High-quality annotations for deep learning enabled plaque analysis in SCAPIS cardiac computed tomography angiography

Heliyon 9 (2023) e16058

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

High-quality annotations for deep learning enabled plaque analysis in SCAPIS cardiac computed tomography angiography

Erika Fagman [a,b], Jennifer Alvén [c,d], Johan Westerbergh [e], Pieter Kitslaar [f], Michael Kercsik [c,g], Kerstin Cederlund [h], Olov Duvernoy [i], Jan Engvall [j,k], Isabel Gonçalves [l,m], Hanna Markstad [m,n], Ellen Ostenfeld [o], Göran Bergström [c,p], Ola Hjelmgren [c,p,*]

[a] Department of Radiology, Institute of Clinical Sciences, University of Gothenburg, Sweden
[b] Department of Radiology, Sahlgrenska University Hospital, Gothenburg, Sweden
[c] Department of Molecular and Clinical Medicine, Institute of Medicine, University of Gothenburg, Sweden
[d] Computer Vision and Medical Image Analysis, Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden
[e] Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden
[f] Medis Medical Imaging Systems BV, Leiden, the Netherlands
[g] Department of Radiology, Alingsås Hospital, Alingsås, Sweden
[h] Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden
[i] Section of Radiology, Department of Surgical Sciences, Uppsala University, Sweden
[j] Department of Clinical Physiology and Department of Health, Medicine and Caring Sciences, Linkoping University, Linkoping, Sweden
[k] CMIV – Center for Medical Image Science and Visualization, Linkoping University, Linkoping, Sweden
[l] Department of Cardiology, Skane University Hospital, Lund, Sweden
[m] Cardiovascular Research Translational Studies, Clinical Sciences Malmö, Lund University, Sweden
[n] Department of Clinical Sciences Lund, Diagnostic Radiology, Lund University, Skane University Hospital, Lund, Sweden
[o] Department of Clinical Sciences Lund, Clinical Physiology, Lund University, Skane University Hospital, Lund, Sweden
[p] Department of Clinical Physiology, Sahlgrenska University Hospital, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Coronary Computed Tomography Angiography
Coronary plaque analysis
Annotated dataset
Deep Learning

ABSTRACT

Background: Plaque analysis with coronary computed tomography angiography (CCTA) is a promising tool to identify high risk of future coronary events. The analysis process is time-consuming, and requires highly trained readers. Deep learning models have proved to excel at similar tasks, however, training these models requires large sets of expert-annotated training data. The aims of this study were to generate a large, high-quality annotated CCTA dataset derived from Swedish CArdioPulmonary BioImage Study (SCAPIS), report the reproducibility of the annotation core lab and describe the plaque characteristics and their association with established risk factors.

Methods and results: The coronary artery tree was manually segmented using semi-automatic software by four primary and one senior secondary reader. A randomly selected sample of 469 subjects, all with coronary plaques and stratified for cardiovascular risk using the Systematic Coronary Risk Evaluation (SCORE), were analyzed. The reproducibility study (n = 78) showed an agreement for plaque detection of 0.91 (0.84–0.97). The mean percentage difference for plaque volumes was −0.6% the mean absolute percentage difference 19.4% (CV 13.7%, ICC 0.94). There

* Corresponding author. Department of Clinical Physiology, Sahlgrenska University Hospital, SE-413 45, Gothenburg, Sweden.
E-mail address: ola.hjelmgren@wlab.gu.se (O. Hjelmgren).

https://doi.org/10.1016/j.heliyon.2023.e16058
Received 4 April 2023; Received in revised form 2 May 2023; Accepted 4 May 2023
Available online 11 May 2023

was a positive correlation between SCORE and total plaque volume (rho = 0.30, $p < 0.001$) and total low attenuation plaque volume (rho = 0.29, $p < 0.001$).
Conclusions: We have generated a CCTA dataset with high-quality plaque annotations showing good reproducibility and an expected correlation between plaque features and cardiovascular risk. The stratified data sampling has enriched high-risk plaques making the data well suited as training, validation and test data for a fully automatic analysis tool based on deep learning.

