Auto 2c-Auto 3a), (Auto 2c-Auto 3b), and (Auto 2c-Auto 3c) of LD measurements using AtheroCloud software.

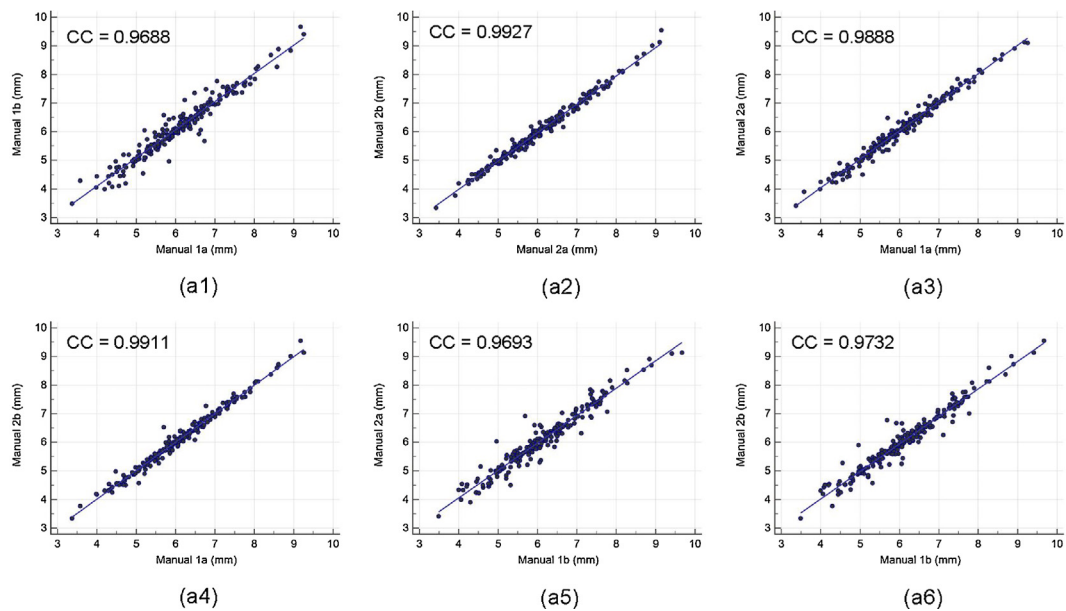


Fig. A5. The regression plot for intra- and inter-observer variability. Figure a1, a2 shows regression plot for intra-observer variability between (Manual 1a-Manual 1b), (Manual 2a- Manual 2b), and figure a3, a4, a5, and a6 for inter-observer variability between (Manual 1a- Manual 2a), (Manual 1a- Manual 2b), (Manual 1b- Manual 2a), and (Manual 1b- Manual 2b) of LD measurements using AtheroCloud software.

Appendix B1. MAD, MAE, PoM and FoM computation

This section presents a brief derivation for computation of *MAD*, *MAE*, *PoM* and *FoM* computation for AtheroCloud LD measurements.

Given the LD_{Auto}^{far} and MA_{Auto}^{near} are the interfaces computed using AtheroCloud automated method, we compute AtheroCloud LD using the polyline distance method and is given as:

$$LD_{Auto} = PDM\left(LD_{Auto}^{far}, MA_{Auto}^{near}\right) \quad (B.1)$$

Similarly, using the definition of PDM, we can compute the LD measurements using manual tracings, given as:

$$LD_{Manual} = PDM\left(LD_{Manual}^{far}, MA_{Manual}^{near}\right) \quad (B.2)$$

If a database of N images is considered, then the overall system's performance can be computed using Mean absolute difference (*MAD*), Mean absolute error (*MAE*), Precision-of-Merit (*PoM*) and Figure-of-Merit (*FoM*) in percentage as:

(i) Mean absolute difference (*MAD*)

$$MAD = \frac{1}{N} \sum_{j=1}^N |LD_{Auto} - LD_{Manual}| \quad (B.3)$$

(ii) Mean absolute error (*MAE*)

$$MAE = \frac{1}{N} \sum_{j=1}^N \left(\frac{|LD_{Auto} - LD_{Manual}|}{LD_{Manual}} \right) * 100 \quad (B.4)$$

(iii) Precision-of-Merit (*PoM*)

$$PoM = \frac{1}{N} \sum_{j=1}^N 100 - \left[\left(\frac{|LD_{Auto} - LD_{Manual}|}{LD_{Manual}} \right) * 100 \right] \quad (B.5)$$

(iv) Figure-of-Merit (*FoM*)

