Title: **Effect of High-Intensity Statin Therapy on Atherosclerosis (IBIS-4): Manual Versus Automated Methods of IVUS Analysis**

Authors: Ronald D. Bass BA[a*]; Héctor M. García-García MD, PhD[b*]; Yasushi Ueki, MD[c]; Lene Holmvang, MD[d]; Giovanni Pedrazzini, MD[e]; Marco Roffi, MD[f]; Konstantinos C Koskinas, MD[f], Hiroki Shibutani MD, PhD[c], Sylvain Losdat BSc[g]; Paulo G.P. Ziemer, PhD[h], Pablo J. Blanco PhD[h]; Molly B. Levine, BSc[b]; Christos V. Bourantas MD, PhD[i,j]; Lorenz Räber MD, PhD[c]

[a] School of Medicine, Georgetown University, Washington DC

[b] Interventional Cardiology, MedStar Washington Hospital Center, Washington, DC

[c] Department of Cardiology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[d] Cardiac Catheterization Laboratory, Rigshospitalet, Copenhagen, Denmark

[e] Cardiocentro, Lugano, Switzerland

[f] Division of Cardiology, University Hospital, Geneva, Switzerland

[g] CTU, Bern, Switzerland

[h] National Laboratory for Scientific Computing, Petrópolis, Brazil

[i] Department of Cardiology, Barts Heart Centre, Barts Health NHS Trust, London, UK

[j] Centre for Cardiovascular Medicine and Devices, William Harvey Research Institute, Queen Mary University of London, UK

* These authors contributed equally

Word Count: 3624

Corresponding Author:

Prof. Héctor M. García-García MD, PhD

Division of Interventional Cardiology of MedStar Cardiovascular Research Network at MedStar Washington Hospital Center

110 Irving Street, Suite 4B-1, Washington, DC, 20010, USA

Phone: (202) 877-2812

Fax: (202) 877-0859

Email: hector.m.garciagarcia@medstar.net; hect2701@gmail.com

Author Email Addresses:

rb1631@georgetown.edu; Hector.M.GarciaGarcia@medstar.net; yasushi522@shinshu-u.ac.jp;lene.holmvang@regionh.dk;Giovanni.Pedrazzini@eoc.ch;Marco.Roffi@hcuge.ch; Konstantinos.Koskinas@insel.ch; Hiroki.shibutani@insel.ch; sylvain.losdat@ctu.unibe.ch;pgpz23@gmail.com;pablo.j.blanco@gmail.com;Molly.B.Levine@medstar.net; cbourantas@gmail.com; Lorenz.Raeber@insel.ch

Abbreviations:

CAD - Coronary Artery Disease

CL - Core Laboratory

CNN - Convolutional Neural Network

GT - Ground Truth

IBIS-4 - Integrated biomarker imaging study

IRA - Infarct-related arteries

IVUS - Intravascular Ultrasound

MFCNN - Multi-frame Convolutional Neural Network

ML - Machine Learning

PAV - Percent Atheroma Volume

PCI - Percutaneous Coronary Intervention

ROI - Region of Interest

Abstract:

Aims: Standard manual analysis of IVUS to study the impact of anti-atherosclerotic therapies on the coronary vessel wall is done by a Core Laboratory (CL), the ground truth (GT). Automatic segmentation of IVUS with a machine learning (ML) algorithm has the potential to replace manual readings with an unbiased and reproducible method. The aim is to determine if results from a CL can be replicated with ML methods.

Methods: This is a post-hoc, comparative analysis of the IBIS-4 (Integrated Biomarkers and Imaging Study-4) study (NCT00962416). The GT baseline and 13-month follow-up measurements of lumen and vessel area and percent atheroma volume (PAV) after statin induction were repeated by the ML algorithm.

Results: The primary endpoint was change in PAV. PAV as measured by GT was 43.95% at baseline and 43.02% at follow-up with a change of -0.90% (p=0.007) while the ML algorithm measured 43.69% and 42.41% for baseline and follow-up, respectively, with a change of -1.28% (p<0.001). Along the most diseased 10mm segments, GT-PAV was 52.31% at baseline and 49.42% at follow-up, with a change of -2.94% (p<0.001). The same segments measured by the ML algorithm resulted in PAV of 51.55% at baseline and 47.81% at follow-up with a change of -3.74% (p<0.001).

Conclusions: PAV, the most used endpoint in clinical trials, analyzed by the CL is closely replicated by the ML algorithm. ML automatic segmentation of lumen, vessel and plaque effectively reproduces GT and may be used in future clinical trials as the standard.