1. Introduction

Coronary artery disease (CAD) is the most common cause of morbidity and mortality in the western world [1]. The underlying substrate for CAD is in most cases atherosclerosis of the coronary arteries. With the advent of modern imaging techniques such as coronary computed tomography angiography (CCTA), coronary plaques can now be visualized, and distribution of disease can be described as well as plaque features including size, shape and composition [2]. It is believed that a detailed plaque analysis using image data from CCTA has the potential to improve risk prediction of CAD beyond that of traditional risk scores and coronary artery calcifications, to better identify individuals at high risk of suffering future coronary events [3]. A number of studies have also shown that CCTA not only provides important information on the degree of stenosis but that certain high-risk plaque (HRP) features are associated with increased risk of acute coronary syndrome [4–6].

Previous studies on CCTA plaque analysis were mainly designed as observational studies on patient populations with suspected chronic or acute coronary syndrome [4,7–13]. In the SCOT-HEART trial, studying patients with chronic coronary syndrome, low-attenuation plaque (LAP) burden was the strongest predictor of future myocardial infarction beyond coronary artery calcium score (CACS) and coronary artery stenosis [4]. In patients with acute coronary syndrome (ROMICAT 1 and 2 trial), CT plaque characteristics (positive remodelling and spotty calcifications) and quantitative plaque variables (LAP volume and stenosis length) had a good discriminatory value for the prediction of acute coronary syndrome incremental to the degree of stenosis [8–10]. However, there are very limited data on risk prediction from CCTA plaque analysis in populations free of previously known CAD and with asymptomatic coronary artery atherosclerosis.

Using data from the Swedish CArdioPulmonarybioImage Study (SCAPIS, n = 30,154) we have previously used CCTA to show that silent coronary artery atherosclerosis is common in the general population [14]. In SCAPIS, CCTA was visually assessed by trained readers focusing on basic plaque characterisation and obstructive disease, limiting the level of granularity in the data and not assessing HRP features in a systematic way. Therefore, access to an automatized technique for CCTA plaque analysis is needed to facilitate generation of such data in a large cohort. However, to develop a deep learning based model for automated CCTA analysis we need access to large and well annotated datasets for model training and validation.

Here we report the first step in such development, presenting data on coronary atherosclerosis on 469 participants in SCAPIS with detailed manual segmentations of the coronary arteries and coronary artery plaques. The cohort was selected from participants in SCAPIS having signs of coronary atherosclerosis on visual assessment and stratified for their 10-year risk of developing cardiovascular disease using the Systematic Coronary Risk Evaluation (SCORE) algorithm [15]. The stratification was performed to generate a dataset containing a large amount of plaques, which is crucial for efficient training of a deep learning model designed to find and correctly outline coronary plaques.

1.1. The specific aims of this study were to

(1) Generate a large, high-quality annotated image dataset that enables development of a fully automatic deep learning model for plaque analysis.
(2) Report the reproducibility and feasibility of the plaque analysis.
(3) Describe plaque characteristics in the dataset and study how different plaque characteristics associate with established risk factors for CAD and with CACS.

2. Materials and methods

2.1. Study population

The study subjects were recruited from SCAPIS. The design of SCAPIS has been previously described [16]. Briefly, between 2013 and 2018, 30,154 SCAPIS participants (50–64 years old) were randomly recruited from the census register and invited to a comprehensive examination including CCTA (n = 27,385). All CCTA examinations performed in SCAPIS were visually assessed by radiologists at each site and per-segment status of the coronary vessel was defined as: no atherosclerosis; 1–49% stenosis; ≥50% stenosis (significant stenosis). Coronary artery calcium score (CACS) was calculated with the Agatston method [17]. The SCAPIS study was approved as a multi-center study by the ethical review board in Umeå (# 2010-228-31 M). The participants gave written informed consent. The current sub-study was approved by the ethical review board in Gothenburg, (#570-18).

