Inter- and intra-observer variability analysis of completely automated cIMT measurement software (AtheroEdge™) and its benchmarking against commercial ultrasound scanner and

Original
Inter- and intra-observer variability analysis of completely automated cIMT measurement software (AtheroEdge™) and its benchmarking against commercial ultrasound scanner and expert Readers / Luca, Saba; Molinari, Filippo; Meiburger, KRISTEN MARIKO; U., Rajendra Acharya; Andrew, Nicolaides; Jasjit S., Suri. - In: COMPUTERS IN BIOLOGY AND MEDICINE. - ISSN 0010-4825. - ELETTRONICO. - 43:9(2013), pp. 1261-1272. [10.1016/j.compbiomed.2013.06.012]

Publisher:
ELSEVIER

Published
DOI:10.1016/j.compbiomed.2013.06.012

Measurement Error and Inter- and Intra- Observer Variability Analysis of Completely Automated cIMT Measurement Software (AtheroEdge™) and its Benchmarking against Commercial Ultrasound Scanner and Expert Readers

Luca Saba[1,] Filippo Molinari*[,2],Kristen M. Meiburger[2], U. Rajendra Acharya[3],Andrew Nicolaides[4],and Jasjit S. Suri[5,6,7]

[1] Department of Radiology, A.U.O. Cagliari, Cagliari, Italy
[2] Biolab, Dept. Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
[3]Department of ECE, Ngee Ann Polytechnic, Singapore, Singapore
[4] Vascular Diagnostic Center, Nicosia, Cyprus
[5]Fellow AIMBE, CTO,Diagnostic and Monitoring Division, AtheroPoint LLC, CA, USA
and [6]Biomedical Engineering Division, Global Biomedical Technologies Inc., CA, USA
[7] Biomedical Engineering Department, University of Idaho (Aff.), ID, USA

*Corresponding Author:

Professor Filippo Molinari, PhD

Biolab - Department of Electronics and Telecommunications - Politecnico di Torino

CorsoDucadegli Abruzzi, 24; 10129 Torino, Italy;

(e-mail: filippo.molinari@polito.it, tel: +39 11 564 4135)

Abstract

The purpose of this study was to evaluate the measurement error and inter- and intra- observer variability of completely off-line automated and semi-automated carotid intima-media thickness (cIMT) measurement software (AtheroEdge™).

Two hundred carotid ultrasound images from 50 asymptomatic women were analyzed. AtheroEdge™ was benchmarked against a commercial system (Syngo, Siemens) using automated and semi-automated modes. The measurement error and inter- and intra- observer variability of AtheroEdge™ were tested using three readings.

The measurement error of AtheroEdge™ compared to the commercial software was 0.002±0.019 mm (r = 0.99) in the automated mode and -0.001±0.004 mm in the semi-automated mode (r = 0.99). The measurement error of AtheroEdge™ compared to the mean value of the three expert Readers (cIMT bias) for the automated and semi-automated methods was -0.0004±0.158 mm and -0.008±0.157 mm, respectively. The Figure-of-Merit was 99.8% and 99.9% when compared to the commercial ultrasound scanner (using the automated and semi-automated method, respectively) and was 99.9% and 98.9% when compared to the mean value of the three expert Readers. Regarding inter- and intra- observer variability, the intra-class correlation coefficient of the three independent users using the semi-automated AtheroEdge™ was 0.98.

AtheroEdge™ showed a measurement performance comparable to the commercial ultrasound scanner software and the expert Readers' tracings. AtheroEdge™ belongs to a class of automated systems that could find application in processing large datasets for common carotid arteries, avoiding subjectivity in cIMT measurements.

Introduction

The carotid artery intima-media thickness (cIMT) is the most widely used and accepted marker of atherosclerosis [1-3]. The increase of the carotid cIMT was correlated to the incidence of stroke risk even in absence of atherosclerotic plaques [4]. Generally speaking, cIMT is clinically used to assess the subject's cardiovascular risk.

High-resolution ultrasound imaging allows the visualization of the carotid artery and, particularly, the distal carotid wall. It is therefore possible to manually measure the carotid IMT value by delineating the lumen-intima (LI) and media-adventitia (MA) borders and then measuring the distance between the LI and the MA interfaces, the so-called IMT. Clinically, the IMT is manually measured by the sonographer, who puts the calipers at the far wall of the vessel. Recently, some other features of the carotid wall were correlated to risk factors. Lau *et al.*[5] studied the intima thickness, showing correlation with risk score. Kazmierski*et al.*[6] showed that the adventitial thickness was correlated to the cIMT, even though the associated cardiovascular risks were slightly different. In the reported studies, the measurements were manually performed by experts. Manual measurements are time consuming, tedious and subjected to intra- and inter-observer variability.

There has been a growing interest in computerized techniques aiding the clinicians in the IMT measurement based on ultrasound images. The most widely used and performing techniques have been reviewed by Molinari *et al.* in 2010 [7]. From a clinical point of view, recently, Polak *et al.* showed that the common cIMT and internal cIMT values had slightly different predictive values even if both were independently associated with cardiovascular disease [8]. Therefore, it is clearly emerging that in advanced clinical studies, improved measurement techniques are required.

Some recent studies explored the clinical relationships between the cIMT and cardiovascular risk. Particularly, Lau *et al.*[5] studied the repeatability and reproducibility of the intima thickness measurement and showed that it was lower than in cIMT measurements, even though clinically acceptable. This study was based on manual measurements. Polak *et al.*[9] used semi-automated IMT measurement software and showed that manual and computer-based cIMT values both maintained the subjects' cardiovascular risk. They proposed that, in addition to the low error in cIMT measurement, a computerized technique can be of clinical utility if it maintained the subject risk score and proved to be reproducible [9]. The same team also showed that inter-reader variability could preclude the cIMT clinical use in evaluating the cardiovascular risk [10]. They demonstrated that the inter-reader variability is mainly due to inaccuracies in MA tracings that could be spotted by an expert reader. Clearly, an automated computer system could help in decreasing the cIMT inter-reader variability, thus enhancing the clinical applicability of this marker for risk stratification.

The aim of this paper was threefold: First, to compare the cIMT measurement error of the AtheroEdge™ software against the commercial ultrasound scanner using the Syngo software. Second, to compute the inter- and intra- observer variability by calculating the measurement error of AtheroEdge™ software for cIMT measurements against those of three expert Readers who traced the LI/MA borders all along the CCA. Third, three inexperienced users used the AtheroEdge™ system in a semi-automated (user-dependent) mode for computing the inter- and intra- observer variability of AtheroEdge™.

Methods

Patients characteristics and image acquisition protocol

The 50 studied subjects were all asymptomatic postmenopausal Chinese women who were part of the cohort of a longitudinal study to look for subclinical atherosclerosis, but agreed to participate in this study. Each subject was instructed about the aim of the present study and signed an informed consent prior of undergoing ultrasound examination. Ethical approval was given by the The Chinese University of Hong Kong, Hong Kong. The age was 60.2 ± 4.9 years (mean ± SD), ranged between 54 and 67 years. Of these 50 females, 22 had normal blood pressure, total cholesterol and glucose levels in fasting blood. Of the remaining 28 females, one was diabetic, three had hypertension, 15 had hypercholesterolemia, seven had both hypertension and hypercholesterolemia, and two had all three abnormalities. Hypertension was defined as a systolic blood pressure of at least 140 mmHg and/or diastolic blood pressure of at least 90 mmHg and/or pharmacologic treatment. Hypercholesterolemia was defined as a total cholesterol level of at least 5.2 mmol/L and/or pharmacologic treatment. Diabetes was defined as fasting blood glucose level of at least 7.0 mmol/L and/or pharmacologic treatment.

Intima-media thickness of the carotid arteries was examined using a commercial 13-5 MHz linear transducer of the Sonoline Antares ultrasound scanner (Siemens, USA). All images were relative to the longitudinal view of the distal common carotid arteries (both left and right) of the subjects acquired using two different insonation angles: anterior and antero-lateral. The total number of images was 200 (50 subjects, two arteries, and two insonation angles). We used ECG gating for the image acquisition. The end-diastolic cIMT of each image was selected at the R-wave on the ECG waveforms. The one-cm plaque-free cIMT was then computed by the automatic edge detection tracing at the far wall of the common carotid artery just proximal to the dilatation of the bulb [11]. All images were exported to an external computer in lossless JPEG format.

