Carotid Ultrasound Boundary Study (CUBS): An Open Multicenter Analysis of Computerized Intima–Media Thickness Measurement Systems and Their Clinical Impact

Ultrasound in Med. & Biol., Vol. 00, No. 00, pp. 1–14, 2021

0301-5629/$ - see front matter

https://doi.org/10.1016/j.ultrasmedbio.2021.03.022

● *Original Contribution*

CAROTID ULTRASOUND BOUNDARY STUDY (CUBS): AN OPEN MULTICENTER ANALYSIS OF COMPUTERIZED INTIMA–MEDIA THICKNESS MEASUREMENT SYSTEMS AND THEIR CLINICAL IMPACT

Kristen M. Meiburger,* Guillaume Zahnd,† Francesco Faita,‡ Christos P. Loizou,§ Catarina Carvalho,¶ David A. Steinman,‖ Lorenzo Gibello,# Rosa Maria Bruno,**,†† Francesco Marzola,* Ricarda Clarenbach,‡‡ Martina Francesconi,‡,** Andrew N. Nicolaides,§§ Aurelio Campilho,¶,¶¶ Reza Ghotbi,†† Efthyvoulos Kyriacou,‖‖ Nassir Navab,†,## Maura Griffin,*** Andrie G. Panayiotou,††† Rachele Gherardini,** Gianfranco Varetto,# Elisabetta Bianchini,‡ Constantinos S. Pattichis,‡‡‡ Lorenzo Ghiadoni,** José Rouco,§§§,¶¶¶ and Filippo Molinari*

* PolitoBIOmed Lab, Biolab, Department of Electronics and Communications, Politecnico di Torino, Torino, Italy; † Computer Aided Medical Procedures, Technische Universität München, München, Germany; ‡ Institute of Clinical Physiology, Italian National Research Council, Pisa, Italy; § Department of Computer Science, University of Cyprus, Nicosia, Cyprus; ¶ INESC Technology and Science, Porto, Portugal; ‖ Biomedical Simulation Lab, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada; # Dipartimento di Scienze Chirurgiche, University of Torino, Torino, Italy; ** Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy; †† INSERM U970, Paris Cardiovascular Research Centre—PARCC and Université de Paris, Paris, France; ‡‡ Helios Klinikum München West, München, Germany; §§ Cyprus Cardiovascular Disease Educational Research Trust, Nicosia, Cyprus; ¶¶ FEUP—Faculty of Engineering, University of Porto, Porto, Portugal; ‖‖ Department of Computer Science and Engineering, Frederick University, Limassol, Cyprus; ## Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, Maryland, USA; *** The Vascular Screening and Diagnostic Centre, Nicosia, Cyprus; ††† Cyprus International Institute for Environmental and Public Health, Cyprus University of Technology, Limassol, Cyprus; ‡‡‡ Department of Computer Science and Biomedical Engineering Research Center, University of Cyprus, Nicosia, Cyprus; §§§ Research Center of Information and Communication Technologies, UDC, A Coruña, Spain; and ¶¶¶ Department of Computer Science, University of A Coruña, A Coruña, Spain

(*Received* 15 *September* 2020; *revised* 11 *March* 2021; *in final from* 19 *March* 2021)

Abstract—Common carotid intima–media thickness (CIMT) is a commonly used marker for atherosclerosis and is often computed in carotid ultrasound images. An analysis of different computerized techniques for CIMT measurement and their clinical impacts on the same patient data set is lacking. Here we compared and assessed five computerized CIMT algorithms against three expert analysts' manual measurements on a data set of 1088 patients from two centers. Inter- and intra-observer variability was assessed, and the computerized CIMT values were compared with those manually obtained. The CIMT measurements were used to assess the correlation with clinical parameters, cardiovascular event prediction through a generalized linear model and the Kaplan–Meier hazard ratio. CIMT measurements obtained with a skilled analyst's segmentation and the computerized segmentation were comparable in statistical analyses, suggesting they can be used interchangeably for CIMT quantification and clinical outcome investigation. To facilitate future studies, the entire data set used is made publicly available for the community at http://dx.doi.org/10.17632/fpv535fss7.1. (E-mail: kristen.meiburger@polito.it) **© 2021 The Author(s). Published by Elsevier Inc. on behalf of World Federation for Ultrasound in Medicine & Biology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).**

Key Words: **Carotid artery intima–media thickness, Segmentation, Ultrasound imaging, Atherosclerosis, Cardiovascular events, Open-source database.**

Address correspondence to: Kristen M. Meiburger, c/o Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy. E-mail: kristen.meiburger@polito.it

INTRODUCTION

The common carotid intima−media thickness (CIMT) is measured by delineating the intima−media complex contours and is commonly used as a surrogate marker for atherosclerosis. The CIMT can be estimated via B-mode ultrasound imaging of the common carotid artery (CCA) (Stein et al. 2008), and increased CIMT has been associated with augmented cardiovascular risk (Oren et al. 2003). Importantly, large epidemiological studies have reported the predictive value of CIMT for myocardial infarction or stroke independently of traditional cardiovascular risk factors (Engelen et al. 2013). Furthermore, increased CIMT was associated with future cardiovascular disease events in high-risk participants (Lorenz et al. 2018).

Negative results on the independent predictive value of CIMT for cardiovascular events, however, have been also reported (Plichart et al. 2011). The additional value of the CIMT measurement in risk classification on top of existing risk scores was found to not be significant in one study (Bots et al. 2014). No association between CIMT progression over time and cardiovascular risk in the general population was observed in one study (Lorenz et al. 2018), partly because changes in the vessel wall over time are too small and their assessment could be complicated by measurement error (Lorenz et al. 2018).

A commonly reported limitation of CIMT measurement techniques is the heterogeneity in methodological approaches (Lorenz et al. 2018). CIMT estimation can be influenced by measurement location (*e.g.,* CCA vs. other locations), the protocol used (*e.g.,* single vs. multiple angles acquisition) and reading system features (*e.g.,* transducer central frequency). Furthermore, ultrasound equipment setup (*e.g.,* filters, image gain, depth and time-gain compensation [TGC]) may affect the robustness (Potter et al. 2008) and should be standardized when using B-mode-based systems (Bianchini et al. 2013). Additionally, the majority of large prospective cohort studies on the predictive value of CIMT used manual measurements with calipers, which can be subject to inter-analyst/intra-analyst variability. Reducing manual measurement variability likely comes at the cost of increased time and effort in the assessment and offline reading.

Many computerized methods have been proposed to extract the contours of the intima−media complex and measure CIMT in longitudinal CCA ultrasound images (Meiburger et al. 2018) (Fig. 1), typically reporting an increase in measurement robustness and faster analysis times (Saba et al. 2018). Segmenting the CCA can be described as determining the position of the lumen−intima (LI) and media−adventitia (MA) anatomical interfaces. CIMT quantification is usually performed on the CCA far wall, as recommended by the Mannheim consensus (Touboul et al. 2012). Many segmentation methods have been extensively described, categorized and compared in several dedicated reviews (Molinari et al. 2012a; Loizou 2014; Meiburger et al. 2018). Briefly, methods are typically based on edge operators (Faita et al. 2008; Rocha et al. 2012); active contours, also referred to as snakes (Loizou et al. 2007; Molinari et al. 2012b); dynamic programming (Ilea et al. 2013; Zahnd et al. 2017); adaptive thresholding (Ilea et al. 2013); fuzzy *c*-means clustering (Hassan et al. 2014); and more recently, machine learning approaches (Biswas et al. 2018; Zhou et al. 2019). Although many computerized methods exist, an analysis of different methods and their clinical impacts on the same patient data set is lacking.

The objective of this study was to compare five different computerized CIMT measurement methods on the same large database. Their correlation with clinical parameters used to assess cardiovascular risk was investigated, and potential differences in their ability to predict cardiovascular events were addressed. The main focus of the study was addressing whether the various computerized methods provide measurements comparable with those performed manually. The database used in this study, including the images used, clinical data, manual segmentations and computerized segmentations of all analyzed methods, has been made publicly available (http://dx.doi.org/10.17632/fpv535fss7.1).