We included a selection of 500 CCTA examinations from the total of 27,385 CCTA examinations performed in SCAPIS. Participants were eligible for selection when having at least one coronary plaque in the visual assessment in the SCAPIS study and having all data

available for calculation of SCORE (age, sex, smoking, systolic blood pressure and total cholesterol) [15]. Exclusion criteria were; previous myocardial infarction, previous stroke or previous cardiac procedures (coronary artery bypass grafting or percutaneous coronary intervention). From all eligible participants, a randomly selected sample, stratified for risk of future CAD was then selected using the SCORE algorithm that represents the estimated 10-year risk (%) of fatal cardiovascular disease. Only image sets acquired using a 100 kV tube voltage were included in the study (representing 94.6% of all examinations). An equal number of subjects were randomized from each of six strata of SCORE: 0–1%; 2%: 3%; 4%; 5%;≥6% reaching a total number of 500 CCTA examinations.

To perform a reproducibility analysis, a total of 78 unique examinations were sampled in a similar way across strata of SCORE. However, to test the ability of the image laboratory to detect coronary plaques, 36 of the examinations had no coronary plaques, based on the visual reading in SCAPIS.

2.2. Data acquisition

The CT protocol in SCAPIS has been described in detail previously [16]. In short, CT was performed on dedicated dual-source CT scanners (Somatom Definition Flash, Siemens Medical Solution, Forchheim, Germany) and included a non-contrast-enhanced ECG--gated scan used for calcium scoring [17] and iodine contrast-enhanced (Omnipaque, GE Healthcare, 350 mgI/ml) CCTA using five different protocols depending on heart rate, heart rate variability, presence of calcifications, and body weight. A β-blocker (metoprolol) and sublingual glyceryl nitrate were administrated to decrease the heart rate and dilate the coronary arteries.