Carotid intima-media thickness (cIMT) measurement techniques

Since the aim of this paper was to quantify the measurement error and inter- and intra- observer variability of AtheroEdge™ software for IMT measurement, we therefore processed all the images of the database using both the AtheroEdge™ software and the commercially available software from Siemens scanner in both (a) automated and (b) semi-automated modes. The comparison of these automatically obtained results with manual measurements by expert readers was used to evaluate the cIMT measurement error of AtheroEdge™. To assess the inter- and intra- observer variability, three different users used AtheroEdge™ in the semi-automated (user-dependent) mode, providing three separate readings. The following subsections will describe the measurement techniques in detail.

cIMT measurements using the Commercial Ultrasound Scanner

The first set of values was acquired by using the Syngo software (Siemens, USA). This software can be used in two modes: (a) fully-automated LI/MA computation in a user defined region of interest and (b) semi-automated, where the user can adjust the LI/MA borders in the known region of interest. The procedure consists of the selection of the region of interest (ROI) 1 cm proximal to the bifurcation and then creating a rectangular box around that region of interest, which was then processed by Syngo commercial software. The Syngo software then automatically searches for the points corresponding to the LI and MA boundaries within this ROI. The distance between the two points is the estimation of the cIMT value. We labeled these results as Siemens (Auto) for our analysis. When used in the semi-automated mode, the Syngo software allows the user to manually adjust the LI/MA points as needed, if the automatic delineation is suboptimal. For analysis, we called this set of cIMT values as Siemens (Semi).

cIMT measurements taken using the AtheroEdge™ Software

AtheroEdge™ is a fully automated platform for the carotid wall delineation and cIMT measurement (from AtheroPoint™ LLC, Roseville, CA, USA). It is a standalone system, which can process images acquired from any scanner in automated and semi-automated modes. In this study, we used AtheroEdge™ to process the images acquired using Siemen's commercial ultrasound scanner while adapting the same protocol as discussed above. We measured a set of cIMT values in the fully automated mode (indicated as AtheroEdge™ (Auto)). As the name says, this is an automated system and hence no user is involved. AtheroEdge™ incorporated an intelligent system that first automatically recognizes the far wall of the carotid artery using a multi-resolution approach followed by automated delineations of LI/MA borders based on a combination of gradient and heuristic information [12] (fig. 1). Multi-resolution is used to exploit the hyper-echoic appearance of the far adventitia layer, thus enabling automatic carotid localization in the presence of the near wall of the CCA and the far and near walls of the Jugular Vein. The actual segmentation stage consists of tracing of the LI/MA boundaries that is based on an integrated approach which combines an edge enhancer using the first-order absolute moment [13] followed by a heuristic approach for LI/MA segmentation and reconstruction [12]. Further, AtheroEdge™ can run in batch mode by processing large databases. Other features include: the ability to read DICOM, BMP and JPEG images; a DBMS system for entering patient demographics, acquisition parameters along with patient history review process images present in the database; the printing of patient reports; saving screen shots of processed images.

In the semi-automated measurement mode, AtheroEdge™ allows the user to place the ROI manually, just like in the commercial semi-automated mode, and then compute the LI/MA borders automatically in this ROI. The user can then adjust any spike or bump in the LI/MA borders using the spike removal option. Manual interaction was made possible by a custom-made graphical user interface, which is currently installable on Windows-based architecture with Java interface. Three inexperienced users measured the cIMT values in the semi-automated mode. The average of the three measurements for each image was considered as the AtheroEdge™ (Semi) value set. The three semi-automated cIMT values by the inexperienced users were further used to assess the inter- and intra- observer variability of the AtheroEdge™ software.

cIMT measurements taken by Three expert Readers using Manual Tracings (Ground Truth)

Three experts sonographers manually traced the LI/MA interfaces of the far wall and then measured the cIMT values of all the images by using a custom made software [14] (ImgTracer™, Global Biomedical Technologies, Inc., Roseville, California, USA). Overall, we had three sets of cIMT values measured by the Readers (we call them as: expert Reader 1, expert Reader 2, and expert Reader 3 from here on).

The cIMT values were calculated starting from the LI/MA profiles by using a specific metric (called Polyline Distance Metric), which computes the average distance between two interfaces [15, 16]. The cIMT value was then converted in mm by calculating the pixel density of the images. We used an automated procedure that scanned the vertical depth scale of the ultrasound image and measured the number of pixels in 1 mm. The inverse of the pixel density is the conversion factor, expressed in mm/pixel, which was multiplied to the value in pixels to express the cIMT values in mm. Overall, for these images the average pixel density was 17.5 ± 24.2 pixel/mm, corresponding to a conversion factor of 0.057 ± 0.041 mm/pixel. We compared AtheroEdge™ to the single expert Reader as well as to the average of the three expert Readers.

Statistical analysis

The statistical structure of our analysis was similar to that of other studies about measurement error and inter- and intra- observer variability of the cIMT measurements[17]. Continuous data were described as the mean value ± Standard Deviation (SD). The Kolmogorov-Smirnov *Z* test was used to test for the distribution normality of each continuous variable group. In case of non-normal distributions, the Wilcoxon test was used to compare the difference among the cIMT sets, whereas if the distributions were normal, we used the Student's paired t-test. The Fisher F-test was used to test the difference in the standard deviation values. The significance level of all the tests was set to 95%. The correlation between groups was calculated by using the Pearson *r* statistic. We evaluated the inter-method agreement using a Bland–Altman analysis. The inter- and intra- observer variability of the measures obtained by using AtheroEdge™ in the semi-automated version was measured by testing the paired difference of the datasets, the limits of agreement, and the intra-class correlation coefficient [18]. *R* software (www.r-project.org) was employed for statistical analyses.

We also analyzed the performance in two modes: automated and semi-automated. We considered a subject as having a high cardiovascular risk factor if their cIMT value (measured by the expert Readers) was equal or higher than 1 mm [19], whereas if a subject had a cIMT lower than 1 mm we considered them as having a lower cardiovascular risk. We then computed the number of patients correctly classified by AtheroEdge™ as having either high or low risk and we computed the following performance indicators [20-22]: sensitivity; specificity; diagnostic accuracy (DA).

Results

Comparison of AtheroEdge™ against the Commercial Ultrasound Scanner

Table 1 reports the comparison of average cIMT values between AtheroEdge™ and the commercial Ultrasound Scanner (Siemens) in both automated and semi-automated modes. As can be noticed from table 1, the difference in the mean IMT value among the techniques is very low and no set was statistically different from the others ($p > 0.05$). Comparing the average cIMT value of AtheroEdge™ and the commercial ultrasound scanner from Siemens in the automated mode, we found a difference of 0.14%, computed as $(100 \cdot (0.704 - 0.703)/0.703)$; coincidently, the percentage difference for the semi-automated method was also 0.14%, computed as ($100 \cdot (0.697 - 0.698)/0.698)$. The bottom row of table 1 reports the 95% limits of agreement for the comparison of the two systems in automated and semi-automated modes. The limits of agreement are larger for the automated mode than for the semi-automated one, but the length of the agreement interval is small compared to the average size of the IMT and to the experimental variability. Therefore, we can conclude that the performance of AtheroEdge™ is nearly identical to the commercial ultrasound scanner cIMT readings.

In Table 2 we compared AtheroEdge™ to the commercial ultrasound scanner on three attributes: (a) cIMT Bias, (b) Correlation Coefficient (CC) and (c) Figure of Merit for the automated and semi-automated modes. Column 1 of Table 2 reports the differences in cIMT measurement for the two sets. The cIMT measurement difference (column two) is of 1 or 2 µm, which is a very encouraging result when compared to the average cIMT value of 0.7 mm (700 µm) (Table 1). Thus, the cIMT bias was 0.3%, a very negligible quantity. The CC was 0.99 with a very small confidence interval, in both modes. The last column of Table 2 reports the Figure-of-Merit (FoM), which is the percent agreement between measurement sets. The FoM can be defined as[23]:

$$FoM = 100 - \left| \frac{\overline{cIMT}_{AtheroEdge} - \overline{cIMT}_{Siemens}}{\overline{cIMT}_{Siemens}} \right| \cdot 100 \tag{1}$$

where $\overline{cIMT}_{AtheroEdge}$ is the average cIMT value measured by AtheroEdge™ and $\overline{cIMT}_{Siemens}$ is the average cIMT value measured by commercial Siemens software. The FoM for AtheroEdge™ was 99.8% with respect to the Siemens system in automated mode and 99.9% in semi-automated mode (last column of Table 2). This showed a nearly perfect agreement between AtheroEdge™ and the Siemens system.

Figure 2 reports the correlation plots comparing the results of the four cIMT sets between AtheroEdge™ and the commercial ultrasound Siemens scanner output for the automated and semi-automated modes. As seen in the correlation plots, 99% of the points are along the linear line in the automated mode and nearly 100% of the points are along the linear line in the semi-automated mode. This further demonstrates the coherency of AtheroEdge™ system compared to the Siemens software.

Figure 3 reports the Bland-Altman plots for AtheroEdge™ w.r.t the commercial ultrasound scanner IMT values. The plots demonstrated the absence of any measurement bias and the cIMT measurement difference was not dependent on the cIMT nominal value. Overall, the graphs showed very high agreement between the AtheroEdge™ software and the commercial ultrasound scanner readings. It can be seen that a one-sided band gap was 0.038/2 mm (19 μm) in automated and 0.008/2 mm (4 μm) in semi-automated mode. The spread function of the cIMT estimates was very low and this clearly showed the robustness of the AtheroEdge™ system for IMT measurement in clinical environments.

Comparison of AtheroEdge™ against Three Expert Readers

Table 3 summarizes the average cIMT values measured by AtheroEdge™ and the average of three expert Readers. AtheroEdge™ (automated and semi-automated) vs. mean of the three expert Readers showed mean cIMT values over 200 images as: 0.704±0.169 mm, 0.697±0.159 mm and 0.705±0.185 mm, respectively. There was a difference in 0.14% between the automated AtheroEdge™ and the average expert Reader values and 1.13% for semi-automated AtheroEdge™ and average expert Reader values.