METHODS

Database description

One thousand eighty-eight participants were analyzed in this study, which complies with the Declaration of Helsinki. CCA ultrasound images were acquired from both sides of the neck (2176 total images). The images were acquired in the villages of Pedoulas, Nissou and Kambos in Cyprus between 2003 and 2007 (694 participants, 1388 images) or at the Hypertension Outpatient Clinic of the University in Pisa between 2011 and 2014 (394 participants, 788 images). All inhabitants from the three villages in Cyprus were identified through the population list, and those over the age of 40 were invited to participate. The ethics committee of the Cyprus Institute of Neurology and Genetics approved the study. The baseline results from the first two villages were published as a cross-sectional study (Griffin et al. 2009). The participants from Pisa were originally enrolled in two studies, which were approved by the institutional ethics committee and whose results have already been published (Bruno et al. 2017, 2018). No clinical follow-up was planned for the Pisa studies. All participants

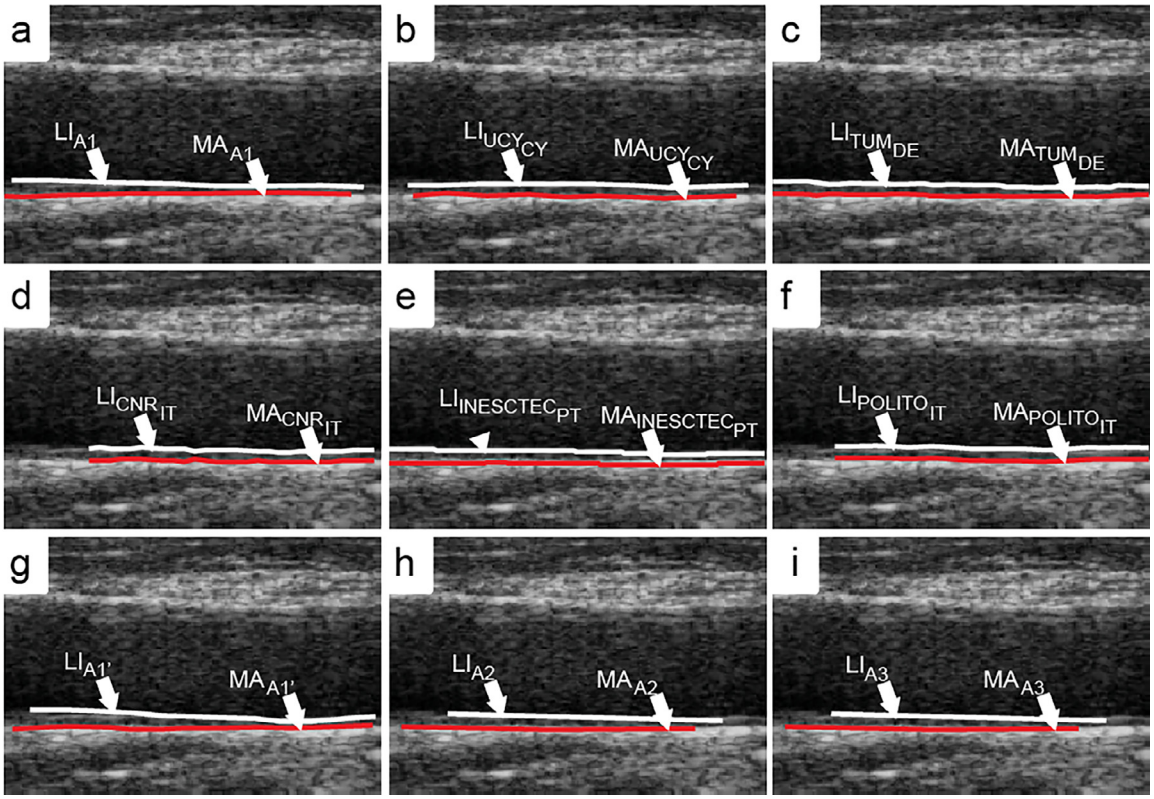


Fig. 1. Example of manual and computerized segmentation results (lumen−intima [LI] border: *white*, media−adventitia [MA] border: *red*). A1, A2 and A3 = manual segmentations of analysts 1 (a), 2 (h) and 3 (panel i), respectively; A1′ = segmentations of analyst 1, traced 1 month after A1 (g); IMT = intima−media thickness; UCY_{CY} = computerized method from University of Cyprus; TUM_{DE} = method from Technische Universität München; CNR_{IT} = method from Consiglio Nazionale delle Ricerche; $INESCTEC_{PT}$ = method from INESC Technology and Science; $POLITO_{IT}$ = method from Politecnico di Torino.

Table 1. Clinical characteristics of all participants

Parameter	Entire database	Cyprus	Pisa
No. of participants (images)	1088 (2176)	694 (1388)	394 (788)
Ultrasound instrumentation		Philips (ATL) HDI 5000 duplex scanner	MyLab25 (Esaote)
Age (y)	62 ± 11	61 ± 10	64 ± 13
Sex	546 females (50%)	316 females (46%)	226 females (57%)
Smoking (pack-years)*	418 (unknown)	260 (36 ± 33)	158 (unknown)
Hypertension	582 (53%)	242 (35%)	340 (86%)
Diabetes	167 (15%)	91 (13%)	76 (19%)
Body mass index (kg/m^2)	28.02 ± 4.48	28.05 ± 4.49	27.97 ± 4.47
Glucose (mg/dL)	104.49 ± 31.31	103.72 ± 27.67	106.06 ± 37.64
Total cholesterol (mg/dL)	218.47 ± 42.16	226.64 ± 42.76	202.61 ± 36.08
HDL cholesterol (mg/dL)	51.55 ± 14.07	50.35 ± 12.53	53.89 ± 16.44
LDL cholesterol (mg/dL)	131.48 ± 32.05	136.08 ± 30.64	122.45 ± 32.88
Triglycerides (mg/dL)	146.35 ± 85.93	151.64 ± 90.97	135.99 ± 74.14
Creatinine (mg/dL)	0.93 ± 0.24	0.93 ± 0.25	0.92 ± 0.23
Apolipoprotein A1 (mg/L)	—	1.44 ± 0.24	—
Apolipoprotein B (mg/L)	—	1.20 ± 0.24	—
Follow-up cardiovascular events	—	127 (18%)	—

Values are expressed as the number (%) or mean ± standard deviation unless stated otherwise.
HDL = high-density lipoprotein; LDL: low-density lipoprotein.
* Number of cigarette packs smoked per day multipled by years of smoking.

provided written informed consent including acceptance that the images can be used for future studies. The Mannheim consensus guidelines for image acquisition were followed for all participants (Touboul et al. 2012). Other clinical parameters were also recorded for each participant (Table 1). The Cyprus participants were followed up until 2017 (mean follow-up: 11 ± 3 y), and follow-up cardiovascular events or deaths were registered. All scans from Cyprus were performed using a Philips (ATL) HDI 5000 duplex scanner (Seattle, WA, USA), with a broadband L12-5 MHz linear array transducer. Several images were saved for each subject, and the best ones were chosen by the operator who acquired the images and then used in this study. Ultrasound parameters were preset (central TGC, 170-dB dynamic range, low persistence, high frame rate) and remained constant for all acquisitions. All scans from Pisa were obtained by a trained operator using a MyLab25 device (ESAOTE, Florence, Italy) with a LA523 4-13 MHz linear array transducer. For each side, one anterior scan and one lateral scan using 10-s clips were saved, and the best frame was extracted for analysis. A dedicated preset (central TGC, average: 96 dB, range: 84 dB−102 dB, dynamic range, no persistence) was created and used for all exams. For both data acquisition protocols, the best image or frame chosen was one that presented the highest qualitative visual contrast between the carotid lumen and intima−media complex.

The pixel dimension of the images included in the entire data set presented a mode equal to 0.064 mm/pixel, with a minimum of 0.038 mm/pixel and a maximum of 0.267 mm/pixel. The pixel dimension, also referred to as the calibration factor, of each image is included in the publicly available data set.

Manual and automated CIMT measurement methods

For all 2176 images, both manual and computerized measurement methods were performed.

Manual measurements. A gold standard reference was generated to evaluate the accuracy of each segmentation method, despite the lack of absolute ground truth inherent to ultrasound *in vivo* data. Manual reference tracing annotations were performed on each image by an experienced analyst A1 (L.Gi. from Torino, >10 y of experience in carotid sonography) and were considered the gold standard. The full exploitable width of each image was determined to (i) exclude regions of potential poor image quality, and (ii) follow the Mannheim consensus guidelines (Touboul et al. 2012). Then, the contours of both LI and MA anatomical interfaces were manually traced within the previously determined width. The entire procedure was performed again by A1 1 month later to assess intra-analyst variability (denoted as A1′) and by two other expert analysts, A2 (G.V., >25 y of experience in carotid sonography) and A3 (for the Cyprus database: M.G., >25 y of experience in carotid sonography; for the Pisa database: L.Gh., >25 y of experience in carotid sonography) to assess inter-analyst variability. The annotations performed by A1, A1′, A2 and A3 were blinded from each other. Manual segmentations were performed with care using a graphical interface developed specifically for this purpose, with the exception of A3 (see Supplementary Data, online only).