Keywords: Intravascular Ultrasound, Coronary Artery Disease, Machine Learning, Lumen Segmentation, Vessel Segmentation

Introduction:

Acute coronary syndromes (ACS), a symptomatic subcategory of coronary artery disease (CAD), remains a major cause of morbidity and mortality around the world.[1] Treatment and secondary prevention has focused on reducing clinically significant adverse events with statins and other medical therapy. Several mechanistic intravascular imaging studies have demonstrated the effects of statin use with plaque-level analyses. Specifically, high-intensity rosuvastatin therapy has been shown to provide the most benefit.[2]

Intravascular ultrasound (IVUS) has become the most established technique to determine the effect of anti-atherosclerotic therapies on the vessel wall as it allows a precise and reproducible assessment of plaque burden.[3,4,5] Typically, IVUS image analysis in clinical trials is conducted by an independent Core Laboratory (CL) that is regarded as the reference standard for IVUS segmentation. This is currently performed manually by tracing frame by frame the lumen and vessel areas. Manual segmentation of IVUS images is a time-consuming analysis plagued by human variability with associated high costs. Recent studies have shown that newer methods of automatic segmentation with machine learning (ML) algorithms improve reproducibility, accuracy, and precision.[6,7,8] Few studies have compared results of ML algorithms to the CL.

The IBIS-4 (Integrated Biomarkers and Imaging Study-4) (NCT00962416) study was a serial IVUS study performed in STEMI patients receiving high-intensity statin therapy.[9] The objective of this study was to explore whether the results of the IBIS-4 serial IVUS study could be reproduced with a previously validated fully automated ML algorithm and were comparable to the gold standard CL results.

Materials and Methods:

Study Design and Patient Population

This is a post-hoc, comparative analysis of the IBIS-4 study.[9] The design of this study is represented in **Figure 1**. The IBIS-4 was a prospective cohort study that evaluated whether high-intensity statin therapy can reduce disease burden in proximal segments of non-infarct-related arteries (IRA) within 13 months of use after successful primary percutaneous coronary intervention (PCI). IVUS determined the effect of statin-use on plaque burden and phenotype at baseline and follow-up. Patients enrolled in the IBIS-4 study met the following criteria: inclusion in the COMFORTABLE-AMI trial plus age <90 years, preserved renal and liver function, hemodynamic stability, TIMI flow $\geq$2 of the IRA at completion of primary PCI, and suitable coronary anatomy for intravascular imaging. Rosuvastatin treatment was initially started at 20 mg once daily for the first 2 weeks and subsequently increased to 40 mg.

All patients in the IBIS-4 were provided written informed consent, and the study was approved by the institutional review boards of all participating centers.

Procedures: Image Acquisition and Analysis

Intracoronary imaging was first performed at baseline of the region of interest (ROI) which included the proximal 50mm of two non-IRA. At follow-up, 13 months after initiation of rosuvastatin, IVUS imaging was repeated of the same ROI. A 20-MHz IVUS catheter (Eagle Eye, Volcano Corporation, Rancho Cordova, CA) was used with a pullback speed of 0.5 mm/s after administration of 200 μg of intracoronary nitroglycerine.[9] For ground

truth (GT), image recordings were sent to an independent CL (Cardialysis B.V., Rotterdam, The Netherlands).[9]

The original report included the analysis of the radiofrequency frames which are a subset of the grayscale complete dataset **(Figure 2).** Both the GT baseline and 13-month follow-up measurements of region length, average lumen area, average vessel area and percent atheroma volume (PAV), or plaque burden, were repeated by the ML algorithm as depicted in **Figure 1**. The ML algorithm was applied to the blinded IVUS datasets which were not labeled baseline or follow-up but instead were labeled with a coding only known by the statistician of the study.

Machine Learning Algorithm

The current study focuses on replicating the ground truth IVUS results with an automated segmentation method. The artificial intelligence software used in this study was proposed by Blanco et al.[6] The ML approach automatically extracts lumen and vessel boundaries from IVUS datasets, relying on a multi-frame convolutional neural network (MFCNN). The MFCNN is coupled to a Gaussian process regressor that delivers consistent contour representation (**Figure 2).** Training, validation, and testing for the ML algorithm was all based on the ground truth. See the *supplementary material* for a more detailed account of the ML algorithm.

Statistical Analysis

The statistician received the results from the ML algorithm which were blinded for the analysis phase (baseline/follow-up). The IBIS-4 primary endpoint was change in the

PAV and the secondary efficacy endpoint was normalized atheroma volume. PAV was calculated as lumen area subtracted from vessel area, divided by vessel area, multiplied by 100. For more detailed definitions of these 2 parameters, please refer to the original publication.

Only the statistician knew the key for deciphering the analysis phase of the IVUS results.

Baseline categorical variables are presented as absolute values and percentages and were compared between groups using Fisher's exact tests. Baseline continuous variables are expressed as mean ± standard deviation or median (lower quartile-upper quartile) and were analyzed using Student's t-tests or Kruskall-Wallis tests, as appropriate. Serial IVUS analyses were conducted using mixed-effect models on the difference between follow-up and baseline, including patient identity as random intercept to correct for the multiple vessels per patient. P-values were two-tailed and the significance level was set to 0.05 in all analyses. Statistical analyses were performed using R Studio version 1.2.5033 (RStudio, PBC, Boston, MA).

Results:

Baseline and demographic characteristics of patients are found in **Table 1**. A total of 103 patients (207 vessels) were originally eligible for IBIS-4 imaging. Serial imaging results for both the GT and ML algorithm analysis was available for 82 patients with 146 vessels. Serial IVUS imaging was not available in 21 patients and thus was not eligible for this analysis. Of the 82 patients included in the current study, 92.7% were male, 47.6% had hypertension, 43.9% had hypercholesterolemia, 2.4% had a previous myocardial infarction and 1.2% had a previous PCI. The baseline demographics were similar between the patients included and those that were excluded in the analysis.