Table 4 reports the overall performance of AtheroEdge™ in the automated and semi-automated modes in comparison to the average expert Readers' values. The cIMT bias between AtheroEdge™ and the average expert Reader was -0.0004±0.158 mm and -0.008±0.157 mm for automated and semi-automated methods, respectively. It is interesting to see how the expert Readers could sometimes differ. Expert Reader 1 and 3 had a positive bias compared to expert Reader 2 in the automated method. Note that the expert Readers were all experienced radiologists from Italy and were asked to trace the LI/MA borders independently.

Figure 4 and fig. 5 reports the correlation plots of AtheroEdge™ values compared to the expert Readers when compared against the automated and semi-automated modes, respectively. These graphs demonstrate the measurement error of the AtheroEdge™ system in automated and semi-automated modes. There is an overall very encouraging correlation between AtheroEdge™ and the expert Readers' cIMT values. The trend and behavior of AtheroEdge™ in the automated and semi-automated modes in terms of measurement error performance is very comparable.

Figures 6 and 7 report the Bland-Altman plots of AtheroEdge™ compared to the expert Readers, when compared against automated and semi-automated modes, respectively. The Bland-Altman plots show the absence of any trend in the cIMT measurements. This comparison of AtheroEdge™ against the expert Readers' values demonstrated the system performance in the measurement of cIMT.

Inter- and intra-observer variability of AtheroEdge™

The cIMT average values that were obtained by the inexperienced users that used AtheroEdge™ in the semi-automated mode were the following: 0.700 ± 0.158 mm for the first, 0.696 ± 0.162 mm for the second, and 0.696 ± 0.157 mm for the third user. Table 5 summarizes the cIMT values compared to the value obtained by the Siemens commercial system in semi-automated mode.

The three measurement sets were not statistically different (always $p < 10^{-7}$), and they were also not statistically different from the Siemens semi-automated cIMT values ($p < 10^{-12}$). The FoM calculated with respect to the Siemens values was equal to 99.7% for all three users. The correlation coefficient between the users cIMT values and the Siemens (Semi) values was always equal to 0.99 with a confidence interval equal to [0.998 – 0.999].

To assess the inter- and intra-observer variability of the system, we measured the intra-class correlation coefficient, which was equal to 0.98, showing a very low variability or high performance of AtheroEdge™. Also, Table 6 shows the cIMT bias among the three users' readings and the corresponding limits of agreement and correlation coefficients. It can be observed that all cIMT biases are very small (4 μm, 3.9 μm and 0.4 μm, respectively, corresponding to the three users) and that correlation is always higher than 0.98. The highest 95% limit of agreement between two measurement sets was 0.057 mm (57 μm), which is a very low value, indicating the system's high performance.

Preservation of the cardiovascular risk

Even though the issue related to the inter-operator variability could not be the first reason to choose an automated system, nevertheless, Polak *et al.*[9] advised to use computer algorithms to reduce the inter-reader bias, provided that the adopted computer system preserved the cardiovascular risk factor. We therefore analyzed if our AtheroEdge™ system maintained the risk factor of the subjects. We considered a subject as having a high risk factor if its cIMT value (measured by the expert Readers) was equal or higher than 1 mm [19]. Then we observed the AtheroEdge™ cIMT values and we computed the number of subjects classified as high risk. We obtained the cross-tables reported in Table 8. It can be shown that AtheroEdge™ in both the automated and semi-automated modes showed a very good performance in terms of specificity, with values of 95% and 97%, and acceptable performance in terms of sensitivity (values of 52% and 51%). The overall diagnostic accuracy was 90%, which is compatible with a clinical use of this software.

Discussion

Our study was designed in order to evaluate the cIMT measurement error and inter- and intra-observer variability of AtheroEdge™ in automated and semi-automated modes. This system performs the carotid far wall delineation and cIMT measurement in fully automated and semi-automated modes. We benchmarked AtheroEdge™ against the commercially available software by Siemens, which can be used in both an automated and semi-automated mode as well. Further, three expert operators manually segmented the 200 images of the database in order to obtain mean expert Reader's value of the cIMT. Three readings were taken using AtheroEdge™ software in the semi-automated mode for evaluating the inter- and intra- observer variability of the AtheroEdge™ software.

The AtheroEdge™ and Siemens system showed cIMT values that were highly correlated. The correlation between the cIMT values from three users was always higher than 0.99. When the systems were compared in semi-automated modes, the correlation coefficient was very close to 1.0 (Table 2).

Similar performance was obtained when comparing AtheroEdge™ values with the mean of the three expert Readers values. When compared to expert Reader 1, AtheroEdge™ over-estimated the cIMT by 0.058 ± 0.146 mm; compared to expert Reader 2, it under-estimated the cIMT by -0.060 ± 0.170 mm; compared to expert Reader 3 it overestimated by 0.002±0.176 mm (see Table 3). On an average, the AtheroEdge™ bias compared to the average expert Reader's was -0.0004±0.158mm and -0.008±0.157 mm for automated and semi-automated methods, respectively (gray rows of Table 3). Thus the AtheroEdge™ performance was also in agreement with expert Readers. The Bland-Altman plots can be seen in fig. 6 and 7. These values are in line with most of the previously published semi-automated algorithms [24-28] and slightly better than some other recent techniques [23, 25].

Table 7 reports the performance indicators (cIMT bias, correlation coefficient, and FoM) for the Commercial system. Siemens (Auto) software showed a cIMT bias equal to 0.056 ± 0.144 mm, -0.062 ± 0.169 mm, -0.0001±0.175 mm, and -0.002±0.157 mm when compared to expert Reader 1, expert Reader 2, expert Reader 3, and the average of the three expert Readers, respectively. The Siemens (Semi) software showed a cIMT bias equal to 0.052 ± 0.144 mm, -0.067 ± 0.171 mm, -0.004±0.175 mm and -0.007±0.157 mm when compared to expert Reader 1, expert Reader 2, expert Reader 3, and the average of the three expert Readers, respectively. The performance was similar to that of AtheroEdge™ both for the cIMT bias average value and for the standard deviation ($p > 0.05$). Moreover, both the correlation coefficients and the FoM values were identical to those of

AtheroEdge™. The cIMT values of the expert Reader 1 differed from that of the expert Reader 2 of -0.118 ± 0.079 mm. The difference between the expert Reader 1 and the expert Reader 3 was -0.056 ± 0.086 mm, whereas the difference between the expert Reader 2 and the expert Reader 3 was 0.062 ± 0.073 mm. The main reason for the expert Readers to have a difference in their LI/MA manual tracings was due to the noise level in these images. This was caused by the particular acquisition procedure adopted. In fact, all arteries were scanned from two different insonation angles. For some patients, such angles were suboptimal and caused a high level of blood backscattering. Such condition brought to an increase in the variability of human tracings [29].

Figure 8 shows an example of AtheroEdge™ LI/MA tracings compared to expert Reader 1 (fig. 8.A), and expert Reader 2 (fig. 8.B). The white dashed rectangle is placed proximal to the artery bifurcation. Figures 8.C and 8.D show the zoomed area of fig. 8.A, and 8.B, respectively. It can be noticed that the main differences between AtheroEdge™ and the expert Reader's tracing can be seen near the bulb. This was caused by the curvature of the arterial wall at the bifurcation. Even though this problem was observed in less than 10 out of 200 images, this brought a slight increase in the overall cIMT bias. Figure 8 also showed that the expert Readers traced different LI/MA profiles. Specifically, expert Reader 1 traced the LI/MA profiles in correspondence to the bifurcation zone (fig. 8.C), thus increasing the overall CIMT values, which was equal to 0.830 mm. Since expert Readers 2 and 3 did not trace the LI/MA profiles in proximity of the bifurcation, the cIMT values they measured were 0.711 mm and 0.710 mm, respectively. Therefore, in this image, the difference among the Readers was equal to about 0.119 mm. Despite having some critical issues in some aspects, we nevertheless thought this dataset was a suitable benchmark for our system, because it came from a real clinical screening study where standard echo-cardiographic acquisition approaches were followed but still the acquired images were suboptimal images.

From a technical point of view, AtheroEdge™ is a computer-based platform to process images coming from any ultrasound scanner. This is an advantage with respect to most of the currently available systems, which are usually customized for a specific scanner. Also, compared to the Siemens cIMT system, AtheroEdge™ from AtheroPoint™ can automatically identify the carotid artery far wall in the presence of the jugular vein and automatically delineate LI/MA interfaces with very high precision. The AtheroEdge™ has already been tested on a multi-institutional and multi-ethnic database, showing a cIMT bias of 0.078 ± 0.112 mm when tested on a database of 365 images acquired by four different institutions [12]. We did not include all of the 365 images in this study because 165 images were not acquired by the Siemens ultrasound system and therefore we did not have the Siemens' cIMT values to compare. Also, the other images were not processed by multiple expert Readers in order to test the inter- and intra- observer variability. Thus, only 200 images were used for this special study and comparison. However, in the future we intend to install our AtheroEdge™ system in clinics for large population analysis and standard comparisons. The evaluation on a large and heterogeneous image dataset is a necessary and sufficient condition to assess the actual clinical usefulness of a measurement system. The current results are very encouraging since, to the best of our knowledge, this is the first study for a completely automated system and its bench testing against a commercial ultrasound system like Siemens.