Computerized measurements. Five computerized segmentation methods developed by the authors coming from five different research groups were employed on all images. All methods produce the LI and MA tracings; some require user interaction, while others are completely automatic. For simplicity, each technique is named by the institution it derives from with a subscript that indicates the country where the research group is located. The methods are based on the first-order absolute moment [CNR_{IT}] (Faita et al. 2008; Bianchini et al. 2013), anisotropic Gaussian derivative filters [$INESCTEC_{PT}$] (Rocha et al. 2012; Rouco et al. 2018), dynamic programming [TUM_{DE}] (Zahnd et al. 2017, 2019), snakes [UCY_{CY}] (Loizou et al. 2007) and dual snakes [$POLITO_{IT}$] (Molinari et al. 2012b).

Detailed descriptions of each method are provided in the Supplementary Data (online only).

By use of all manual and computerized segmentations, the final CIMT value was computed using the polyline distance method (see Supplementary Data, online only). The coefficient of variation (CV), defined as the ratio between the standard deviation and the mean, was computed for each image to estimate the percentage variation of the CIMT measurement. Three cases were analyzed: (i) all nine CIMT measurement methods, (ii) four manual CIMT measurements and (iii) five computerized CIMT measurements.

Statistical analysis

All statistical analysis in this study was done using R statistical software Version 3.6.2. First, the mean signed difference (bias) and absolute bias of the CIMT values against the manual A1 values were computed as

$$CIMT_{bias} = CIMT_{method} - CIMT_{A1} \quad (1)$$

$$CIMT_{AbsBias} = |CIMT_{method} - CIMT_{A1}| \quad (2)$$

where *method* refers to either the manual CIMT values (*i.e.,* A2, A3, A1′) or the automatic CIMT measurements (*i.e.,* UCY_{CY}, TUM_{DE}, CNR_{IT}, $INESCTEC_{PT}$,

$POLITO_{IT}$). A Wilcoxon paired test was also done to test for statistically significant differences between the CIMT measurements.

The remaining statistical analysis can be divided into three main parts:

Inter-analyst/intra-analyst variability. Investigating inter- and intra-analyst variability is relevant as a certain level of discrepancy is expected between different manual tracings and measurements. Although errors are unavoidable when generating manual references, it is insightful to quantify the agreement between different expert analysts, and to compare it with the degree of accuracy from computerized methods. To assess this variability, the difference between CIMT values was computed, and a regression and Bland−Altman analysis were performed.

Correlation of mean CIMT with clinical parameters. The mean CIMT (average between left and right CIMTs of each participant) was correlated with the clinical parameters available for the entire database (Table 1, first column). The Spearman correlation coefficient, ρ, was computed, and a regression analysis was carried out, computing the β coefficients and 95% confidence intervals.

Prediction of cardiovascular events. A stepwise logistic regression analysis was used to model a generalized linear model (GLM) for the prediction of follow-up cardiovascular events for the Cyprus participants who did not have any baseline cardiovascular event (N = 590). The data set contained 1.7% cases with missing values, concentrated mainly in the blood test parameters, and these entries were removed. Because there was a class imbalance between participants (*i.e.,* 18% of the participants presented a subsequent follow-up event), resampling was done with the automatic synthetic minority oversampling technique method (Chawla et al. 2002). The data set was divided into a training set (80%) and a test set (20%), maintaining the same class imbalance ratio. All clinical parameters and the mean CIMT value were initially included, and a logistic regression with stepwise feature selection was employed (residual sum of squares in both directions with Akaike information criterion). To evaluate the predictive value of the CIMT alone, a Kaplan−Meier survival analysis was done, and the hazard ratios were computed using the CIMT values as the explanatory variable and a cardiovascular event as the outcome, and employing the Cox proportional hazards model. The CIMT values were divided into four distinct groups: <650, 650−750, 750−850 and >850 μm. These values were determined by approximating the 25th, 50th and 75th percentiles of all CIMT measurements (equal to 649, 737 and 849 μm, respectively).

RESULTS

CIMT measurements

The CIMT values obtained are reported in Table 2. The overall values are provided, as are the $CIMT_{bias}$ with 95% confidence intervals and $CIMT_{AbsBias}$ when compared with A1. Specifically, the $CIMT_{AbsBias}$ versus A1 was equal to 171 ± 144 μm, 114 ± 117 μm, 120 ± 123 μm, 202 ± 271 μm and 179 ± 138 μm for the computerized methods (respectively UCY_{CY}, TUM_{DE}, CNR_{IT}, $INESCTEC_{PT}$ and $POLITO_{IT}$). Only the TUM_{DE} and CNR_{IT} CIMT measurements did not have statistically significant differences compared with the A1 CIMT measurements. For the other expert manual tracings, the $CIMT_{AbsBias}$ versus A1 was equal to 140 ± 143 μm, 255 ± 230 μm and 188 ± 180 μm for A1′, A2 and A3, respectively, always with a statistically significant difference. Two computerized methods (CNR_{IT} and $INESCTEC_{PT}$) were not able to process all images

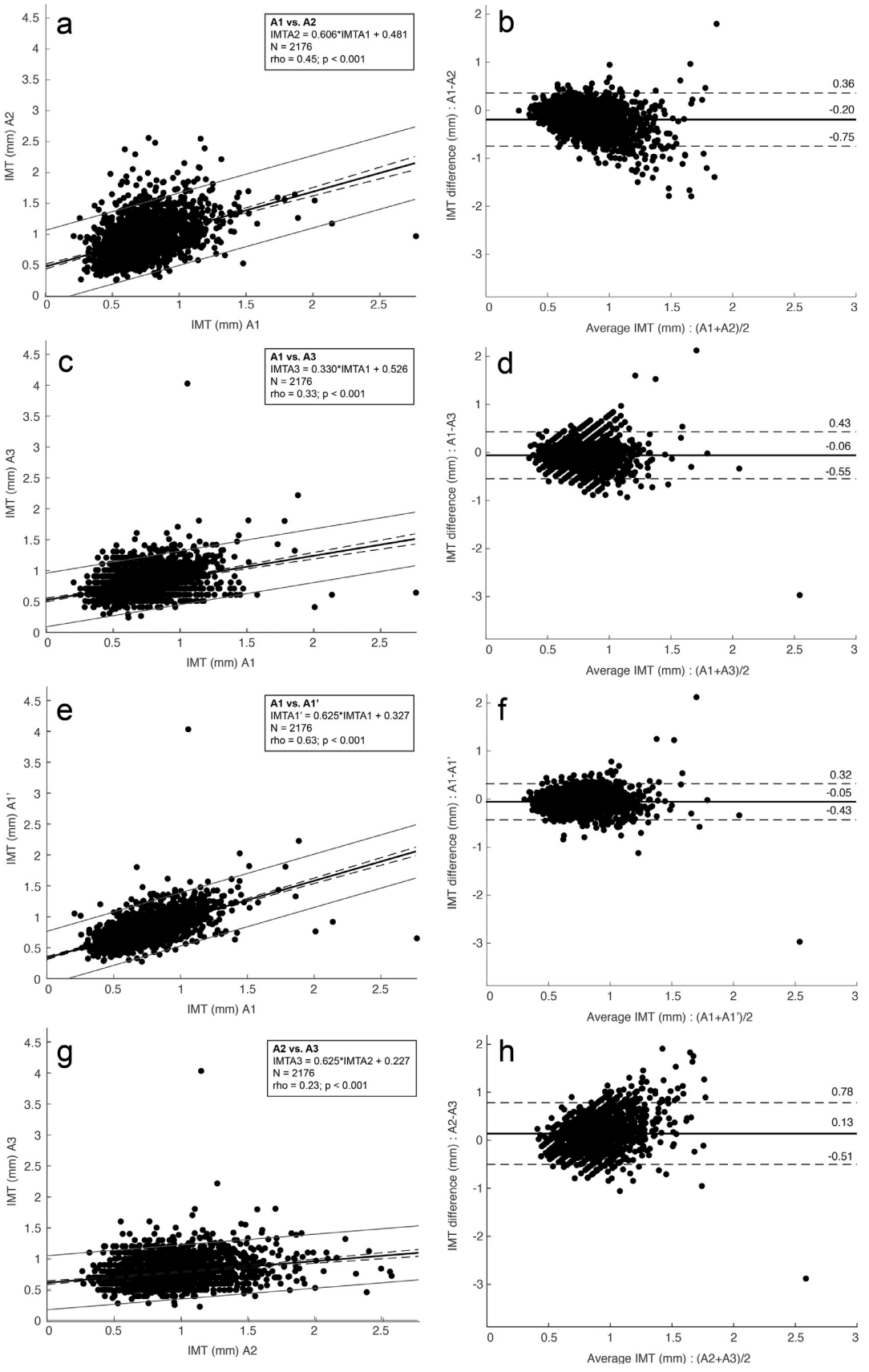
Table 2. CIMT values and mean signed difference (bias)