Parameter measurements from imaging of the entire ROI and the most diseased 10 mm segment were reported as analyzed by GT and the ML algorithm, summarized in **Table 2**. Region length was determined by GT to be 36.15 mm at baseline and 36.21 mm at follow-up.

For the entire region, change in average lumen area between baseline and follow-up as measured by GT was -0.16 mm^2 while change in average lumen area measured by ML was -0.17 mm^2. The change in average lumen area was statistically significant for both GT and ML with a p-value of 0.020 and 0.009, respectively. The change in average vessel area measured by GT was -0.51 mm^2. The same parameter as measured by the ML algorithm was -0.62 mm^2. The regression in vessel area was statistically significant with a p-value of <0.001 in both cases.

The primary endpoint in IBIS-4 was the change in the PAV. Change in PAV as calculated by GT was -0.90% (p=0.007), while change in PAV calculated by the ML algorithm was -1.28% (p<0.001), summarized in **Figure 1**. The secondary efficacy

endpoint was change in normalized total atheroma volume (mm^3): for GT was -12.18 (p<0.001) and for ML was -15.08 (p<0.001).

For the most diseased 10 mm segments, change in average lumen area was 0.22 mm^2 (p=0.061) by GT and 0.28 mm^2 (p=0.005) by the ML algorithm. Change in average vessel area was -0.57 mm^2 (p<0.001) by GT and -0.66 mm^2 (p<0.001) by the ML algorithm. The primary endpoint change in PAV was calculated as -2.94% (p<0.001) by GT and -3.74% p<0.001) by the ML algorithm. The secondary endpoint change in normalized total atheroma volume was -7.73 mm^3 (p<0.001) by GT and -9.44 mm^3 (p<0.001) by the ML algorithm.

The results from each parameter including change in mean lumen, vessel and plaque areas and change in PAV comparing GT to ML are presented in scatter plots in **Supplementary Figure 1**. The correlation coefficients were 0.85, 0.73, 0.54, and 0.76 for lumen area, vessel area, plaque area and PAV, respectively. The agreement between methods is represented with Bland-Altman plots in Supplementary Figure 2. For change in lumen area, the limits of agreement are -8.71 to 9.07 with a bias of 0.18 while for vessel area the limits range from -7.81 to 9.38 with a bias of 0.79. For change in plaque area, the limits of agreement are -19.36 to 22.55 with a bias of 1.59. Lastly, for plaque burden, the limits range from -4.4 to 5.09 with a bias of 0.34.

Discussion:

The main finding of this study is that the serial intracoronary imaging analysis by the CL (i.e., ground truth) and the ML algorithm both show significant regression in plaque burden after high-intensity rosuvastatin therapy. The direction and magnitude of the change were comparable not only for the primary endpoint, PAV, but also for the change in mean lumen, vessel, and plaque areas.

This is the first time a ML algorithm replicated the same IVUS analysis already read by a CL successfully. Comparing the primary endpoint, change in PAV, analyzed by each method for the entire region, both show a significant reduction from baseline to follow-up (-0.90% for GT and -1.28% for ML). A similar trend was seen for the most diseased 10mm segments where GT measured PAV decreased by 2.94% and ML algorithm measured PAV decreased by 3.74%. The ML algorithm “overestimates” PAV which in part is because PAV is computed after taking two differences in the middle of the process (vessel - lumen to define plaque, and follow-up - baseline to define change). When performing these derived entities, it is natural that small errors become amplified. Another plausible contributor is that GT analyses were done on the radiofrequency images (not more in use) and the ML algorithm analyzed the greyscale images. Importantly, the magnitude and the direction of the change are very similar and therefore we believe that these results open a new opportunity to use ML algorithms in prospective clinical trials. These methodologies should still be used within a CL environment so that the CL readers may overread the analysis performed by the software and finalize the case.

The relationship between GT results and the ML results for each parameter were represented in the scatter plots in **Supplementary Figure 1**. The highest correlation is seen with lumen area at 0.85 while the correlation of vessel area is also noteworthy at 0.73. For plaque area and plaque burden (percent atheroma volume), which are derived measurements using vessel and lumen area, the correlations are notable as well at 0.54 and 0.76, respectively. The agreement between methods was further described in Bland-Altman plots in **Supplementary Figure 2.** The methods perform most similarly when calculating lumen area, vessel area and plaque burden with minimal absolute differences of 0.18, 0.79 and 0.34, respectively. An automated ML segmentation method, especially when it is adequately trained and validated, can successfully replicate data from GT and result in the same research conclusions.

A low level of intraobserver variability when comparing IVUS readings from different analysts is desirable for standardization and consistency in clinical practice and research. Unlike ML algorithms which have no variability, reproducibility tested within a CL by Gerstein et al. showed that the intraobserver variability reported mean (SD) differences of −0.02 (0.23) mm2 for lumen area and 0.09 (0.18) mm2 for vessel area.[10] Although these differences seem to be small, another source of variability exists between different Core Laboratories. This has also been tested in Gerstein et al. where same cases were analyzed twice at different Core Laboratories (Cardialysis, Rotterdam, the Netherlands, and MedStar Research Institute, Division of Cardiology, Washington, DC). The mean (SD) differences were −0.07 (0.45) mm2 for lumen area and 0.53 (0.37) mm2 for vessel area.[10] Again, in contrast to manual methods applied by CL, the ML algorithm has no variability and produces the same output for the same input without fail.