Recently, there has been a debate about the advantages of automated systems in atherosclerosis studies. In 2011, Peters *et al.*[30] studied the differences between manual (caliper-based) and semi-automated in cIMT measurements. They used the data coming from the METEOR study [31] and performed cIMT measurements by using semi-automated and manual software. They observed that cIMT semi-automated measurements showed the same measurement error and inter- and intra-observer variability compared to manual measurements. Also, the patients enrolled in the METEOR study were repeatedly scanned during time. Hence, each patient was associated with a specific rate of change of the cIMT, which was computed based on the serial measurements in time. Peters *et al.* showed that the semi-automated software showed same rates of change in time compared to manual measurements. They concluded that the choice between semi-automated and manual reading should be based on costs and time required to complete the analysis, because the quality of the measured data was similar. Compared to Peters *et al.* semi-automated and manual study, we tested our software using both automated and semi-automated modes and on healthy patients. Our results of measurement error and inter- and intra- observer variability are in line with Peter's results, except that we used automated and semi-automated methods and our benchmark testing was against a commercial ultrasound scanner and three expert Readers. Further, we demonstrated that our automated system clinically showed the same performance as the semi-automated system, with a percent difference in the cIMT values equal to 0.14%. Thus, the use of a computer system for cIMT measurement should take into account costs and time issues and we think that the adoption of a fully automated computer system could be considered in large multi-center and/or epidemiological studies. Clearly, full automation allows a substantial saving of time and reduces the costs associated to the time the operators require for manual measurements.

Polak *et al.*[9] advised to use computer algorithms to reduce the inter-reader bias, provided that the adopted computer system preserved the cardiovascular risk factor. AtheroEdge™ in automated and semi-automated modes showed a very good performance in terms of specificity, with values of 95% and 97%. The sensitivity was about 50%, which can be considered a sufficient performance. Haq *et al.*[32] showed that often single indicators of coronary artery disease alone had either sensitivity or specificity equal to about 50% and advised to insert more indexes to correctly score the overall risk. Since the cIMT value is never used alone in scoring the cardiovascular risk of patients, we therefore believe that a sensitivity of 50% is a compatible performance with clinical use.

Conclusions

In this study we compared AtheroEdge™ (courtesy of AtheroPoint LLC, CA, USA), a novel system for cIMT measurement, against a commercially available ultrasound scanner from Siemens.

We showed that the performance of the two systems was comparable, and that there were no differences in the cIMT values measured in the automated and semi-automated modes. The intra-class coefficient of 0.98 showed that the inter- and intra- observer variability of the system was low even when used in semi-automated mode by inexperienced users. We also benchmarked cIMT measurements from AtheroEdge™ against the mean expert Readers cIMT values and showed similar performance to the commercial ultrasound scanner. Further, AtheroEdge™ showed a diagnostic accuracy of 90% when used to score the subject's cardiovascular risk. We conclude that AtheroEdge™ belongs to a class of automated clinical systems that could find application in the automated processing of large datasets of vascular images.

Acknowledgements

The Authors would like to thank Dr. Sin Yee Stella Ho from the Dept. of Imaging and Interventional Radiology, The Chinese University of Hong Kong for having provided the images and the patients' demographics.

References

1. J. Roquer, T. Segura, J. Serena, and J. Castillo, Endothelial dysfunction, vascular disease and stroke: the ARTICO study*,* Cerebrovasc Dis, 27 Suppl 1, 25-37, (2009).
2. P.M. Rothwell, R.J. Gibson, J. Slattery, and C.P. Warlow, Prognostic value and reproducibility of measurements of carotid stenosis. A comparison of three methods on 1001 angiograms. European Carotid Surgery Trialists' Collaborative Group*,* Stroke, 25(12), 2440-4, (1994).
3. P.M. Rothwell and C.P. Warlow, Prediction of benefit from carotid endarterectomy in individual patients: a risk-modelling study. European Carotid Surgery Trialists' Collaborative Group*,* Lancet, 353(9170), 2105-10, (1999).
4. M. Rosvall, L. Janzon, G. Berglund, G. Engstrom, and B. Hedblad, Incidence of stroke is related to carotid IMT even in the absence of plaque*,* Atherosclerosis, 179(2), 325-31, (2005).
5. K.H. Lau, Y.K. Fung, Y.T. Cheung, W.K. Tsang, and M. Ying, Repeatability and reproducibility of ultrasonographic measurement of carotid intima thickness*,* J Clin Ultrasound, 40(2), 79-84, (2011).
6. R. Kazmierski, C. Watala, E. Podsiadly, J. Dorszewska, and W. Kozubski, Association of atherosclerotic risk factors with carotid adventitial thickness assessed by ultrasonography*,* J Clin Ultrasound, 37(6), 333-41, (2009).
7. F. Molinari, G. Zeng, and J.S. Suri, A state of the art review on intima-media thickness (IMT) measurement and wall segmentation techniques for carotid ultrasound*,* Computer Methods and Programs in Biomedicine, 100, 201-221, (2010).
8. J.F. Polak, M.J. Pencina, A. Meisner, K.M. Pencina, L.S. Brown, P.A. Wolf, and R.B. D'Agostino, Sr., Associations of Carotid Artery Intima-Media Thickness (IMT) With Risk Factors and Prevalent Cardiovascular Disease: Comparison of Mean Common Carotid Artery IMT With Maximum Internal Carotid Artery IMT*,* J Ultrasound Med, 29(12), 1759-68, (2010).
9. J.F. Polak, M.J. Pencina, D. Herrington, and D.H. O'Leary, Associations of edge-detected and manual-traced common carotid intima-media thickness measurements with framingham risk factors: the multi-ethnic study of atherosclerosis*,* Stroke, 42(7), 1912-6, (2011).
10. J.F. Polak, L.C. Funk, and D.H. O'Leary, Inter-reader differences in common carotid artery intima-media thickness: implications for cardiovascular risk assessment and vascular age determination*,* J Ultrasound Med, 30(7), 915-20, (2011).
11. R.H. Yu, S.C. Ho, S.S. Ho, S.S. Chan, J.L. Woo, and A.T. Ahuja, Carotid atherosclerosis and the risk factors in early postmenopausal Chinese women*,* Maturitas, 63(3), 233-9, (2009).
12. F. Molinari, C.S. Pattichis, G. Zeng, L. Saba, U.R. Acharya, R. Sanfilippo, A. Nicolaides, and J.S. Suri, Completely automated multiresolution edge snapper--a new technique for an accurate carotid ultrasound IMT measurement: clinical validation and benchmarking on a multi-institutional database*,* IEEE Trans Image Process, 21(3), 1211-22, (2012).
13. M. Demi, M. Paterni, and A. Benassi, The first absolute central moment in low-level image processing*,* Comput Vis Image Und, 80(1), 57-87, (2000).
14. L. Saba, R. Montisci, F. Molinari, N. Tallapally, G. Zeng, G. Mallarini, and J.S. Suri, Comparison between manual and automated analysis for the quantification of carotid wall by using sonography. A validation study with CT*,* Eur J Radiol, 81(5), 911-8, (2012).
15. F. Molinari, G. Zeng, and J. Suri, Inter-Greedy Technique for Fusion of Different Segmentation Strategies Leading to High-Performance Carotid IMT Measurement in Ultrasound Images*,* Journal of Medical Systems, 1-15, (2010).

16. J.S. Suri, R.M. Haralick, and F.H. Sheehan, Greedy algorithm for error correction in automatically produced boundaries from low contrast ventriculograms, Pattern Anal Appl, 3(1), 39-60, (2000).
17. J. Gonzalez, J.C. Wood, F.J. Dorey, T.A. Wren, and V. Gilsanz, Reproducibility of carotid intima-media thickness measurements in young adults, Radiology, 247(2), 465-71, (2008).
18. P.E. Shrout and J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, Psychol Bull, 86(2), 420-8, (1979).
19. B. von Sarnowski, J. Ludemann, H. Volzke, M. Dorr, C. Kessler, and U. Schminke, Common carotid intima-media thickness and framingham risk score predict incident carotid atherosclerotic plaque formation: longitudinal results from the study of health in Pomerania, Stroke, 41(10), 2375-7, (2010).
20. D.G. Altman and J.M. Bland, Diagnostic tests. 1: Sensitivity and specificity, BMJ, 308(6943), 1552, (1994).
21. D.G. Altman and J.M. Bland, Diagnostic tests 2: Predictive values, BMJ, 309(6947), 102, (1994).
22. D.G. Altman and J.M. Bland, Diagnostic tests 3: receiver operating characteristic plots, BMJ, 309(6948), 188, (1994).
23. F. Molinari, W. Liboni, M. Pantziaris, and J.S. Suri, CALSFOAM - completed automated local statistics based first order absolute moment" for carotid wall recognition, segmentation and IMT measurement: validation and bench-marking on a 300 patient database, Int Angiol, 30(3), 227-41, (2011).
24. F. Molinari, G. Zeng, and J.S. Suri, An integrated approach to computer- based automated tracing and its validation for 200 common carotid arterial wall ultrasound images: A new technique, J Ultras Med, 29, 399-418, (2010).
25. F. Molinari, G. Zeng, and J.S. Suri, Intima-media thickness: setting a standard for completely automated method for ultrasound, IEEE Transaction on Ultrasonics Ferroelectrics and Frequency Control, 57(5), 1112-1124, (2010).
26. S. Delsanto, F. Molinari, P. Giustetto, W. Liboni, S. Badalamenti, and J.S. Suri, Characterization of a Completely User-Independent Algorithm for Carotid Artery Segmentation in 2-D Ultrasound Images, Instrumentation and Measurement, IEEE Transactions on, 56(4), 1265-1274, (2007).
27. M.A. Gutierrez, P.E. Pilon, S.G. Lage, L. Kopel, R.T. Carvalho, and S.S. Furuie, Automatic measurement of carotid diameter and wall thickness in ultrasound images, Computers in Cardiology, 29, 359-362, (2002).
28. C.P. Loizou, C.S. Pattichis, M. Pantziaris, T. Tyllis, and A. Nicolaides, Snakes based segmentation of the common carotid artery intima media, Med Biol Eng Comput, 45(1), 35-49, (2007).
29. F. Molinari, K.M. Meiburger, L. Saba, R.U. Acharya, G. Ledda, G. Zeng, S.Y.S. Ho, A.T. Ahuja, S.C. Ho, A. Nicolaides, and J.S. Suri, Ultrasound IMT Measurement on a Multi-Ethnic and Database: Our Review and Experience using Four Fully Automated and one Semi-Automated Methods, Computer Methods and Programs in Biomedicine, ((in press)).
30. S.A. Peters, H.M. den Ruijter, M.K. Palmer, D.E. Grobbee, J.R. Crouse, 3rd, D.H. O'Leary, G.W. Evans, J.S. Raichlen, L. Lind, and M.L. Bots, Manual or semi-automated edge detection of the maximal far wall common carotid intima-media thickness: a direct comparison, J Intern Med, 271(3), 247-56, (2011).
31. J.R. Crouse, 3rd, D.E. Grobbee, D.H. O'Leary, M.L. Bots, G.W. Evans, M.K. Palmer, W.A. Riley, and J.S. Raichlen, Measuring Effects on intima media Thickness: an Evaluation Of Rosuvastatin in subclinical atherosclerosis--the rationale and methodology of the METEOR study, Cardiovasc Drugs Ther, 18(3), 231-8, (2004).