Segmentation method	No. of unprocessed images	CIMT value (μm)	$CIMT_{Bias}$ vs. A1 (μm)*	$CIMT_{AbsBias}$ vs. A1 (μm)
A1	0	725 ± 215	—	—
UCY_{CY}	0	769 ± 162	−44 (−474; 386)†	171 ± 144
TUM_{DE}	0	713 ± 170	12 (−307; 332)	114 ± 117
CNR_{IT}	10 (0.45%)	714 ± 159	10 (−325; 344)	120 ± 123
$INESCTEC_{PT}$	102 (4.69%)	849 ± 374	−125 (−734; 485)†	202 ± 271
$POLITO_{IT}$	0	834 ± 160	−109 (−497; 279)†	179 ± 138
A1′	0	780 ± 220	−55 (−431; 321)†	140 ± 143
A2	0	920 ± 299	−195 (−749; 359)†	255 ± 230
A3	0	780 ± 224	−57 (−552; 438)†	188 ± 180

A1, A2 and A3 = manual segmentations of analysts 1, 2 and 3, respectively; A1′ = segmentations of analyst 1, traced 1 month after A1; UCY_{CY} = computerized method from University of Cyprus; TUM_{DE} = method from Technische Universität München; CNR_{IT} = method from Consiglio Nazionale delle Ricerche; $INESCTEC_{PT}$ = method from INESC Technology and Science; $POLITO_{IT}$ = method from Politecnico di Torino.

* Ninety-five percent confidence interval in parentheses.

† Paired Wilcoxon test p value < 0.05.


a
A1 vs. A2
IMTA2 = 0.606*IMTA1 + 0.481
N = 2176
rho = 0.45; p < 0.001
IMT (mm) A2
IMT (mm) A1
b
IMT difference (mm) : A1-A2
Average IMT (mm) : (A1+A2)/2
0.36
-0.20
-0.75
c
A1 vs. A3
IMTA3 = 0.330*IMTA1 + 0.526
N = 2176
rho = 0.33; p < 0.001
IMT (mm) A3
IMT (mm) A1
d
IMT difference (mm) : A1-A3
Average IMT (mm) : (A1+A3)/2
0.43
-0.06
-0.55
e
A1 vs. A1'
IMTA1' = 0.625*IMTA1 + 0.327
N = 2176
rho = 0.63; p < 0.001
IMT (mm) A1'
IMT (mm) A1
f
IMT difference (mm) : A1-A1'
Average IMT (mm) : (A1+A1')/2
0.32
-0.05
-0.43
g
A2 vs. A3
IMTA3 = 0.625*IMTA2 + 0.227
N = 2176
rho = 0.23; p < 0.001
IMT (mm) A3
IMT (mm) A2
h
IMT difference (mm) : A2-A3
Average IMT (mm) : (A2+A3)/2
0.78
0.13
-0.51

because of either faulty LI and MA profile determination (CNR_{IT}, n = 10) or unsuccessful automatic lumen detection ($INESCTEC_{PT}$, n = 102). The CV for all nine methods employed (*i.e.,* four manual and five computerized CIMT measurements for each image) was 20 ± 10%. The CV considering only the four manual measurements was 21 ± 12%. Interestingly, for the five computerized CIMT measurements, the CV was 17 ± 11%.

To provide a brief outlier analysis, the 75% quartile of all absolute CIMT errors of the computerized methods compared with A1 was computed, and values above this threshold value were considered outliers. In 62 images (<3%), all of the computerized measurement methods produced outlier CIMT measurements. After visual inspection, the two main reasons were a hypo-echoic LI border or a noisy lumen close to the LI border. In a few cases, an overestimation/underestimation of the manual A1 CIMT measurement based on visual analysis proved to be faulty. Finally, the impossibility of completely respecting the Mannheim consensus guidelines (Touboul et al. 2012) (*i.e.,* curved/inclined artery, slight presence of plaque) was the reason for failure in a few cases. Interestingly, the manual measurement CV for the outlier images (29 ± 17%) was much higher than that for the non-outlier images (21 ± 11%), suggesting that the images were difficult to evaluate manually and that the discrepancy in the computerized technique's measurements is a feature of these images and not an overall defect.

Inter-observer/intra-observer variability

On the left side of Figure 2 (a, c, e and g) are the regression analysis results for inter/intra-observer variability. The Spearman correlation coefficient between A1 and A1′ was $\rho = 0.63$; it decreased to $\rho = 0.45$ between A1 and A2, to $\rho = 0.33$ between A1 and A3 and to $\rho = 0.23$ between A2 and A3. On the right side of Figure 2 are the Bland–Altman analyses, where the CIMT bias and 95% confidence interval can be appreciated.

The regression and Bland–Altman analysis of the computerized methods versus A1 CIMT measurements are provided in Figures 3 and 4, respectively.

Correlation results

Spearman correlation coefficients between the clinical parameters and mean CIMT values are listed in Supplementary Tables S1 and S2 (online only). The mean CIMT values computed with manual and computerized methods correlated significantly with all the clinical parameters, except low-density lipoproteins (LDL), for at least four methods; age was moderately positively correlated for all methods. Other clinical parameters exhibiting strong correlations with the CIMT measurements were sex, hypertension and creatinine levels.

Prediction results

Table 3 provides the area under the curve (AUC) results obtained using a stepwise logistic regression model to predict cardiovascular events. The mean CIMT value was chosen as a significant parameter in eight of the nine cases analyzed (88.9%). The right side of Table 3 depicts the clinical parameters and which parameters were selected or rejected. Three clinical parameters were always chosen: age, packs of cigarettes smoked per day × years of smoking (pack-years), and high-density lipoprotein. To further verify if there is an added predictive value of the CIMT measurements, the GLM analysis was also done without including the CIMT parameter and computing again the AUC. The results can be seen in the last row of Table 3.

In Figure 5 are the Kaplan–Meier cardiovascular event analysis plots. A net increase in cardiovascular event risk with CIMT group increases ($p < 0.05$) is illustrated for all methods with the exception of the computerized technique UCY_{CY}. The hazard ratio analysis results obtained by dividing the CIMT measurements into quartiles are reported in Table 4; the highest CIMT quartile was associated with a significantly higher hazard ratio in comparison to the lowest CIMT quartile for all methods, except the computerized technique UCY_{CY} and the manual measurement A3. The manual measurements A1 and A1′ and the computerized techniques TUM_{DE} and CNR_{IT} had statistically significant higher hazard ratios for all three quartiles in comparison to the lowest CIMT quartile.

DISCUSSION

In this study, five different computerized CIMT measurement methods and four manual segmentations were compared in terms of measurement variability and clinical significance.

First, this study indicated that variability is lower between computerized segmentations than skilled analysts' manual segmentations. The inter- and intra-observer reproducibility of manual CIMT measurements is

Fig. 2. Inter-observer/intra-observer analysis. Left: Correlation analyses (*black dotted lines* represent 95% prediction intervals). Right: Bland–Altman analyses (*black solid line* represents the mean difference; *black dotted lines* represent mean ± 1.96 × standard deviation): (a, b) A1 versus A2; (c,d) A1 versus A3; (e,f) A1 versus A1′; (g, h) A2 versus A3. A1, A2 and A3 = manual segmentations of analysts 1, 2 and 3, respectively; A1′ = segmentations of analyst 1, traced 1 month after A1. IMT = intima–media thickness.