The ML algorithm removes the time-consuming nature of manual segmentation. A prior study validated the improved reproducibility and time efficiency inherent in automatic segmentation compared to manual segmentation.[11] They directly compared 10 human operator manual readings from different hospitals to the ML algorithm and a Core Laboratory. Not surprisingly, relative differences revealed considerable variability in manual readings as compared to the Core Lab, especially in lumen area from 0.26% to 12.61%. The ideal scenario is such that human observations consistently match the Core Lab in terms of accuracy and precision. This is difficult with manual readings. Furthermore, while human analysts spent an average 47 minutes to complete the 40-frame analysis, the ML algorithm took on average less than one minute.

Limitations:

A few limitations should be acknowledged: 1. This is a post-hoc analysis of a single arm study; 2. The original report included the analysis of the radiofrequency frames, and the ML algorithm analyzed instead the greyscale frames. Although this may seem a real limitation, the reader may remember that this type of IVUS catheter acquires simultaneously both frames and there is always a set of two frames that correspond to each other; 3. The ML algorithm was originally trained with the IBIS-4 data which could have resulted in overfitting. The model thus may perform better in this setting when compared to a different dataset. However, during training of the algorithm, the full dataset was split into three groups. One of them was kept blindly while the algorithm was trained with the other two groups. This was repeated three times such that three algorithms

performed blind assessment of the entire IBIS-4 database. In this manner, we avoid the overfitting limitation.

Conclusion:

The ML algorithm can replicate results from a Core Laboratory, the often-used gold standard of IVUS readings. Both the automated method and ground truth showed a significant regression in PAV after high-intensity rosuvastatin therapy. Automated methods minimize variability and improve efficiency while providing similar results and conclusions compared to manual analyses. The use of a validated ML algorithm is thus warranted in the catheterization laboratory for clinical and research purposes.

Conflicts of Interest: **Disclosures**

Héctor M. García-García reports the following Institutional grant support: Biotronik, Boston Scientific, Medtronic, Abbott, Neovasc, Shockwave, Phillips and Corflow; Speaker's fee from Boston Scientific and Amgen.

Lorenz Räber reports institutional grant support: Abbott, Biotronik, Boston Scientific, Sanofi, Regeneron, Infraredx and consultation fees by Abbott, Amgen, AstraZeneca, Occlutec, Medtronic, Sanofi, Vifor.

Marco Roffi reports institutional research grants from Terumo, Biotronik, Medtronic and Cordis/Cardinal Health

Author Contributions:

All authors contributed significantly to draft and critically revise the work.

Acknowledgements:

Thank you to the IBIS-4 trial investigators for their collaboration.

References:

1. Sanchis-Gomar F, Perez-Quilis C, Leischik R, Lucia A. Epidemiology of coronary heart disease and acute coronary syndrome. *Ann Transl Med* 2016;**4**:256.

2. Nicholls SJ, Brandrup-Wognsen G, Palmer M, Barter PJ. Meta-analysis of comparative efficacy of increasing dose of Atorvastatin versus Rosuvastatin versus Simvastatin on lowering levels of atherogenic lipids (from VOYAGER). *Am J Cardiol* 2010;**105**:69–76.

3. Garcia-Garcia HM, Costa MA, Serruys PW. Imaging of coronary atherosclerosis: intravascular ultrasound. *European Heart Journal* 2010;**31**:2456–2469.

4. Nissen SE, Tuzcu EM, Schoenhagen P, Brown BG, Ganz P, Vogel RA, Crowe T, Howard G, Cooper CJ, Brodie B, Grines CL, DeMaria AN, REVERSAL Investigators. Effect of intensive compared with moderate lipid-lowering therapy on progression of coronary atherosclerosis: a randomized controlled trial. *JAMA* 2004;**291**:1071–1080.

5. Jensen LO, Thayssen P, Pedersen KE, Stender S, Haghfelt T. Regression of Coronary Atherosclerosis by Simvastatin. *Circulation* 2004;**110**:265–270.

6. Blanco PJ, Ziemer PGP, Bulant CA, Ueki Y, Bass R, Räber L, Lemos PA, García-García HM. Fully automated lumen and vessel contour segmentation in intravascular ultrasound datasets. *Medical Image Analysis* 2022;**75**:102262.

7. Ziemer PGP, Bulant CA, Orlando JI, Maso Talou GD, Álvarez LAM, Guedes Bezerra C, Lemos PA, García-García HM, Blanco PJ. Automated lumen segmentation using

multi-frame convolutional neural networks in intravascular ultrasound datasets. *European Heart Journal - Digital Health* 2020;**1**:75–82.

8. Li K, Tong J, Zhu X, Xia S. Automatic Lumen Border Detection in IVUS Images Using Deep Learning Model and Handcrafted Features. *Ultrason Imaging* SAGE Publications Inc; 2021;**43**:59–73.