32. I.U. Haq, L.E. Ramsay, P.R. Jackson, and E.J. Wallis, Prediction of coronary risk for primary prevention of coronary heart disease: a comparison of methods*,* QJM, 92(7), 379-85, (1999).

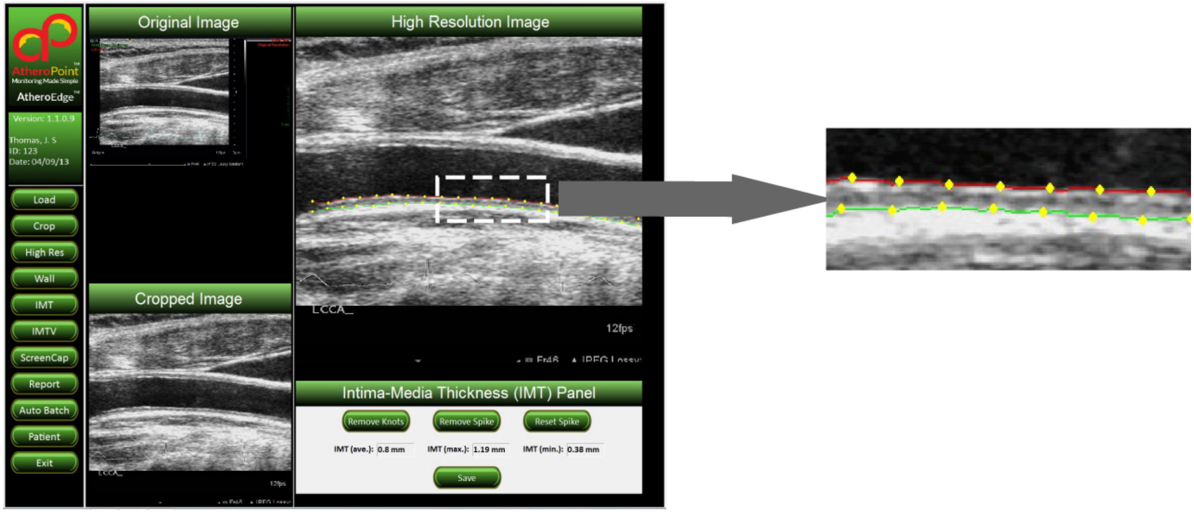
List of Abbreviations

Abbreviation	Meaning
cIMT	Carotid Intima-Media Thickness
LI	Lumen-Intima interface
MA	Media-Adventitia interface
ROI	Region Of Interest
DA	Diagnostic Accuracy
CC	Correlation Coefficient
FoM	Figure-of-Merit

Figure Legends

Figure 1

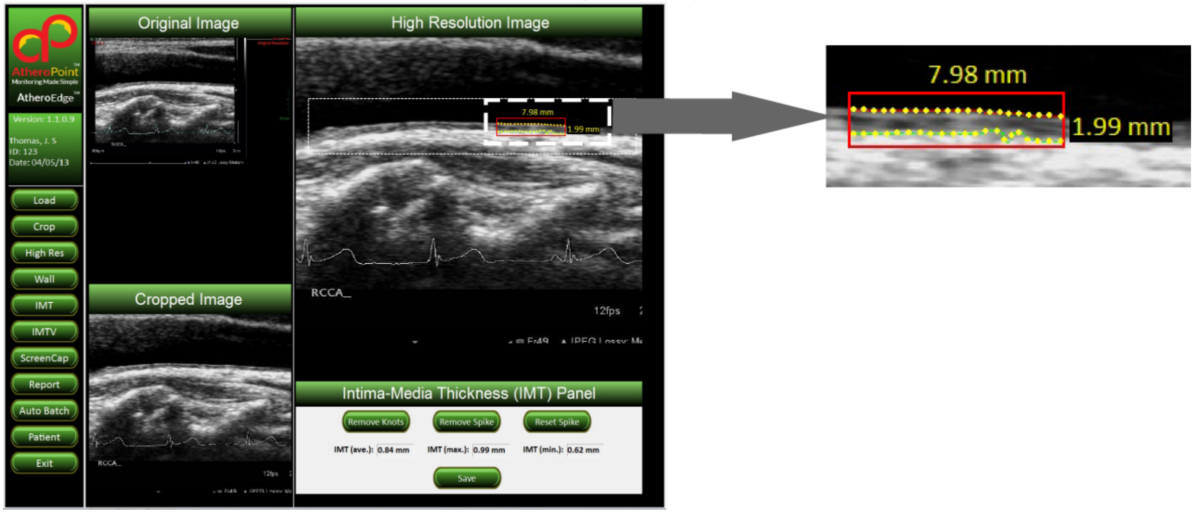
AtheroEdge(TM) (Automated Mode)



Thomas, J. S
ID: 123
Date: 04/09/13

IMT (ave.): 0.8 mm
IMT (max.): 1.19 mm
IMT (min.): 0.38 mm

AtheroEdge(TM) (Semi-Automated Mode)



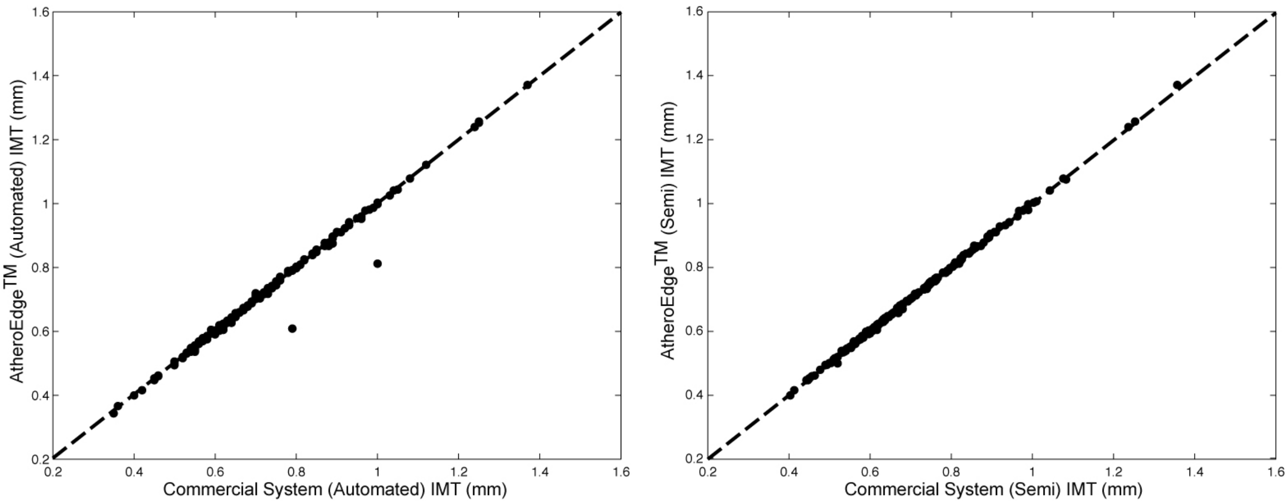
Thomas, J. S
ID: 123
Date: 04/05/13

IMT (ave.): 0.84 mm
IMT (max.): 0.99 mm
IMT (min.): 0.62 mm

Delineation of ultrasound image using AtheroEdge™ software. The figure represents the AtheroEdge™ graphical-user interface (GUI). The top panel is relative to the AtheroEdge™ automated mode, the bottom panel to the AtheroEdge™ Semi-automated mode. The lumen-intima interface is depicted by a red line (LI) and the media-adventitia interfaces by a green line (MA). In the bottom panel, the white dotted box represents the far wall position and the red box is the region-of-interest manually selected by the user. The cIMT measurement is visually presented to the user in the bottom box of the GUI.