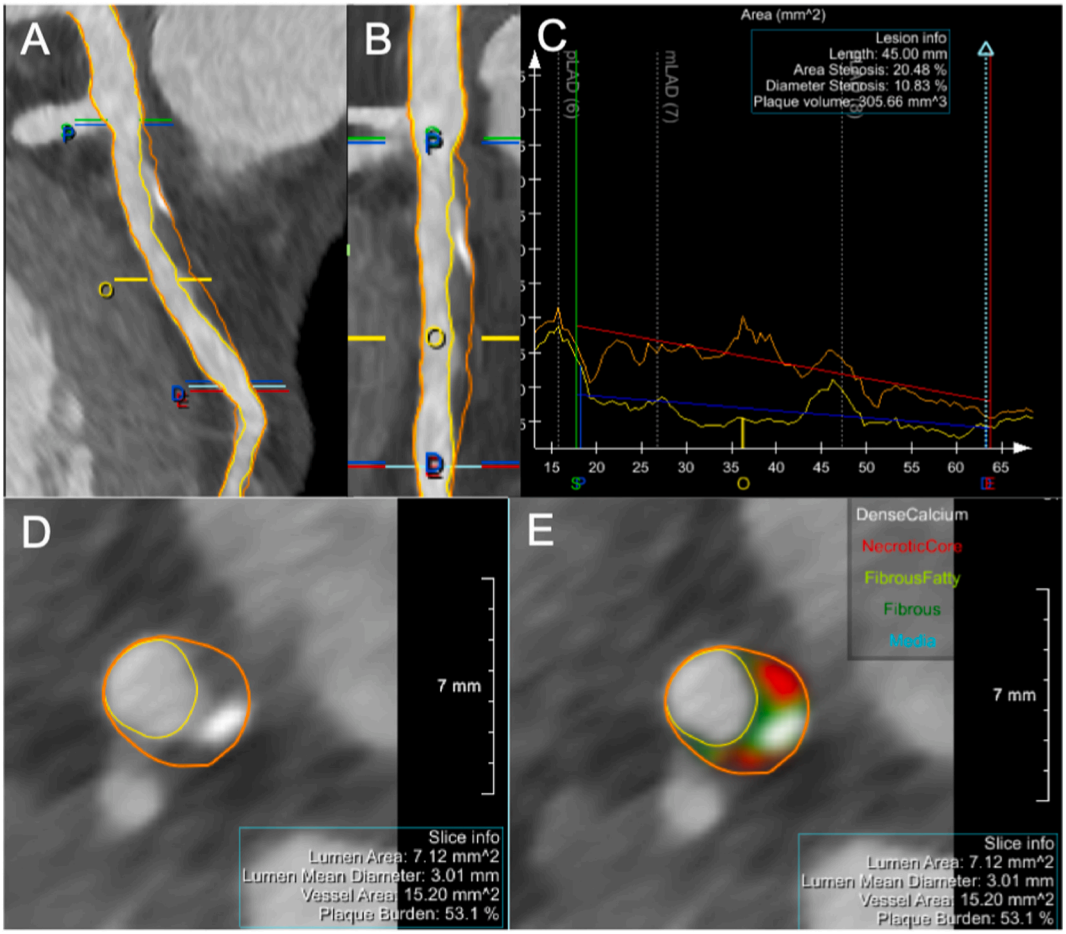

Fig. 1. Semiautomated plaque analysis. Curved MPR (A) and straightened MPR (B) views show a mixed plaque in proximal left anterior descending artery (LAD). Vessel wall contour (orange line) and lumen contour (yellow line) are automatically detected by the software and manually corrected if needed. Markers P and D are placed at start and end of the plaque and marker O the point of maximal stenosis. C is a graphic representation of the plaque area along the vessel length with calculation of total plaque volume. D and E show a representative cross section of the plaque where plaque components are color coded in E (Dense Calcium: HU 351–2048; Low attenuation plaque: 30-30 HU; Fibrous fatty tissue: 31–150 HU; Fibrous tissue: 151–350 HU). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.3. Image analysis

For plaque analysis, a dedicated cardiac image analysis laboratory was established. The analysis was performed in a two-step process with four primary readers and one expert secondary reader (author EF, senior cardiac radiologist with 13 years of experience on cardiac CT level 3). The primary readers were four radiographers, that had more than seven years of previous experience of performing CCTA examinations, but no previous experience of analyzing CCTA images. Before the start of the study, the primary readers underwent a training program including a two-day course on CCTA protocols, plaque analysis and the software application followed by hands-on training on a training dataset. All study examinations were first analyzed by one of the four primary readers and then the analysis was finalized by the expert reader by correcting the annotations and contours in the dataset if needed. The reader group had weekly recurring feedback meetings in which difficult or interesting cases were demonstrated to the primary readers.

Plaque analysis was performed using dedicated workstations (Medis suite version 3.1.16.8) and a semiautomated plaque analysis software (QAngio CT RE Version 3.1.4.2, both from Medis Medical Imaging Systems, Leiden, the Netherlands). The annotation process was performed in standardized steps and is described in detail in Appendix 1. In summary, all coronary vessels down to a diameter of 2 mm were annotated with vessel wall and lumen contours and all plaques were marked manually (Fig. 1A–E).

The CCTA annotations were exported and converted to labeled volumes suitable for deep learning model development. The details of this procedure are described in Appendix 2.

2.4. Reproducibility study

The primary aim of the reproducibility study was to assess the ability of the cardiac image analysis laboratory to separate subjects with coronary disease (one or more plaques) from healthy (no plaques). Secondary aims were to assess the reproducibility of plaque volumes and plaque characteristics. Power calculations for the primary aim are presented in Appendix 3. The 78 examinations included in the reproducibility study were duplicated and tagged with separate codes. To assure a blinded analysis, the examinations were shuffled and presented in a random order to the two primary readers and all cases were then corrected by the senior reader. The reproducibility analysis was performed on the output of the analysis by the senior reader (analysis 1 vs analysis 2). Reproducibility was analyzed on a per-individual level for the detection of any plaque (yes/no) and on the total plaque volume in an individual. Sub analysis was performed on the three main vessels LAD, RCA and CX. Furthermore, reproducibility was performed on a per-plaque level for matched plaques. Two plaques were defined as matching when they had at least 60% overlap in position along the vessel centerline. In matched plaques, reproducibility of plaque volume and plaque components were analyzed.