9. Räber L, Taniwaki M, Zaugg S, Kelbæk H, Roffi M, Holmvang L, Noble S, Pedrazzini G, Moschovitis A, Lüscher TF, Matter CM, Serruys PW, Jüni P, Garcia-Garcia HM, Windecker S, for the IBIS 4 (Integrated Biomarkers and Imaging Study-4) Trial Investigators (NCT00962416). Effect of high-intensity statin therapy on atherosclerosis in non-infarct-related coronary arteries (IBIS-4): a serial intravascular ultrasonography study. *European Heart Journal* 2015;**36**:490–500.

10. Gerstein HC, Ratner RE, Cannon CP, Serruys PW, García-García HM, Es G-A van, Kolatkar NS, Kravitz BG, Miller DM, Huang C, Fitzgerald PJ, Nesto RW. Effect of Rosiglitazone on Progression of Coronary Atherosclerosis in Patients With Type 2 Diabetes Mellitus and Coronary Artery Disease. *Circulation* 2010;**121**:1176–1187.

11. Bass RD, García-García HM, Sanz-Sánchez J, Ziemer PGP, Bulant CA, Kuku KK, Kahsay YA, Beyene S, Melaku G, Otsuka T, Choi J, Fernández-Peregrina E, Erdogan E, Gonzalo N, Bourantas CV, Blanco PJ, Räber L. Human vs. Machine vs. Core Lab for the Assessment of Coronary Atherosclerosis with Lumen and Vessel Contour Segmentation with Intravascular Ultrasound. *Int J Cardiovasc Imaging* 2022; **38**:1431-1439.

Table 1: Baseline Characteristics

	All IBIS4 patients	No Machine Learning	Machine Learning	P-value
	N=103	N=21	N=82	
Age at date of informed consent	58.2 (10.5)	57.1 (12.9)	58.5 (9.9)	0.63
Gender (female)	10 (9.7)	4 (19.0)	6 (7.3)	0.12
BMI	27.8 (4.2)	29 (5.5)	27.5 (3.8)	0.27
Diabetes mellitus (yes)	13 (12.6)	4 (19.0)	9 (11.0)	0.46
Hypertension (yes)	48 (46.6)	9 (42.9)	39 (47.6)	0.81
Hypercholesterolemia (yes)	42 (40.8)	6 (28.6)	36 (43.9)	0.22
Current smoker (yes)	47 (45.6)	12 (57.1)	35 (42.7)	0.33
Family history of coronary artery disease	31 (30.1)	6 (28.6)	25 (30.5)	1.00
Renal failure (eGFR) (yes (<60 eGFR))	5 (4.9)	2 (9.5)	3 (3.7)	0.28
Previous myocardial infarction (yes)	2 (1.9)	0 (0.0)	2 (2.4)	1.00
Previous PCIs (yes)	1 (1.0)	0 (0.0)	1 (1.2)	1.00
Time from symptom onset to balloon inflation (min)	258 (170;472)	238 (168;374)	262 (170;476)	0.73
Left ventricular ejection fraction (angiography)	47.8 (9.4)	49.4 (12)	47.5 (8.8)	0.56
Resuscitation prior to hospital arrival (%)	5 (4.9)	0 (0.0)	5 (6.1)	0.58

IBIS4=integrated biomarker imaging study. BMI=body mass index. eGFR=estimated glomerular filtration rate. PCI=percutaneous coronary intervention