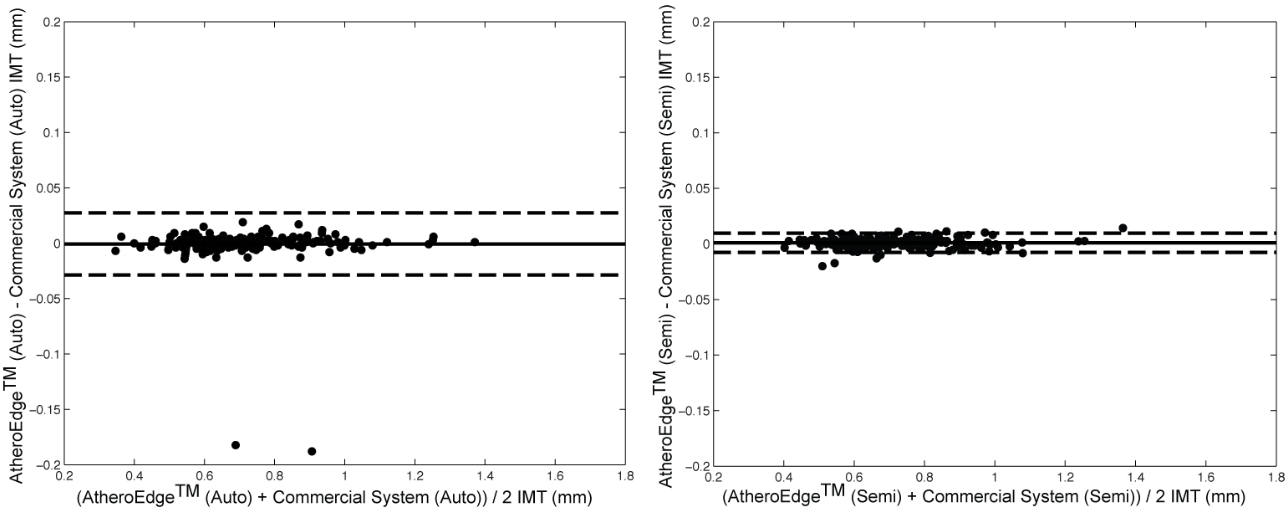
(CCA = common carotid artery)

Figure 2



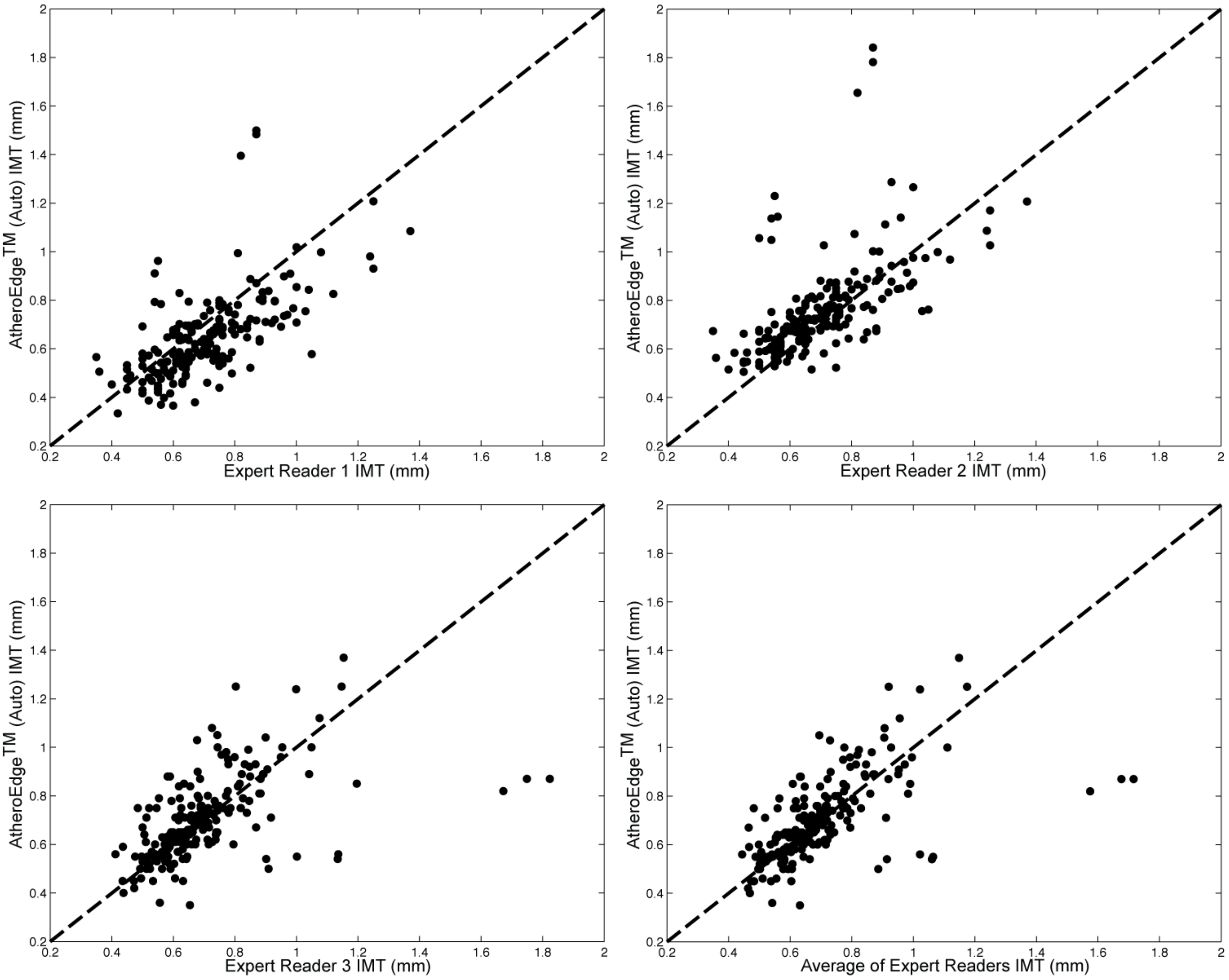
IMT Correlation plots between the AtheroEdge™ cIMT values (vertical axis) and the commercial ultrasound scanner values (horizontal axis). The left panel is relative to the automated mode, the right to the semi-automated mode. The dashed line represents the line of equality.

Figure 3



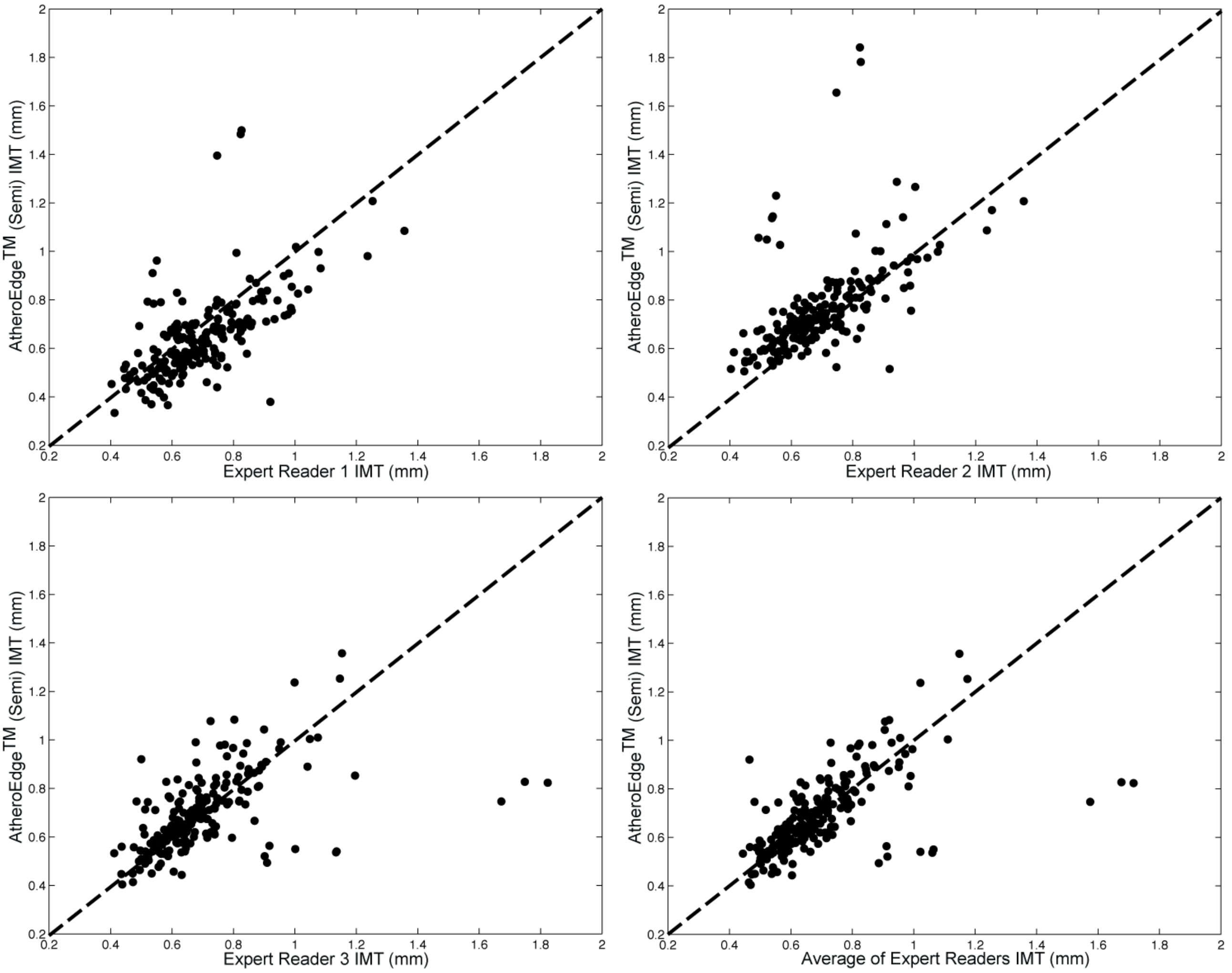
Bland-Altman plots between the AtheroEdge™ cIMT values and the commercial system cIMT values. The left panel is relative to the automated mode, the right to the semi-automated mode.