2.5. Definitions used for plaque characterization

Total plaque volume was defined as the sum of the plaque volumes of all marked plaques in one examination. Quantification of plaque components was automated and based on the following attenuation thresholds in Hounsfield units (HU): low attenuation plaque (LAP)-30-30 HU; fibrous fatty tissue 31–150 HU; fibrous tissue 151–350 HU and calcium 351–2048 HU [5,18,19]. Plaque burden was calculated as the plaque area divided by outer vessel area at the site of maximum stenosis.

2.6. Definitions used for high-risk plaques (HRP)

We used a modified version of the CAD-RADS system for defining HRP [20]. Two of the four suggested HRP features were used: positive remodelling and LAP. The napkin ring sign as well as spotty calcifications were not assessed, due to difficulty to establish an objective and reproducible way to assess these features. The remodelling index was calculated as the ratio between the outer vessel wall area at the site of maximum stenosis and the mean vessel area at the two reference points. A remodelling index $\geq$1.1 was defined as positive remodelling and considered a HRP feature [20]. If the LAP volume was $>$1.31 mm^3, the plaque was considered a LAP and defined as a HRP [13]. The number of HRP features were calculated for each plaque and the individual was classified according to the plaque with most HRP features.

3. Statistics

Subject characteristics were summarized using frequencies for categorical variables and median and interquartile range (IQR) for continuous variables. The agreement for plaque detection was evaluated using Cohen's Kappa and by calculating the proportion of agreement. The agreement for plaque volumes was analyzed both with regard to systematic error and with regard to precision. The former was done by calculating the % difference (the difference between the raters divided by the mean volume of the raters) and the later by calculating the absolute % difference, the coefficients of variation (CV%) and for matched samples also the interclass correlation (ICC). The associations between plaque volumes and SCORE as well as plaque volumes and CACS were evaluated using the Spearman correlation test (rho). The tests were 2-tailed and performed at the 0.05 significance level. R version 3.6.1 was used for all analyses.

4. Results

4.1. Study cohort characteristics

Of the 27,385 participants in SCAPIS who performed CCTA, 10,580 were eligible for inclusion in the current study according to the inclusion criteria with the following distribution among the 6 strata of SCORE: SCORE 0–1 52.1%; SCORE 2 25.2%; SCORE 3 12.0%; SCORE 4 5.5%; SCORE 5 2.7% and SCORE ≥6 2.5%. From this cohort, 500 cases were randomized. An initial image quality check was performed by the senior reader of all cases randomized for the study. In 35 cases, the image quality was considered too poor to perform plaque analysis, and these cases were replaced with new randomized cases until reaching 500 cases. In 31 of the 500 cases, no plaques were detected in the analysis. Therefore, the cohort for this study consists of 469 examinations. Characteristics of these subjects, and of the entire SCAPIS cohort, are shown in Table 1.

Women are less common in this study (14.9%) than in the SCAPIS study (50.3%), and the majority of women are sampled in stratum 1 of SCORE. Age, blood pressure and cigarette smoking increase with increasing SCORE strata, as expected. The stratified selection of participants in the current study had an average risk of 10-year fatal cardiovascular event (SCORE) of 3.6% while all participants who performed CCTA in SCAPIS had an average SCORE of 1.4%.

4.2. Feasibility of the cardiac image analysis laboratory

The annotation work was performed between May 2019 and September 2021. At the end of the study, the primary readers spent on average 60 min on analyzing one examination. This time was significantly longer during the first half of the study, averaging 130 min. The senior secondary reader spent on average 30 min on each examination. Cases without coronary plaques required significantly less time, both in primary and secondary reading (average 25 and 10 min, respectively).

5. Reproducibility study

5.1. Plaque detection and total plaque volume

Out of 78 subjects selected for the reproducibility study, 10 cases were excluded due to poor image quality and 68 subjects were included in the analysis. One or more coronary plaques was detected in 41 out of 68 cases (analysis 1) and 39 out of 68 cases (analysis 2) respectively. The proportion of agreement was 0.91 (0.84–0.97) with a kappa value of 0.82 (0.68–0.96). Bland-Altman plots for total

Table 1
Characteristics of included subjects and all SCAPIS participants with CCTA for comparison. Data presented as median (IQR) or counts (%). DLP, dose length product; BMI, body mass index.