Table 2: Serial Intravascular Ultrasound Results

N=82 patients (146 vessels)	Baseline	Follow-up	Change	P-value
Entire region	mean (SD)	mean (SD)	mean (95% CI)	
Region length (mm)	36.15 (16.10)	36.21 (16.05)	0.06 (-1.06 to 1.17)	0.917
Region length by machine learning (mm)	34.77 (16.18)	35.06 (16.15)	0.29 (-0.98 to 1.56)	0.651
Average lumen area (mm^2)	8.64 (3.07)	8.48 (2.89)	-0.16 (-0.29 to -0.03)	0.020
Average lumen area by machine learning (mm^2)	8.88 (3.03)	8.72 (2.87)	-0.17 (-0.29 to -0.04)	0.009
Average vessel area (mm^2)	15.74 (5.64)	15.23 (5.38)	-0.51 (-0.67 to -0.35)	<0.001
Average vessel area by machine learning (mm^2)	16.17 (5.56)	15.54 (5.34)	-0.62 (-0.80 to -0.45)	<0.001
Average atheroma area (mm^2)	7.10 (3.21)	6.75 (3.15)	-0.35 (-0.49 to -0.21)	<0.001
Average atheroma area by machine learning (mm^2)	7.28 (3.13)	6.82 (3.04)	-0.46 (-0.59 to -0.33)	<0.001
Percent atheroma volume (%)	43.95 (9.66)	43.02 (9.82)	-0.90 (-1.56 to -0.25)	0.007
Percent atheroma volume by machine learning (%)	43.69 (9.11)	42.41 (9.17)	-1.28 (-1.84 to -0.72)	<0.001
Normalized total atheroma volume (mm^3)	248.40 (112.69)	235.95 (110.25)	-12.18 (-16.91 to -7.44)	<0.001
Normalized total atheroma volume by machine learning (mm^3)	239.71 (102.89)	224.44 (100.00)	-15.08 (-19.39 to -10.78)	<0.001
Most diseased 10mm segments				
Region length (mm)	9.96 (0.15)	10.25 (2.23)	0.29 (-0.08 to 0.66)	0.118
Region length (mm) by machine learning	9.93 (0.25)	10.43 (2.89)	0.50 (0.02 to 0.98)	0.041
Average lumen area (mm^2)	7.59 (2.97)	7.78 (2.96)	0.22 (-0.01 to 0.45)	0.061
Average lumen area by machine learning (mm^2)	7.80 (2.89)	8.08 (2.99)	0.28 (0.09 to 0.47)	0.005
Average vessel area (mm^2)	16.26 (5.61)	15.67 (5.52)	-0.57 (-0.84 to -0.31)	<0.001
Average vessel area by machine learning (mm^2)	16.44 (5.49)	15.76 (5.47)	-0.66 (-0.91 to -0.42)	<0.001
Average atheroma area (mm^2)	8.67 (3.78)	7.89 (3.53)	-0.78 (-0.99 to -0.56)	<0.001
Average atheroma area by machine learning (mm^2)	8.64 (3.59)	7.69 (3.32)	-0.95 (-1.15 to -0.75)	<0.001
Percent atheroma volume (%)	52.31 (12.17)	49.42 (11.65)	-2.94 (-3.89 to -1.98)	<0.001
Percent atheroma volume by machine learning (%)	51.55 (11.21)	47.81 (10.73)	-3.74 (-4.66 to -2.83)	<0.001
Normalized total atheroma volume (mm^3)	86.44 (37.76)	78.72 (35.08)	-7.73 (-9.86 to -5.59)	<0.001
Normalized total atheroma volume by machine learning (mm^3)	85.78 (35.59)	76.31 (32.93)	-9.44 (-11.44 to -7.44)	<0.001

CASS of Imaged Vessel (n)	DB	IMB	LAD	LCX	RCA
	2	1	50	55	38

Values at baseline and follow-up are mean (SD). Mean change (95%CI) are estimated from mixed-effect models.

DB=diagonal Branch. IMB=intermediate branch. LAD=left anterior descending. LCX=left circumflex.

RCA=right coronary artery.

Figure Titles and Legends:

Figure 1: Study Design and Key Finding

There are two panels (A and B) to represent the main comparison in this study. Panel (A) indicates that the IVUS dataset was first analyzed at baseline and follow-up in a previously published study by a core laboratory consisting of human experts. The lumen and vessel borders were manually extracted to calculate percent atheroma volume (PAV). The core lab assessment yielded a change in PAV of -0.90%. Panel (B) shows that the same dataset was analyzed in this study by a machine learning algorithm which automatically extracted the lumen and vessel borders. The machine calculated PAV at baseline and follow-up, generating a change in PAV of -1.28%. Both studies resulted in the same conclusion that high-intensity statin therapy is associated with regression of coronary atherosclerosis.

Figure 2: Intravascular Ultrasound (IVUS) Image Processing Pipeline

Raw data consists of ground truth (GT) lumen and vessel contours, which were manually segmented using radiofrequency IVUS images. Then, the GT contours were linked to the corresponding grayscale frames present in the gated DICOM dataset, and the masks for the lumen and plaque classes were generated. Grayscale frames and corresponding masks in polar coordinates are the input for training the machine learning (ML) algorithm. The ML strategy consists of a multi-frame convolutional neural network with U-Net architecture coupled to a Gaussian process regressor. In the validation and testing stages, only the polar coordinate grayscale image is fed to the network, and the

ML final prediction is the semantic segmentation that characterizes lumen and vessel contours.