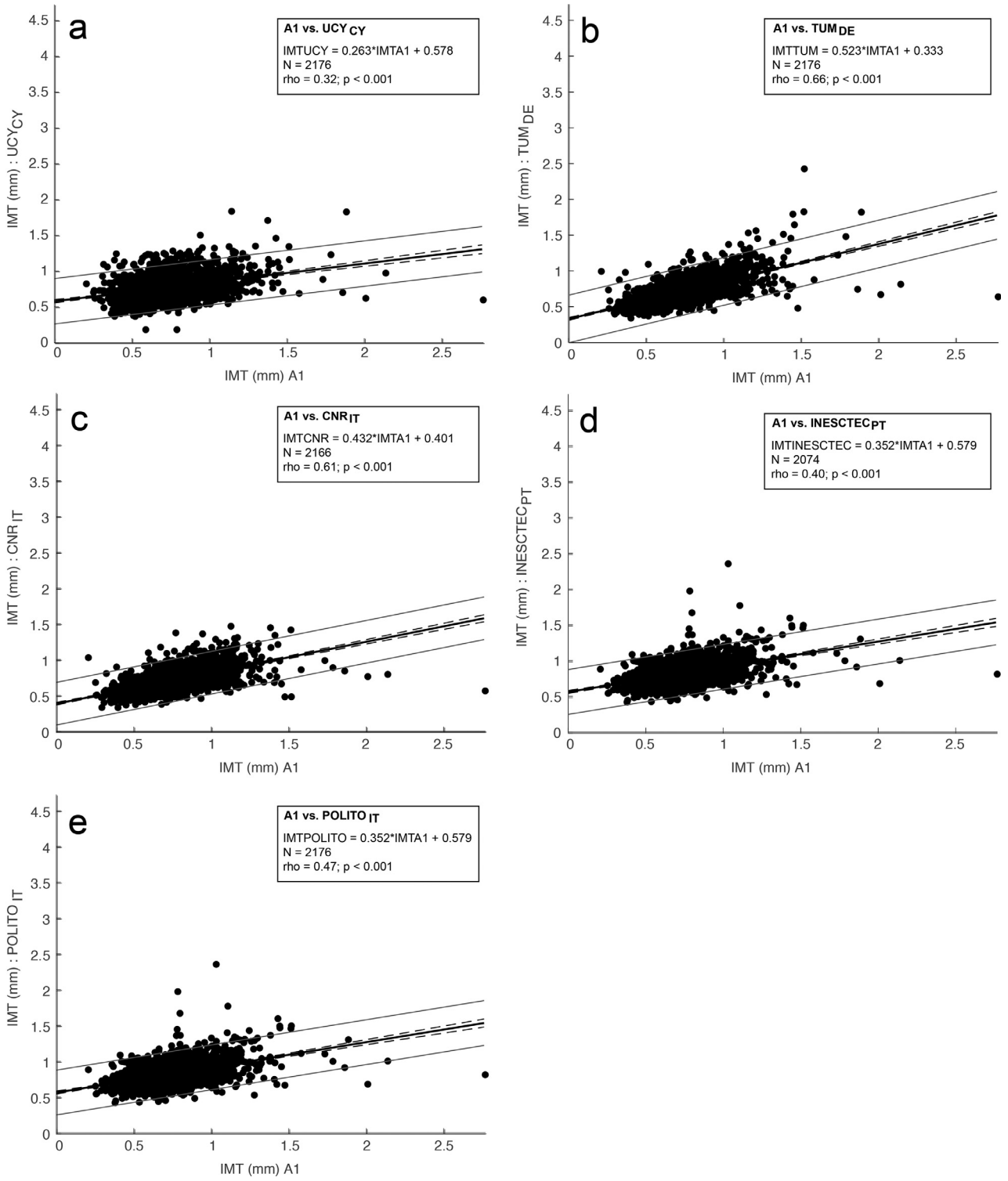


Fig. 3. Correlation analysis of computerized methods compared with expert analyst A1. *Black dotted lines* represent 95% prediction intervals. IMT = intima–media thickness; UCY_{CY} = computerized method from University of Cyprus; TUM_{DE} = computerized method from Technische Universität München (TUM) in Germany; CNR_{IT} = computerized method from Consiglio Nazionale delle Ricerche (CNR) in Pisa, Italy; $INESCTEC_{PT}$ = method from INESC Technology and Science (INESCTEC) in Porto, Portugal; $POLITO_{IT}$ = computerized method from Politecnico di Torino in Turin, Italy.

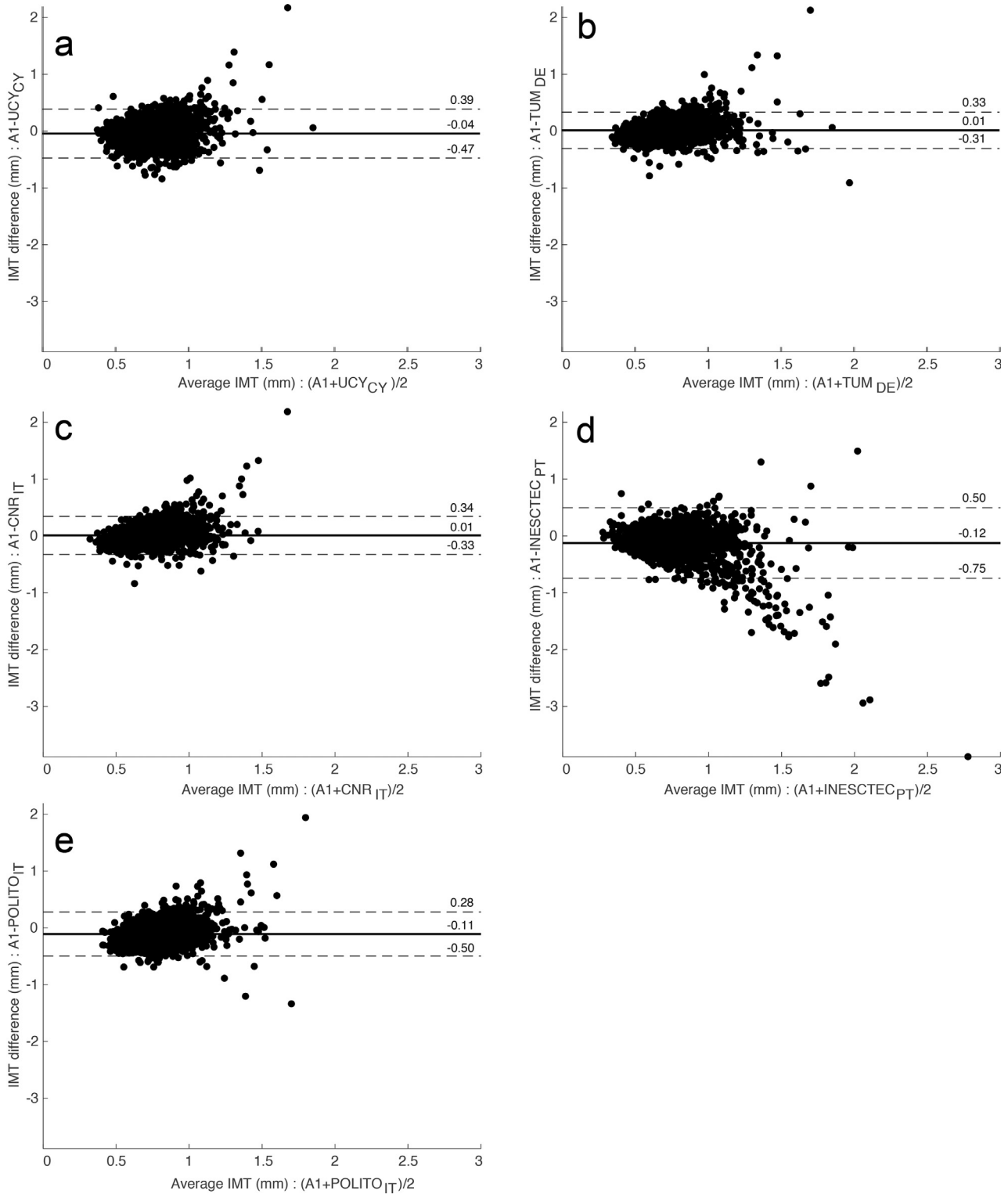


Fig. 4. Bland–Altman analysis of computerized methods compared with expert analyst A1. The *black solid line* represents the mean difference; *black dotted lines* represent the mean ± 1.96 × standard deviation. IMT = intima–media thickness; UCY_{CY} = computerized method from University of Cyprus; TUM_{DE} = computerized method from Technische Universität München (TUM) in Germany; CNR_{IT} = computerized method from Consiglio Nazionale delle Ricerche (CNR) in Pisa, Italy; $INESCTEC_{PT}$ = method from INESC Technology and Science (INESCTEC) in Porto, Portugal; $POLITO_{IT}$ = computerized method from Politecnico di Torino in Turin, Italy.