	All SCAPIS participants N = 27835	Study subjects N = 469	SCORE 0-1 N = 83	SCORE 2 N = 75
Age (years)	57.4 (53.6–61.2)	61.7 (58.7–63.6)	55.9 (52.8–60.1)	60.5 (58.2–62.1)
Female sex	13767 (50.3%)	70 (14.9%)	44 (53.0%)	11 (14.5%)
Weight (kg)	79.2 (69.0–90.0)	82.4 (74.9–91.0)	78.0 (69.0–84.6)	83.7 (75.1–92.6)
BMI (kg/cm2)	26.3 (23.9–29.3)	26.6 (24.3–29.0)	26.1 (23.4–28.4)	27.1 (24.6–29.1)
Total cholesterol (mmol/l)	5.4 (4.80–6.10)	5.8 (5.1–6.5)	5.4 (4.8–6.2)	5.6 (5.0–6.3)
Systolic Blood pressure	124 (114–136)	136 (124–150)	122 (114–132)	128 (120–135)
Diastolic blood pressure	77 (70–84)	82 (74–89)	75 (68–82)	78 (72–84)
Current smoker	3341 (12.2%)	150 (32.9%)	6 (7.7%)	12 (16.4%)
Radiation dose, DLP (mGy*cm)	79 (57–106)	85 (69–105)	75 (58–93)	86 (69–100)
Calcium score >0%	40.5%	94% (441)	89% (74)	97% (73)
Calcium score (Agatston units) For subjects with CACS >0	36 (8–129)	69 (16–200)	22 (7–72)	86 (18–211)

	SCORE 3 N = 88	SCORE 4 N = 73	SCORE 5 N = 69	SCORE≥6 N = 81
Age (years)	62.6 (60.3–63.5)	62.7 (59.7–64.0)	62.2 (61.0–63.6)	63.3 (61.7–64.5)
Female sex	6 (6.9%)	7 (9.6%)	0 (0.0%)	2 (2.5%)
Weight (kg)	83.0 (76.3–93.1)	83.2 (73.4–91.2)	87.6 (78.2–92.8)	81.0 (75.2–90.6)
BMI (kg/cm2)	26.7 (24.8–28.4)	25.7 (23.9–28.4)	27.4 (24.1–29.3)	26.4 (24.1–28.9)
Total cholesterol (mmol/l)	5.7 (5.0–6.3)	6.2 (5.3–6.8)	6.0 (5.2–6.6)	6.1 (5.3–6.7)
Systolic Blood pressure	132 (124–146)	143 (130–154)	146.5 (132–160)	156 (142–169)
Diastolic blood pressure	80 (75–88)	84 (77–92)	86 (76–92)	88 (82–97)
Current smoker	20 (23.5%)	23 (31.9%)	35 (50.7%)	54 (68.4%)
Radiation dose, DLP (mGy*cm)	90 (73–112)	82 (72–98)	89 (74–108)	86 (73–113)
Calcium score >0%	93% (82)	97% (71)	94% (65)	94% (76)
Calcium score (Agatston units) For subjects with CACS >0	83 (18–224)	47 (13–124)	83 (19–232)	117 (39–411)

plaque volume per individual are shown in Fig. 2A. The mean % difference between the two analyses was 13.2% and the mean absolute % difference was 39.4%.

5.2. Matched plaques

A total number of 61 matched plaques were identified. Descriptive statistics for matched plaques are found in Table 2.

Bland Altman plots for the plaque volumes are shown in Fig. 2B. The mean difference between the two readings was −0.6% and the mean absolute difference was 19.4% with a CV% of 18.2 and ICC = 0.98. The agreements of the different plaque components are shown in Table 3. The agreement improved when removing the smallest plaques (volume<25 mm^3), CV% = 15.5.

LAP was the plaque component that showed highest variation between the two readings. Only four plaques had a zero value for LAP. However, the majority of the plaques had small LAP volumes where 66% had a volume <1 mm^3 and 43% had a volume <0.25 mm^3. To further test the reproducibility, a dichotomization process was performed, where three different cut-offs of LAP volume were tested; 0 mm^3, 0.25 mm^3 and 1 mm^3. A cut-off of 0.25 mm^3 yielded the best agreement with a kappa value of 0.93 (0.84–1.00).