Figure 4



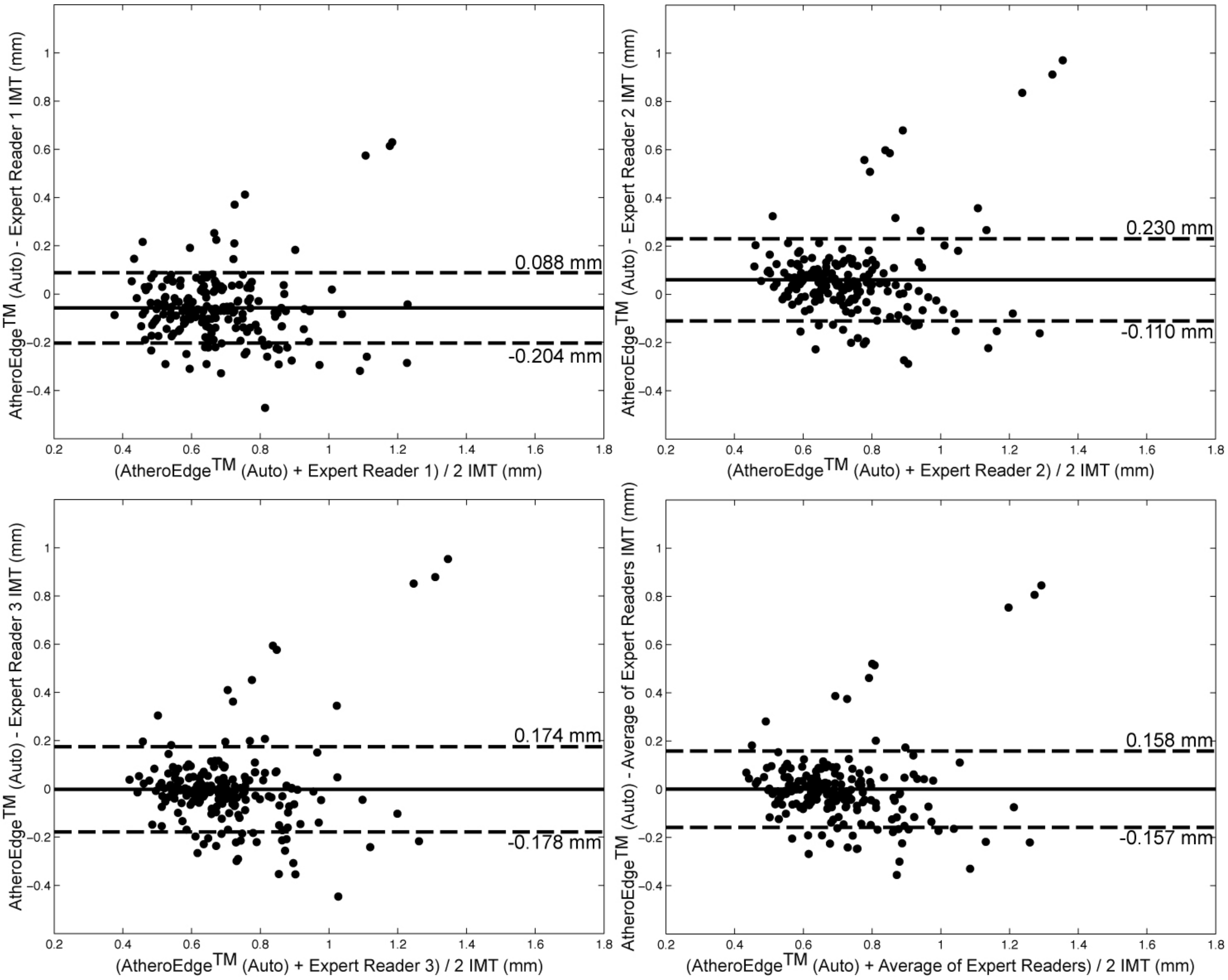
Correlation plots between the AtheroEdge™ automated mode cIMT values (vertical axis) and the Readers' cIMT values (horizontal axis). The dashed line represents the line of equality.

Figure 5



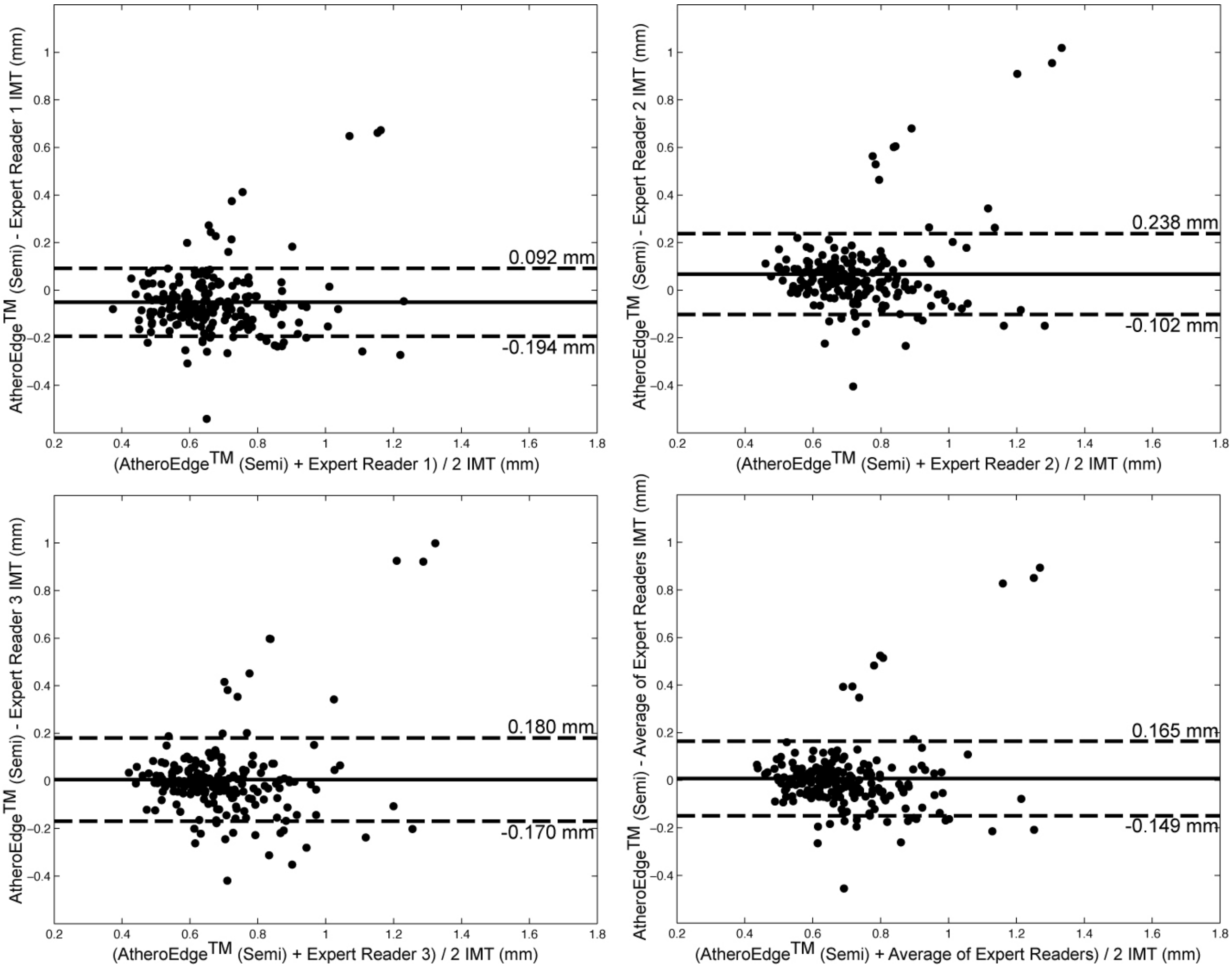
Correlation plots between the AtheroEdge™ semi-automated mode cIMT values (vertical axis) and the Readers' cIMT values (horizontal axis). The dashed line represents the line of equality.