Table 3. Area under the curve values for prediction of cardiovascular events*

CIMT measurement	GLM training AUC	GLM Test AUC	Clinical parameters chosen in GLM: Age	Gender	CV symptoms	Smoking	Pack-years	Hypertension	Diabetes	Body mass index	Glucose	Total cholesterol	HDL	LDL	Triglycerides	Creatinine	Apolipoprotein A1	Apolipoprotein B	CIMT measurement
A1	0.806	0.728	■	■			■		■		■		■		■			■	■
UCY_{CY}	0.803	0.643	■		■		■				■		■	■		■			■
TUM_{DE}	0.803	0.704	■			■	■			■	■	■	■		■	■			■
CNR_{IT}	0.798	0.708	■				■		■		■	■	■			■			■
$INESCTEC_{PT}$	0.762	0.707	■		■		■						■			■		■	■
$POLITO_{IT}$	0.810	0.703	■			■	■			■	■		■	■	■	■			■
A1’	0.804	0.694	■			■	■		■				■	■	■	■			
A2	0.791	0.732	■	■			■		■				■		■			■	■
A3	0.790	0.738	■	■			■		■				■		■			■	■
No CIMT	0.792	0.665	■		■		■			■	■	■	■	■				■	■

CIMT = carotid intima–media thickness; GLM = generalized linear model; A1, A2 and A3 = manual segmentations of analysts 1, 2 and 3, respectively; A1′ = segmentations of analyst 1, traced 1 month after A1; UCY_{CY} = computerized method from the University of Cyprus; TUM_{DE} = method from Technische Universität München; CNR_{IT} = method from Consiglio Nazionale delle Ricerche; $INESCTEC_{PT}$ = method from INESC Technology and Science; $POLITO_{IT}$ = method from Politecnico di Torino. HDL = high-density lipoproteins; LDL = low-density lipoproteins.

* The right side of the table depicts which clinical parameters were chosen (*shaded squares*) in the final GLM model for each CIMT measurement. The *black square* depicts variables that were not included in the GLM model.

typically an obstacle, especially when considering large databases and cohort studies, as it depends on analyst effort and concentration. The variability of the computerized methods versus the gold standard (A1) was similar to that found during the inter-analyst and intra-analyst variability analyses. This finding suggests that the measurements obtained using computerized methods are as accurate as the measurements obtained using a gold standard skilled analyst’s manual segmentation. Indeed, the CV was higher between analysts (21% ± 12%) than between computerized techniques (17% ± 11%). Intra-observer variability of a completely automatic and fully deterministic method will be zero, whereas a semi-automatic method may exhibit some variability because of a different manual initialization, but previous studies reported that this variability is negligible (Saba et al. 2018). Moreover, the segmentation errors found in this study are comparable to those of other computerized techniques that have been employed on other databases (Molinari et al. 2012a; Biswas et al. 2018). More specifically, in Biswas et al. (2018) a CIMT error equal to 126 ± 134 μm was obtained with a deep learning technique on 396 images, whereas in Molinari et al. (2012a), an absolute CIMT error was equal to at best 150 ± 169 μm and at worst 224 ± 252 μm when comparing five computerized techniques on a database of 665 images.

Interestingly, a rather high value of statistically significant differences were found between not only between the CIMT values obtained by the computerized methods and the A1 manual measurements, but also between the manual measurements (*i.e.,* A1 vs. A2, A1 vs. A1′). This again attests to the high variability found between the observers in this study. Two of the five computerized methods (TUM_{DE} and CNR_{IT}) provided CIMT measurements that were not significantly different from the A1 manual measurements, when considering the paired Wilcoxon test. This can also be observed in Figure 3, where these two methods exhibit a higher correlation coefficient, and in Figure 4, where the mean error is equal to 0.1 mm for both of these methods. While exhibiting a statistically significant difference when computing the paired Wilcoxon test, the other methods still manifest a statistically significant correlation with the A1 manual CIMT measurements. As the Bland–Altman plots in Figure 4 indicate, the other three methods (*i.e.,* UCY_{CY}, $INESCTEC_{PT}$, $POLITO_{IT}$) do not indicate a specific trend, but present an overall larger bias as they all tend to overestimate the CIMT value. This can also be observed in Figure 5, where these three computerized methods indicate a larger number of subjects belonging to IMT group 4 (CIMT >850 μm).

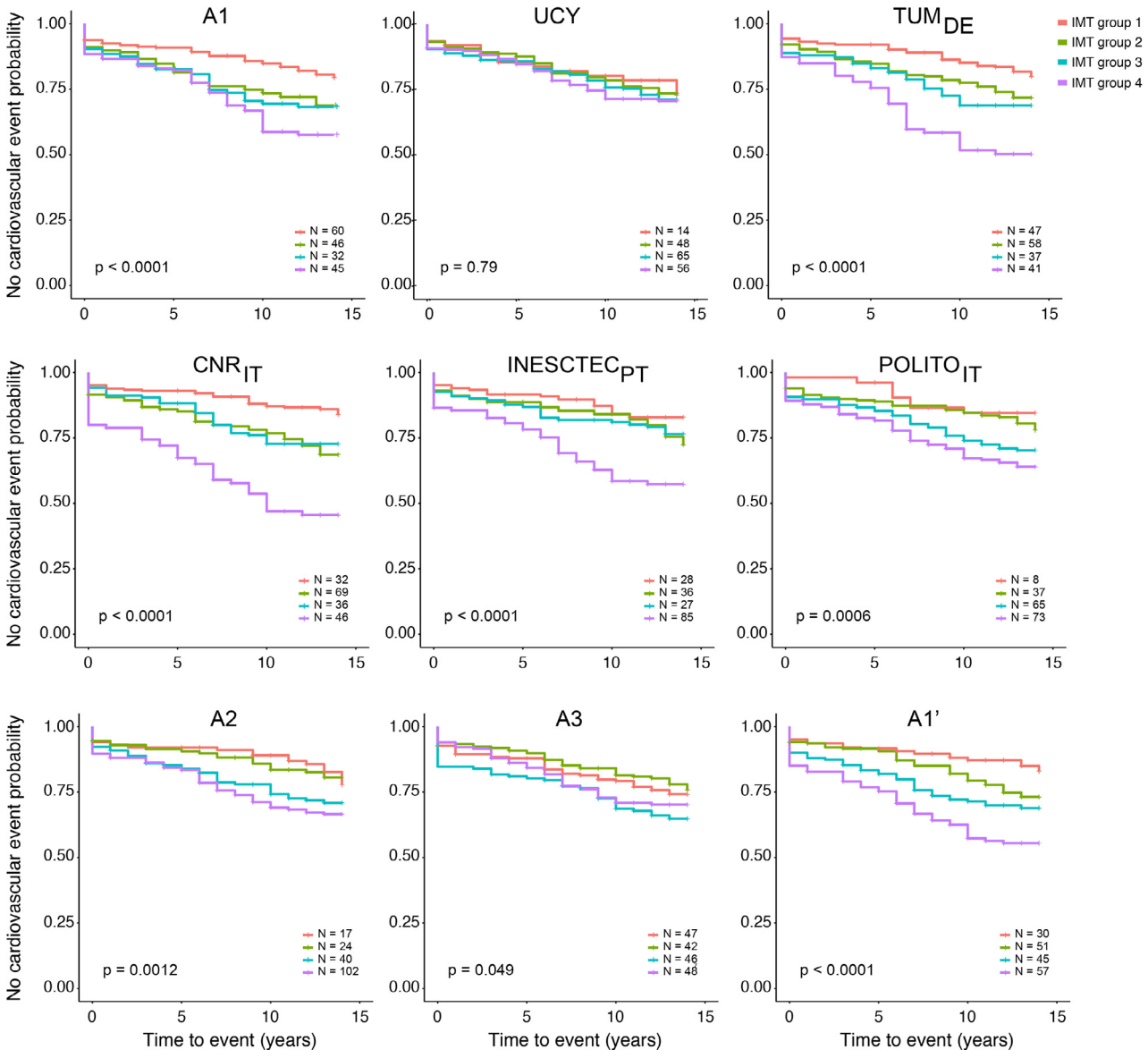


Fig. 5. Kaplan–Meier cardiovascular event analysis plots, considering four CIMT groups: <650 μm, between 650 and 750 μm, between 750 and 850 μm and >850 μm. A1, A2 and A3 = manual segmentations of analysts 1, 2 and 3, respectively. A1′ = segmentations of analyst 1, traced 1 month after A1; UCY_{CY} = computerized method from University of Cyprus; TUM_{DE} = method from Technische Universität München; CNR_{IT} = method from Consiglio Nazionale delle Ricerche; $INESCTEC_{PT}$ = method from INESC Technology and Science; $POLITO_{IT}$ = method from Politecnico di Torino.