5.3. General plaque characteristics

In the 469 CCTA studies, a total of 1640 plaques were found, ranging between 1 and 17 plaques per subject. The distribution of plaques was as follows: Left main: 78 plaques (5%), LAD 771 plaques (47%), CX 306 plaques (19%), RCA 433 plaques (26%) and intermediate branch 38 plaques (2%). A total of 90 plaques (5%) had a degree of stenosis ≥50% found in 63 subjects (13%). These plaques were found in left main (1 plaque), LAD (55 plaques), CX (13 plaques), RCA (17 plaques) and intermediate branch (4 plaques). Median plaque volume was 28.7 mm^3 (13.3–73.1). Total plaque volume per individual was 131 mm^3 (53.2–304.8). The majority of the plaques contained calcifications; 98 plaques (6%) were non-calcified or had less than 1 mm^3 of calcium found in 86 subjects (18%).

5.4. High risk plaque features

We found 547 LAP (33%) in 317 subjects (68%). Positive remodelling was found in 281 plaques (17%) in 197 subjects (42%). At least one of these two HRP features was present in 745 plaques (45%) found in 376 subjects (80%). There was a total of 86 (5%) two-feature positive plaques found in 74 subjects (16%).

5.5. Plaque characteristics and SCORE

There was a positive correlation between total plaque volume and SCORE (rho = 0.30, p < 0.001) and between total LAP volume

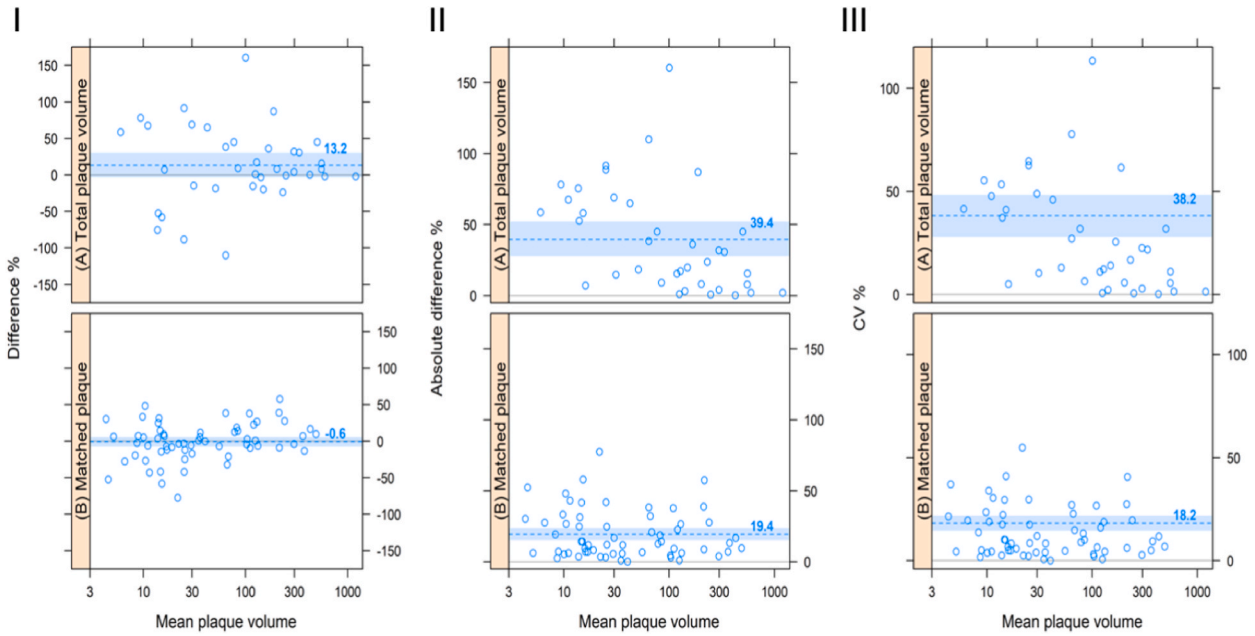

Fig. 2. I: Difference in percent between analysis 1 and analysis 2 for total plaque volume (A) and for plaque volume in matched plaques (B), 95% confidence intervals included. **II:** Absolute difference in percent between analysis 1 and analysis 2 for total plaque volume (A) and for plaque volume in matched plaques (B), 95% confidence intervals included. **III:** Variation in CV% between analysis 1 and analysis 2 for total plaque volume (A), and for plaque volume in matched plaques (B), 95% confidence intervals included.