Figure 1

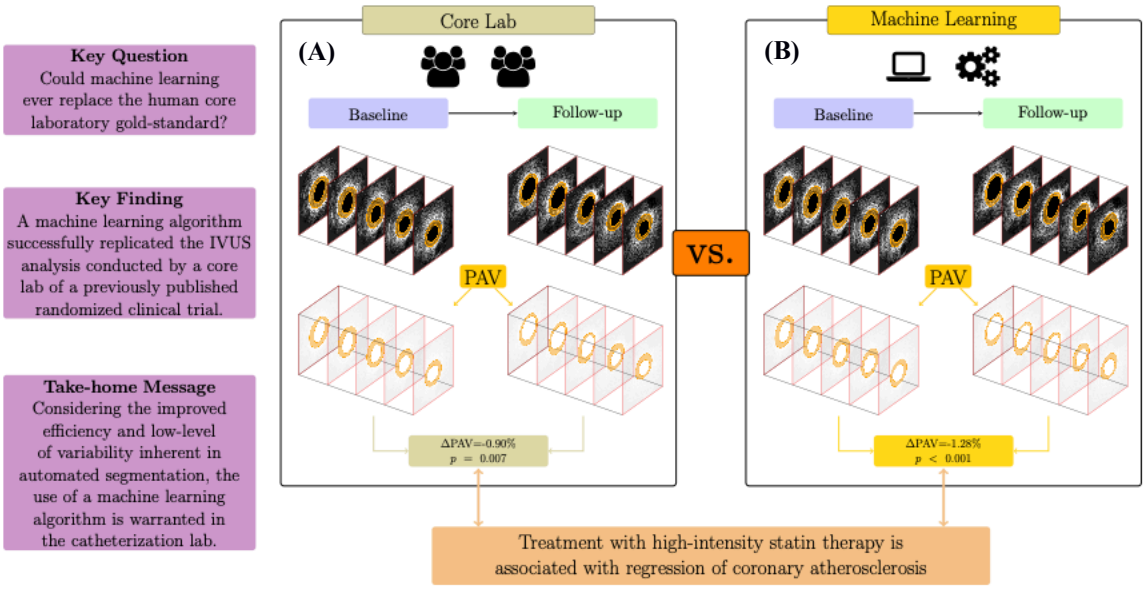
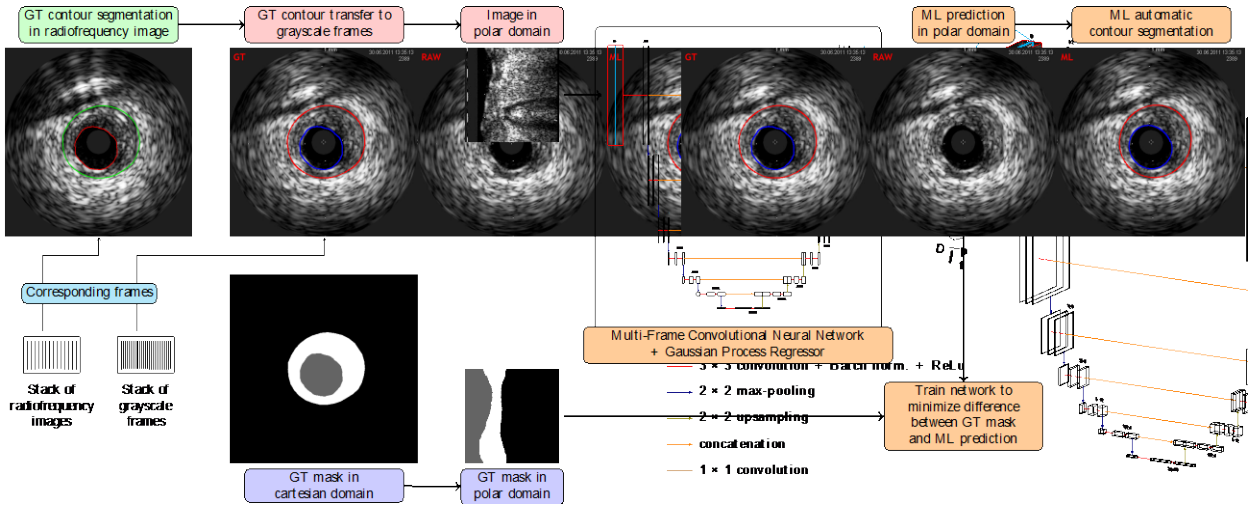

Key Question
Could machine learning ever replace the human core laboratory gold-standard?
Key Finding
A machine learning algorithm successfully replicated the IVUS analysis conducted by a core lab of a previously published randomized clinical trial.
Take-home Message
Considering the improved efficiency and low-level of variability inherent in automated segmentation, the use of a machine learning algorithm is warranted in the catheterization lab.
Core Lab
(A)
Baseline
Follow-up
PAV
ΔPAV=-0.90%
p = 0.007
VS.
Machine Learning
(B)
Baseline
Follow-up
PAV
ΔPAV=-1.28%
p < 0.001
Treatment with high-intensity statin therapy is associated with regression of coronary atherosclerosis

Figure 2


GT contour segmentation in radiofrequency image
GT contour transfer to grayscale frames
Image in polar domain
ML prediction in polar domain
ML automatic contour segmentation
Corresponding frames
Stack of radiofrequency images
Stack of grayscale frames
GT mask in cartesian domain
GT mask in polar domain
Multi-Frame Convolutional Neural Network + Gaussian Process Regressor
3 × 3 convolution + Batch norm. + ReLu
2 × 2 max-pooling
2 × 2 upsampling
concatenation
1 × 1 convolution
Train network to minimize difference between GT mask and ML prediction

Supplementary Material

Supplementary Methods: Description of the Machine Learning (ML) Algorithm

Next, we provide a description of the machine learning (ML) algorithm proposed by Blanco et al.[6] which was used in the present study for the automated segmentation of lumen and vessel contours in IVUS datasets.