Figure 6



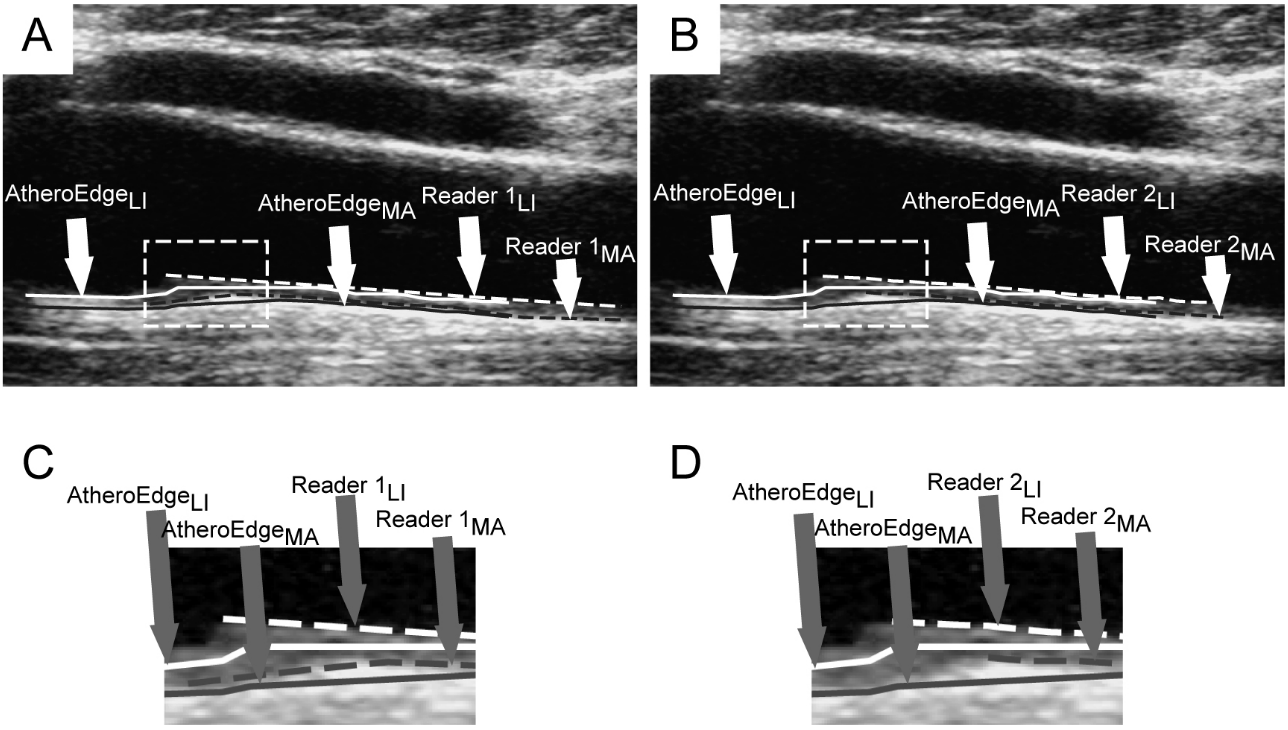
Bland-Altman plots of AtheroEdge™ automated mode cIMT values and the expert Readers' cIMT values.

Figure 7



Bland-Altman plots of AtheroEdge™ semi-automated mode cIMT values and the expert Readers' cIMT values.

Figure 8



Sample of AtheroEdge™ LI/MA tracings (automated mode) compared to the expert Readers LI/MA profiles. The LI profiles are depicted in white, the MA profiles in black. A) Original B-Mode image with AtheroEdge™ LI/MA tracings (continuous lines) compared to expert Reader 1 (dashed lines). B) Original B-Mode image with AtheroEdge™ LI/MA tracings (continuous lines) compared to Reader 2 (dashed lines). C) Expanded view of the carotid bifurcation (bulb – dashed rectangle of panel A) region comparing AtheroEdge™ and expert Reader 1 LI/MA tracings. D) Expanded view of the carotid bifurcation (bulb – dashed rectangle of panel B) region comparing AtheroEdge™ and expert Reader 2 LI/MA tracings.

Table 1. Comparison of mean IMT values for AtheroEdge™ and the commercial Ultrasound Scanner in automated and semi-automated modes. The bottom row reports the 95% limits of agreement between the two measurement sets.

	AtheroEdge™ (Auto)	**Siemens (Auto)**	**AtheroEdge™ (Semi)**	**Siemens (Semi)**
cIMT value (mm)	0.704±0.169	0.703±0.169	0.697±0.159	0.698±0.159
95% Limits of agreement (mm)	[-0.036; 0.039]		[-0.007; 0.009]	

Table 2. Comparison of three attributes: (a) cIMT Bias, (b) Correlation Coefficient and (c) Figure of Merit for AtheroEdge™ and commercial Ultrasound Scanner in automated and semi-automated modes. The confidence interval for the correlation coefficient is reported between brackets.

	cIMT bias (mm)	**Correlation Coefficient**	**FoM**
AtheroEdge™ Auto Vs. Siemens Auto	0.002±0.019	0.99 [0.992-0.995]	99.8%
AtheroEdge™ Semi Vs. Siemens Semi	-0.001±0.004	0.99 [0.999-1.000]	99.9%

Table 3. Comparison of cIMT values from AtheroEdge™ (automated and semi-automated methods) against expert Reader 1, expert Reader 2, expert Reader 3, and the average of the three expert Readers.

	AtheroEdge™ Auto	**AtheroEdge™ Semi**	**expert Reader 1**	**expert Reader 2**	**expert Reader 3**	**Average of the expert Readers**
cIMT value (mm)	0.704±0.169	0.697±0.159	0.647±0.176	0.765±0.198	0.703±0.196	0.705±0.185

Table 4. Comparison of three attributes: (a) cIMT Bias, (b) Correlation Coefficient and (c) Figure of Merit for AtheroEdge™ and expert Readers in automated and semi-automated modes. The confidence interval for the correlation coefficient is reported between brackets.

	cIMT bias (mm)	**Correlation coefficient**	**FoM**
AtheroEdge™ Auto Vs. expert Reader 1	0.058±0.146	0.64 [0.56-0.72]	91.1%
AtheroEdge™ Auto Vs. expert Reader 2	-0.060±0.170	0.58 [0.48-0.66]	92.1%
AtheroEdge™ Auto Vs. expert Reader 3	0.002±0.176	0.54 [0.44-0.63]	99.8%
AtheroEdge™ Auto Vs. Average of the expert Readers	-0.0004±0.158	0.60 [0.51-0.68]	99.9%
AtheroEdge™ Semi Vs. expert Reader 1	0.051±0.143	0.64 [0.55-0.72]	92.2%
AtheroEdge™ Semi Vs. expert Reader 2	-0.068±0.170	0.56 [0.46-0.65]	91.2%
AtheroEdge™ Semi Vs. expert Reader 3	-0.005±0.175	0.53 [0.42-0.62]	99.2%
AtheroEdge™ Semi Vs. Average of the expert Readers	-0.008±0.157	0.59 [0.50-0.68]	98.9%

Table 5. cIMT values for the three independent inexperienced users for inter- and intra-variability estimation using AtheroEdge™ in comparison to Siemens cIMT (semi-automated mode).

	User 1 (semi)-mm	**User 2 (semi)-mm**	**User 3 (semi)-mm**	**Siemens (Semi)-mm**
cIMT value (mm)	0.700±0.158	0.696±0.162	0.696±0.157	0.698±0.159

Table 6. Inter- and Intra- variability measures of the AtheroEdge™ users' cIMT values.

	cIMT bias (mm)	**95% Superior Limit of Agreement (mm)**
User 1 Vs. User 2	0.0040±0.023	0.045
User 1 Vs. User 3	0.0039±0.021	0.041
User 2 Vs. User3	0.0004±0.029	0.057

Table 7. Comparison of three attributes: (a) cIMT Bias, (b) Correlation Coefficient and (c) Figure of Merit for the Commercial System and expert Readers in automated and semi-automated modes. The confidence interval for the correlation coefficient is reported between brackets.

	cIMT bias (mm)	**Correlation coefficient**	**FoM**
Siemens Auto Vs. expert Reader 1	0.056±0.144	0.65 [0.57-0.73]	91.3%
Siemens Auto Vs. expert Reader 2	-0.062±0.169	0.59 [0.49-0.67]	91.8%
Siemens Auto Vs. expert Reader 3	-0.0001±0.175	0.55 [0.44-0.64]	99.9%
Siemens Auto Vs. Average of the expert Readers	**-0.002±0.157**	**0.61 [0.52-0.69]**	**99.6%**
Siemens Semi Vs. expert Reader 1	0.051±0.143	0.64 [0.55-0.72]	92.0%
Siemens Semi Vs. expert Reader 2	-0.067±0.171	0.56 [0.46-0.65]	91.3%
Siemens Semi Vs. expert Reader 3	-0.004±0.175	0.53 [0.42-0.62]	99.4%
Siemens Semi Vs. Average of the expert Readers	**-0.007±0.157**	**0.59 [0.49-0.67]**	**99.1%**

Table 8. Cross-table for the risk factor assessment of AtheroEdge™ in automated and semi-automated modes. The Diagnostic Accuracy (DA) was between 0 and 1.

	Sensitivity	**Specificity**	**PPV**	**NPV**	**DA**
AtheroEdge™ (Auto)	0.52	0.95	0.63	0.92	0.89
AtheroEdge™ (Semi)	0.51	0.97	0.75	0.92	0.90