Beyond the accuracy of contour segmentation and CIMT quantification, it is crucial to determine the reliability of clinical information that can be inferred from either manual tracings or computerized measurements. The results here demonstrate that:

- Both manual and computerized CIMT measurements exhibited a similar correlation with clinical parameters (Supplementary Tables S1 and S2, online only).
- All approaches yielded a comparable parameter set when constructing a GLM (Table 3).
- The hazard ratio analysis and Kaplan–Meier cardiovascular event analysis plots revealed overall comparable results between computerized and manual CIMT measurements, with one computerized measurement (UCY_{CY}) and one manual measurement (A3) exhibiting slightly lower performance (Table 4 and Fig. 5).

Table 4. Hazard ratio results for cardiovascular events (N = 183).

Hazard group		A1	UCY_{CY}	TUM_{DE}	CNR_{IT}	$INESCTEC_{PT}$	$POLITO_{IT}$	A1′	A2	A3
Group 1	% Events	0.33	0.08	0.26	0.17	0.15	0.04	0.16	0.09	0.26
IMT <650 μm	H ratio	1	1	1	1	1	1	1	1	1
	95% CI	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	p Value	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	$CIMT_{AbsBias}$ (μm)	0 ± 0	209 ± 110	78 ± 64	83 ± 65	138 ± 112	187 ± 94	108 ± 87	217 ± 162	185 ± 151
Group 2	% Events	0.25	0.26	0.32	0.38	0.20	0.20	0.28	0.13	0.23
IMT ≥650 μm	H ratio	1.7	1.1	1.6	2.3	1.4	1.3	1.8	1.1	0.84
IMT <750 μm	95% CI	1.2−2.5	0.6−2.0	1.1−2.4	1.5−3.5	0.86−2.3	0.6−2.8	1.1−2.8	0.6−2.1	0.55−1.3
	p Value	0.006	0.769	0.014	<0.001	0.179	0.514	0.010	0.736	0.402
	$CIMT_{AbsBias}$ (μm)	0 ± 0	135 ± 112	70 ± 58	65 ± 53	141 ± 191	123 ± 91	92 ± 71	202 ± 170	143 ± 136
Group 3	% Events	0.17	0.36	0.20	0.20	0.15	0.36	0.25	0.22	0.25
IMT ≥750 μm	H ratio	1.8	1.2	1.9	2.1	1.4	2.1	2.3	1.8	1.47
IMT <850 μm	95% CI	1.2−2.8	0.68−2.2	1.2−2.9	1.3−3.3	0.81−2.3	1.0−4.4	1.5−3.7	1.0−3.2	0.98−2.2
	p Value	0.005	0.502	0.003	<0.001	0.233	0.490	<0.001	0.039	0.063
	$CIMT_{AbsBias}$ (μm)	0 ± 0	110 ± 105	71 ± 52	69 ± 50	111 ± 154	91 ± 71	105 ± 79	169 ± 131	149 ± 105
Group 4	% Events	0.25	0.31	0.22	0.25	0.46	0.40	0.31	0.56	0.26
IMT ≥850 μm	H ratio	2.5	1.3	3.4	5.1	3.0	2.6	3.7	2.2	1.21
	95% CI	1.7−3.7	0.71−2.3	2.2−5.2	3.3−8.1	1.93−4.5	1.3−5.5	2.4−5.7	1.3−3.7	0.81−1.8
	p Value	<0.001	0.413	<0.001	<0.001	<0.001	0.009	<0.001	0.003	0.346
	$CIMT_{AbsBias}$ (μm)	0 ± 0	202 ± 391	173 ± 402	183 ± 399	186 ± 408	166 ± 387	170 ± 386	214 ± 396	268 ± 398

A1, A2 and A3 = manual segmentations of analysts 1, 2 and 3, respectively; A1′ = segmentations of analyst 1, traced 1 month after A1; UCY_{CY} = computerized method from University of Cyprus; TUM_{DE} = method from Technische Universität München; CNR_{IT} = method from Consiglio Nazionale delle Ricerche; $INESCTEC_{PT}$ = method from INESC Technology and Science; $POLITO_{IT}$ = method from Politecnico di Torino; IMT = intima−media thickness; H ratio = Hazard ratio; CI = confidence interval.

With respect to the GLM analysis, it can be observed that when excluding the CIMT variable, the test set AUC is equal to 0.665, which is overall lower compared with the value when the GLM analysis is done using the computerized and manual CIMT values. Hence, we can infer that the CIMT provides added predictive value for cardiovascular events, but considering the data set presented here in this study, the added predictive value is not significant. According to the Kaplan−Meier cardiovascular event analysis, a slight difference can be observed with respect to the TUM_{DE} and CNR_{IT} methods, as they exhibit a less than 50% event-free probability by year 10. This could be owing to the strict IMT group limits used for all methods and the inclusion of a larger number of subjects that did have an event throughout the study time in IMT group 4 (CIMT >850 μm) for these two methods. Still, the majority of the other methods exhibit a similar trend, albeit less noticeable, and also still generally present low *p* values, attesting that this observation cannot really differentiate these two methods from the others (excluding UCY_{CY} and A3) statistically.

Finally, it is noted that no individual computerized method investigated in this study was found to always perform substantially better or worse than the others. The TUM_{DE} and CNR_{IT} methods exhibited promising results in particular when considering the nominal CIMT value by itself, as there was no statistically significant difference with A1 (paired Wilcoxon test). The other methods also exhibited promising results as none tended to systematically underestimate CIMT, which is important when screening for cardiovascular disease. Moreover, all methods exhibited high linear correlations with the expert manual segmentations, which is significant when considering the relevance of following up patients and observing not only a one-time CIMT value but also its evolution.

CONCLUSIONS

In this study, five different computerized CIMT measurement methods were compared on the same large database. The comparison was done both in terms of correlation with clinical parameters that are used to assess cardiovascular risk and in terms of similarities or differences in their ability to predict cardiovascular events.

The main concluding finding of this study is the fact that computerized methods can be used instead of a skilled analyst's manual segmentation for CCA segmentation, CIMT quantification and clinical outcome investigation. Furthermore, it was found that the two approaches (*i.e.,* manual and computerized) performed similarly, while computerized methods have the advantages of a favorable time efficiency, full reproducibility and easy standardization.

The full data set involved in this study (1088 participants and 2176 images, clinical information, manual reference tracings and segmentation contours from all compared methods) has been made publicly available for the community. It is the authors' intention to

facilitate future studies by different groups—potentially also in the field of deep learning—by providing a large annotated data set.

Acknowledgments—We are grateful to the CCDER Trustees for permission to use the images from the Cyprus Epidemiological Study.

Conflict of interest disclosure—F.F. and E.B. are both co-founders of Quipu Srl, Pisa, Italy, a spin-off company of the Italian National Research Council and the University of Pisa. L.G. reports other income from Quipu srl, outside the submitted work. C.C., A.C. and J.R. report ECHOCAD CAROTID (IGAC Software Registration 185/2918).

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ultrasmedbio.2021.03.022.

REFERENCES

Bianchini E, Giannarelli C, Maria Bruno R, Armenia S, Landini L, Faita F, Gemignani V, Taddei S, Ghiadoni L. Functional and structural alterations of large arteries: Methodological issues. Curr Pharm Des 2013;19:2390–2400.

Biswas M, Kuppili V, Araki T, Edla DR, Godia EC, Saba L, Suri HS, Omerzu T, Laird JR, Khanna NN, Nicolaides A, Suri JS. Deep learning strategy for accurate carotid intima-media thickness measurement: An ultrasound study on Japanese diabetic cohort. Comput Biol Med 2018;98:100–117.

Bots ML, Groenewegen KA, Anderson TJ, Britton AR, Dekker JM, Engström G, Evans GW, de Graaf J, Grobbee DE, Hedblad B, Hofman A, Holewijn S, Ikeda A, Kavousi M, Kitagawa K, Kitamura A, Ikram MA, Lonn EM, Lorenz MW, Mathiesen EB, Nijpels G, Okazaki S, O'Leary DH, Polak JF, Price JF, Robertson C, Rembold CM, Rosvall M, Rundek T, Salonen JT, Sitzer M, Stehouwer CDA, Franco OH, Peters SAE, den Ruijter HM. Common carotid intima-media thickness measurements do not improve cardiovascular risk prediction in individuals with elevated blood pressure. Hypertension 2014;63:1173–1181.

Bruno RM, Cartoni G, Stea F, Armenia S, Bianchini E, Buralli S, Giannarelli C, Taddei S, Ghiadoni L. Carotid and aortic stiffness in essential hypertension and their relation with target organ damage: The CATOD study. J Hypertens 2017;35:310–318.