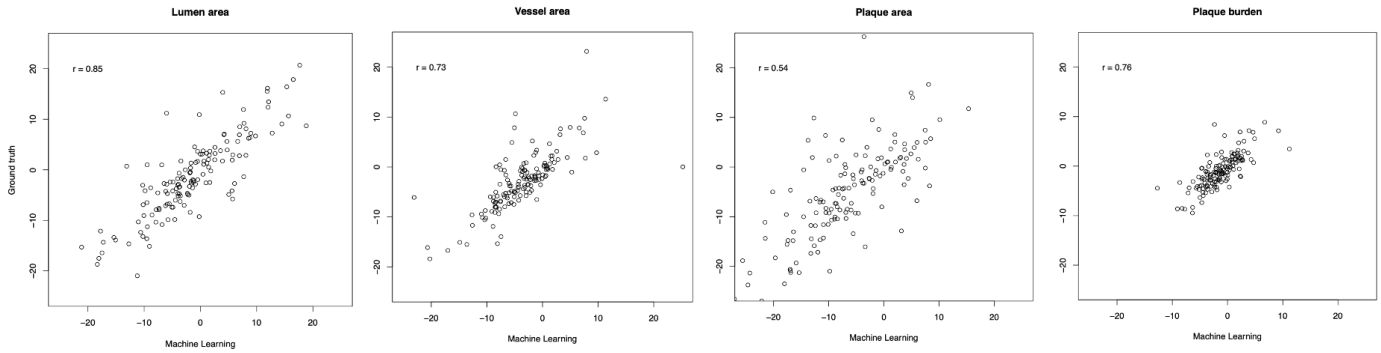
The ML algorithm consists of a deep learning strategy that produces the semantic segmentation of lumen and plaque classes, coupled to a Gaussian process regressor that yields the final lumen and vessel contours. The deep learning block is a convolutional neural network (CNN) which is fed with the frame of interest to be segmented and its neighboring frames in polar coordinates. The CNN architecture consists of a typical U-shaped network formed by combining convolution, batch normalization, and rectified linear unit activation functions with max-pooling (contracting branch) and up-sampling (expanding branch), as seen in **Figure 2**. The use of the neighborhood of the given frame (use of multiple frames) allows the methods to provide a more consistent segmentation along the longitudinal direction and characterizes the approach as a multi-frame CNN (MFCNN).

The entire dataset consists of 22089 IVUS frames manually annotated, regarded as the ground truth (GT) data. We divide, at per-pullback level, the entire dataset into training (60%) validation (20%) and testing (20%). A three-fold cross validation is employed so that the entire dataset is used in the study. The multi-class cross-entropy loss function was minimized using the ADAM method with learning rate initialized at 0.0001. Hyperparameter selection and model training were performed following conventional practices.[6]

The output of the MFCNN consists of a three-class image describing the lumen, plaque, and background in polar coordinates **(Figure 2)**. Still in the polar frame of

reference, the set of points that characterize the position of the lumen and vessel contours is regarded as a noisy periodic signal. Thus, a Gaussian process (GP) regressor is employed to create a periodic and smoothed signal that provides the final characterization of both contours. Mathematically, the GP algorithm is a non-parametric regression method based on the definition of a kernel function, taken to be the exp-sine-squared kernel, which is able to model periodic functions.[6] Using the GP algorithm we construct the final shape of the contours from which we compute the lumen area and vessel area which is then also used to derive plaque area and plaque burden from both GT and ML-generated data.

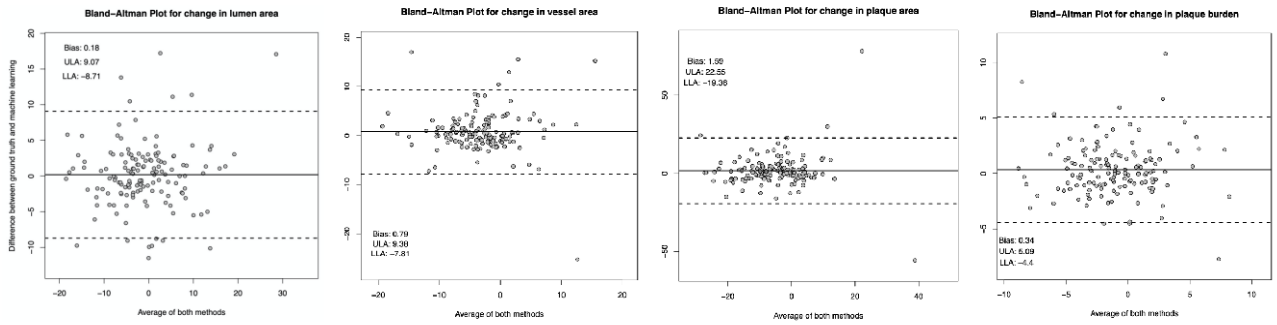
Supplementary Figure 1



Supplementary Fig. 1: Vessel-Level Correlations between Intravascular Ultrasound (IVUS) Endpoints Processed Using Machine Learning (ML) versus Ground Truth (GT)

From left to right, all scatter plots show the change of quantities between follow-up and baseline for the entire dataset (n=146) as computed from ground truth segmentations and from machine learning automatic segmentations. Positive values denote increase, and negative values, decrease. Percent of change is computed as follows: lumen area relative to baseline, vessel area relative to baseline, plaque area relative to vessel area at baseline, and plaque burden only difference between follow-up and baseline.

Supplementary Figure 2



Supplementary Fig. 2: Agreement between Ground Truth (GT) and Machine Learning (ML) Methods for Lumen Area, Vessel Area, Plaque Area and Plaque Burden

From left to right, all Bland-Altman plots display the change in quantities between baseline and follow-up for each measured parameter including lumen area, vessel area, plaque area and plaque burden. The y-axis shows the difference between the ground truth and machine learning measurements while the x-axis shows the average of both methods.