Let LD_{Auto_i} be the LD value automatically computed by the proposed AtheroCloud \hat{O} system on the i^{th} image of the database of N images. Now the overall mean AtheroCloud \hat{O} LD can be computed as:

$$\overline{LD}_{Auto} = \frac{1}{N} \sum_{i=1}^N LD_{Auto_i} \quad (B.6)$$

Correspondingly, if LD_{Manual_i} is the LD value computed from the traced manual measurements on the i^{th} image of the database of N images. Then, the overall mean manual LD can be computed as:

$$\overline{LD}_{Manual} = \frac{1}{N} \sum_{i=1}^N LD_{Manual_i} \quad (B.7)$$

The system's performance can finally be computed using Figure-of-Merit (*FoM*) in percentage as:

$$FoM = 100 - \left[\left(\left| \overline{LD}_{Auto} - \overline{LD}_{Manual} \right| \right) * 100 \right] \quad (B.8)$$

References

- Mozaffarian D, Benjamin EJ, Go AS, et al. Executive Summary: Heart Disease and Stroke Statistics–2016 Update: a report from the American Heart Association. *Circulation*. 2016;133:447–454.
- Ross R. Cell biology of atherosclerosis. *Annu Rev Physiol*. 1995;57:791–804.
- Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature*. 2011;473:317.
- Strandness DE, Eikelboom BC. Carotid artery stenosis – where do we go from here? *Eur J Ultrasound*. 1998;7:S17–S26.
- Silver FL, Mackey A, Clark WM, et al. Safety of stenting and endarterectomy by symptomatic status in the Carotid Revascularization Endarterectomy Versus Stenting Trial (CREST). *Stroke*. 2011;42:675–680.
- Saba L, Banchhor SK, Suri HS, et al. Accurate cloud-based smart IMT measurement, its validation and stroke risk stratification in carotid ultrasound: A web-based point-of-care tool for multicenter clinical trial. *Comput Biol Med*. 2016;75:217–234.
- Sanches JM, Laine AF, Suri JS. *Ultrasound Imaging*. New York: Springer; 2012.
- Saba L, Sanches JM, Pedro LM, Suri JS. *Multi-modality Atherosclerosis Imaging and Diagnosis*. New York: Springer; 2014.
- Polak JF. *Peripheral Vascular Sonography*. Lippincott Williams & Wilkins; 2004.
- Polak JF, Wong Q, Johnson WC, Bluemke DA, et al. Associations of cardiovascular risk factors, carotid intima-media thickness and left ventricular mass with inter-adventitial diameters of the common carotid artery: the Multi-Ethnic Study of Atherosclerosis (MESA). *Atherosclerosis*. 2011;218:344–349.
- Polak JF, Sacco RL, Post WS, Vaidya D, et al. Incident stroke is associated with common carotid artery diameter and not common carotid artery intima-media thickness. *Stroke*. 2014;45:1442–1446.
- Naim C, Douziech M, Therasse É, Robillard P, et al. Vulnerable atherosclerotic carotid plaque evaluation by ultrasound, computed tomography angiography, and magnetic resonance imaging: an overview. *Can Assoc Radiol J*. 2014;65:275–286.
- Sahani AK, Joseph J, Radhakrishnan R, Sivaprakasam M. Automatic measurement of end-diastolic arterial lumen diameter in ARTSENS. *J Med Devices*. 2015;9:041002.
- Araki T, Kumar AM, Krishna Kumar P, et al. Ultrasound-based automated carotid lumen diameter/stenosis measurement and its validation system. *J Vasc Ultrasound*. 2016;40:120–134.
- Davies KN, Humphrey PR. Complications of cerebral angiography in patients with symptomatic carotid territory ischemia screened by carotid ultrasound. *J Neurol Neurosurg Psychiatry*. 1993;56:967–972.
- Featherstone RL, Brown MM, Coward LJ. International carotid stenting study: protocol for a randomised clinical trial comparing carotid stenting with endarterectomy in symptomatic carotid artery stenosis. *Cerebrovasc Dis*. 2004;18:69–74.
- Nicolaides A, Beach KW, Kyriacou E, Pattichis CS. *Ultrasound and Carotid Bifurcation Atherosclerosis*. Springer Science & Business Media; 2011.
- Saba L, Montisci R, Molinari F, et al. Comparison between manual and automated analysis for the quantification of carotid wall by using sonography A validation study with CT. *Eur J Radiol*. 2012;81:911–918.
- Araki T, Kumar PK, Suri HS, et al. Two automated techniques for carotid lumen diameter measurement: regional versus boundary approaches. *J Med Syst*. 2016;40:1–19.
- Saba L, Araki T, Krishna Kumar P, et al. Carotid inter-adventitial diameter is more strongly related to plaque score than lumen diameter: An automated tool for stroke analysis. *J Clin Ultrasound*. 2016;44:210–220.
- Dey N, Bose S, Das A, et al. Effect of watermarking on diagnostic preservation of atherosclerotic ultrasound video in stroke telemedicine. *J Med Syst*. 2016;40:91.
- Ikeda N, Araki T, Dey N, et al. Automated and accurate carotid bulb detection, its verification and validation in low quality frozen frames and motion video. *Int Angiol*. 2014;33:573–589.
- de Korte CL, Hansen HH, van der Steen AF. Vascular ultrasound for atherosclerosis imaging. *Interface Focus*. 2011;1:565–575.
- Stadler RW, Karl WC, Lees RS. New methods for arterial diameter measurement from B-mode images. *Ultrasound Med Biol*. 1996;22:25–34.
- Beux F, Carmassi S, Salvetti MV, et al. Automatic evaluation of arterial diameter variation from vascular echographic images. *Ultrasound Med Biol*. 2001;27:1621–1629.
- Newey VR, Nassiri DK. Online artery diameter measurement in ultrasound images using artificial neural networks. *Ultrasound Med Biol*. 2002;28:209–216.
- Gutierrez MA, Pilon PE, Lage SG, et al. Automatic measurement of carotid diameter and wall thickness in ultrasound images. *Comput Cardiol*. 2002;29:359–362.
- Potter K, Green DJ, Reed CJ, et al. Carotid intima-media thickness measured on multiple ultrasound frames: evaluation of a DICOM-based software system. *Cardiovasc Ultrasound*. 2007;5:29.
- Cinthio M, Jansson T, Ahlgren AR, Lindström K, Persson HW. A method for arterial diameter change measurements using ultrasonic B-mode data. *Ultrasound Med Biol*. 2010;36:1504–1512.
- Sahani AK, Joseph J, Radhakrishnan R, Sivaprakasam M. Automatic measurement of end-diastolic arterial lumen diameter in ARTSENS. *J Med Devices*. 2015;9:041002.
- Sicari R, Galderisi M, Voigt JU, Habib G, et al. The use of pocket-size imaging devices: a position statement of the European Association of Echocardiography. *Eur J Echocardiogr*. 2011;12:85–87.
- Saba L, Banchhor SK, Londhe ND, Araki T, et al. Web-based accurate measurements of carotid lumen diameter and stenosis severity: an ultrasound-based clinical tool for stroke risk assessment during multicenter clinical trials. *Comput Biol Med*. 2017;91:306–317.
- Mughal MM, Khan MK, DeMarco JK, et al. Symptomatic and asymptomatic carotid artery plaque. *Expert Rev Cardiovasc Ther*. 2011;9:1315–1330.
- Polak JF, Kronmal RA, Tell GS, et al. Compensatory increase in common carotid artery diameter. *Stroke*. 1996;27:2012–2015.