Bruno RM, Stea F, Sicari R, Ghiadoni L, Taddei S, Ungar A, Bonuccelli U, Tognoni G, Cintoli S, Del Turco S, Sbrana S, Gargani L, D'Angelo G, Pratali L, Berardi N, Maffei L, Picano E, Andreassi MG, Angelucci A, Baldacci F, Baroncelli L, Begenisic T, Bellinvia PF, Biagi L, Bonaccorsi J, Bonanni E, Borghini A, Braschi C, Broccardi M, Caleo M, Carlesi C, Carnicelli L, Cartoni G, Cecchetti L, Cenni MC, Ceravolo R, Chico L, Cioni G, Costa M, D'Ascanio P, De Nes M, Di Coscio E, Di Galante M, di Lascio N, Faita F, Falorni I, Faraguna U, Fenu A, Fortunato L, Franco R, Gargiulo R, Giorgi FS, Iannarella R, Iofrida C, Kusmic C, Limongi F, Maestri M, Maffei M, Maggi S, Mainardi M, Mammana L, Marabotti A, Mariotti V, Melissari E, Mercuri A, Molinaro S, Narducci R, Navarra T, Noale M, Pagni C, Palumbo S, Pasquariello R, Pellegrini S, Pietrini P, Pizzorusso T, Poli A, Retico A, Ricciardi E, Rota G, Sale A, Scabia G, Scali M, Scelfo D, Siciliano G, Tonacci A, Tosetti M, Turchi S, Volpi L. Vascular function is improved after an environmental enrichment program. Hypertension 2018;71:1218–1225.

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–357.

Engelen L, Ferreira I, Stehouwer CD, Boutouyrie P, Laurent S. Reference intervals for common carotid intima-medi thickness measured with echotracking: Relation with risk factors. Eur Heart J 2013;34:2368–2380.

Faita F, Gemignani V, Bianchini E, Giannarelli C, Ghiadoni L, Demi M. Real-time measurement system for evaluation of the carotid intima-media thickness with a robust edge operator. J Ultrasound Med 2008;27:1353–1361.

Griffin M, Nicolaides A, Tyllis T, Georgiou N, Martin RM, Bond D, Panayiotou A, Tziakouri C, Doré CJ, Fessas C. Carotid and femoral arterial wall changes and the prevalence of clinical cardiovascular disease. Vasc Med 2009;14:227–232.

Hassan M, Chaudhry A, Khan A, Iftikhar MA. Robust information gain based fuzzy *c*-means clustering and classification of carotid artery ultrasound images. Comput Methods Programs Biomed 2014;113:593–609.

Ilea DE, Duffy C, Kavanagh L, Stanton A, Whelan PF. Fully automated segmentation and tracking of the intima media thickness in ultrasound video sequences of the common carotid artery. IEEE Trans Ultrason Ferroelectr Freq Control 2013;60:158–177.

Loizou CP. A review of ultrasound common carotid artery image and video segmentation techniques. Med Biol Eng Comput 2014;52:1073–1093.

Loizou CP, Pattichis CS, Pantziaris M, Tyllis T, Nicolaides A. Snakes based segmentation of the common carotid artery intima media. Med Biol Eng Comput 2007;45:35–49.

Lorenz MW, Gao L, Ziegelbauer K, Norata GD, Empana JP, Schmidtmann I, Lin HJ, McLachlan S, Bokemark L, Ronkainen K, Amato M, Schminke U, Srinivasan SR, Lind L, Okazaki S, Stehouwer CDA, Willeit P, Polak JF, Steinmetz H, Sander D, Poppert H, Desvarieux M, Ikram MA, Johnsen SH, Staub D, Sirtori CR, Iglseder B, Beloqui O, Engström G, Friera A, Rozza F, Xie W, Parraga G, Grigore L, Plichart M, Blankenberg S, Su TC, Schmidt C, Tuomainen TP, Veglia F, Völzke H, Nijpels G, Willeit J, Sacco RL, Franco OH, Uthoff H, Hedblad B, Suarez C, Izzo R, Zhao D, Wannarong T, Catapano A, Ducimetiere P, Espinola-Klein C, Chien KL, Price JF, Bergström G, Kauhanen J, Tremoli E, Dörr M, Berenson G, Kitagawa K, Dekker JM, Kiechl S, Sitzer M, Bickel H, Rundek T, Hofman A, Mathiesen EB, Castelnuovo S, Landecho MF, Rosvall M, Gabriel R, de Luca N, Liu J, Baldassarre D, Kavousi M, de Groot E, Bots ML, Yanez DN, Thompson SG, PROG−IMT study group. Predictive value for cardiovascular events of common carotid intima media thickness and its rate of change in individuals at high cardiovascular risk—Results from the PROG−IMT collaboration. PLoS One 2018;13 e0191172.

Meiburger KM, Acharya UR, Molinari F. Automated localization and segmentation techniques for B-mode ultrasound images: A review. Comput Biol Med 2018;92:210–235.

Molinari F, Meiburger KM, Saba L, Acharya UR, Ledda G, Zeng G, Ho SYS, Ahuja AT, Ho SC, Nicolaides A, Suri JS. Ultrasound IMT measurement on a multi-ethnic and multi-institutional database: Our review and experience using four fully automated and one semi-automated methods. Comput Methods Programs Biomed 2012a;108:946–960.

Molinari F, Meiburger KM, Saba L, Zeng G, Acharya UR, Ledda M, Nicolaides A, Suri JS. Fully automated dual-snake formulation for carotid intima-media thickness measurement: A new approach. J Ultrasound Med 2012b;31:1123–1136.

Oren A, Vos LE, Uiterwaal CSPM, Grobbee DE, Bots ML. Cardiovascular risk factors and increased carotid intima-media thickness in healthy young adults. Arch Intern Med 2003;163:1787.

Plichart M, Celermajer DS, Zureik M, Helmer C, Jouven X, Ritchie K, Tzourio C, Ducimetière P, Empana JP. Carotid intima-media thickness in plaque-free site, carotid plaques and coronary heart disease risk prediction in older adults: The Three-City Study. Atherosclerosis 2011;219:917–924.

Potter K, Reed CJ, Green DJ, Hankey GJ, Arnolda LF. Ultrasound settings significantly alter arterial lumen and wall thickness measurements. Cardiovasc Ultrasound 2008;6:6.

Rocha R, Silva J, Campilho A. Automatic segmentation of carotid B-mode images using fuzzy classification. Med Biol Eng Comput 2012;50:533–545.

Rouco J, Carvalho C, Domingues A, Azevedo E, Campilho A. A robust anisotropic edge detection method for carotid ultrasound image processing. Procedia Comput Sci 2018;126:723–732.

Saba L, Banchhor S, Araki T, Suri H, Londhe N, Laird J, Viskovic K, Suri J. Intra- and inter-operator reproducibility of automated cloud-

based carotid lumen diameter ultrasound measurement. Indian Heart J 2018;70:649–664.

Stein JH, Korcarz CE, Hurst RT, Lonn E, Kendall CB, Mohler ER, Najjar SS, Rembold CM, Post WS. American Society of Echocardiography Carotid Intima-Media Thickness Task Force. Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: A Consensus Statement from the American Society of Echocardiography Carotid Intima-Media Thickness Task Force Endorsed by the Society for Vascular Medicine. J Am Soc Echocardiogr 2008;21:93–111 quiz 189−190.

Touboul P, Hennerici M, Meairs S, Adams H. Mannheim carotid intima-media thickness and plaque consensus (2004−2006−2011). Cerebrovasc Dis 2012;34:290–296.

Zahnd G, Kapellas K, van Hattem M, van Dijk A, Sérusclat A, Moulin P, van der Lugt A, Skilton M, Orkisz M. A fully-automatic method to segment the carotid artery layers in ultrasound imaging: Application to quantify the compression-decompression pattern of the intima-media complex during the cardiac cycle. Ultrasound Med Biol 2017;43:239–257.

Zahnd G, Orkisz M, Dávila Serrano EE, Vray D. CAROLAB—A platform to analyze carotid ultrasound data. 2019 IEEE International Ultrasonics Symposium (IUS), Glasgow, UK. Piscataway, NJ. : IEEE; 2019. p. 463–466.

Zhou R, Fenster A, Xia Y, Spence JD, Ding M. Deep learning-based carotid media-adventitia and lumen-intima boundary segmentation from three-dimensional ultrasound images. Med Phys 2019;46:3180–3193.