35. Eigenbrodt ML, Sukhija R, Rose KM, et al. Common carotid artery wall thickness and external diameter as predictors of prevalent and incident cardiac events in a large population study. *Cardiovasc Ultrasound*. 2007;5:11.
36. Eigenbrodt ML, Bursac Z, Rose KM, et al. Common carotid arterial interadventitial distance (diameter) as an indicator of the damaging effects of age and atherosclerosis, a cross-sectional study of the Atherosclerosis Risk in Community Cohort Limited Access Data (ARICLAD), 1987–89. *Cardiovasc Ultrasound*. 2006;4:1.
37. Jensen-Urstad K, Jensen-Urstad M, Johansson J. Carotid artery diameter correlates with risk factors for cardiovascular disease in a population of 55-year-old subjects. *Stroke*. 1999;30:1572–1576.
38. Molinari FM, Meiburger K, Zeng G, et al. Carotid artery recognition system: a comparison of three automated paradigms for ultrasound images. *Med Phys*. 2012;39:378–391.
39. Molinari F, Krishnamurthi G, Acharya UR, et al. Hypothesis validation of far-wall brightness in carotid-artery ultrasound for feature-based IMT measurement using a combination of level-set segmentation and registration. *IEEE Trans Instrum Meas*. 2012;61:1054–1063.
40. Araki T, Banchhor SK, Londhe ND, et al. Reliable and accurate calcium volume measurement in coronary artery using intravascular ultrasound videos. *J Med Syst*. 2016;40:51.
41. Noor NM, Than JC, Rijal OM, et al. Automatic lung segmentation using control feedback system: morphology and texture paradigm. *J Med Syst*. 2015;39:22.
42. Saba L, Lippo RS, Tallapally N, et al. Evaluation of carotid wall thickness by using computed tomography and semiautomated ultrasonographic software. *J Vasc Ultrasound*. 2011;35:136–142.
43. Saba L, Tallapally N, Gao H, et al. Semiautomated and automated algorithms for analysis of the carotid artery wall on computed tomography and sonography. *J Ultrasound Med*. 2013;32:665–674.
44. Saba L, Gao H, Acharya UR, et al. Analysis of carotid artery plaque and wall boundaries on CT images by using a semi-automatic method based on level set model. *Neuroradiology*. 2012;54:1207–1214.