Table 2
Descriptive statistics for matched plaques. Data presented as median (IQR).

	Analysis 1	Analysis 2
Number of matched plaques	61	
Plaque volume, mm^3	30.2 (13.6–102.5)	24.6 (13.3–105.6)
Low attenuation plaque, mm^3	0.3 (0.1–2.5)	0.4 (0.1–3.2)
Fibrous fatty volume, mm^3	3.9 (1.4–18.9)	4.0 (2.0–20.8)
Fibrous volume, mm^3	9.3 (4.5–30.4)	11.7 (5.8–32.2)
Calcium volume, mm^3	12.1 (6.2–35.6)	8.4 (4.3–32.9)
Plaque length, mm	7.0 (3.5–13.5)	7.0 (4.5–16.0)

Table 3
Agreement of plaque components in matched plaques. **Diff%:** Mean difference in %. **AbsDiff%**: mean absolute difference in %. **CV%**: coefficient of variation. **ICC**: intra class correlation for log2 values. Zero values for low attenuation plaque were set to NA.

All plaques

	Diff%	AbsDiff%	CV%	ICC
Plaque volume	−0.62	19.38	18.17	0.98
Low attenuation plaque	27.02	65.42	49.91	0.91
Fibrous fatty volume	15.92	36.43	33.37	0.94
Fibrous volume	8.08	20.20	18.63	0.97
Calcium volume	−22.68	26.00	26.64	0.96

Plaques with total plaque volume≤25 mm^3

	Diff%	AbsDiff%	CV%	ICC
Plaque volume	−6.49	22.30	20.93	0.80
Low attenuation plaque	21.01	79.55	54.07	0.64
Fibrous fatty volume	10.6	37.57	33.80	0.67
Fibrous volume	9.80	22.70	20.82	0.81
Calcium volume	−32.44	34.17	33.50	0.78

Plaques with total plaque volume > 25 mm^3

	Diff%	AbsDiff%	CV%	ICC
Plaque volume	4.36	16.90	15.46	0.97
Low attenuation plaque	31.75	54.28	47.08	0.89
Fibrous fatty volume	20.44	35.46	33.01	0.90
Fibrous volume	6.62	18.07	16.55	0.96
Calcium volume	−14.39	19.08	18.97	0.98

and SCORE (rho = 0.29, $p < 0.001$) (Fig. 3). Differences in plaque characteristics of the largest plaque in each subject between SCORE strata are shown in Table 4. All variables except remodelling index showed a positive correlation with SCORE strata.

Total calcium volume in plaques showed strong positive correlation to CACS reported from the non-enhanced ECG-gated CT scan in the SCAPIS study (rho = 0.94, $p < 0.001$).

6. Discussion

In this study we present a high-quality dataset of fully annotated CCTA examinations. The reproducibility of the cardiac image analysis laboratory is good and the data contains detailed information on extent, volume and characteristics of coronary artery plaques. The plaque characteristics show correlations to established risk factors as expected. All these findings suggest high validity of the generated data, a factor that is important when the data will be used for developing deep learning models. The stratified sampling strategy performed well in order to enrich the dataset with larger plaques and plaques with vulnerable features.

6.1. Feasibility and reproducibility of the cardiac image analysis laboratory

Plaque assessment and annotation in CCTA images is time-consuming and to be able to analyse large datasets we have built an annotation workflow where the first and most time-consuming step was performed by radiographers without previous training in CCTA analysis. The second step was a less time-consuming assessment by a senior radiologist that focused on plaque detection and contouring corrections. A great deal of effort was put into recurring feedback meetings with the primary readers for education purposes. The work-flow was further supported by a semiautomated plaque analysis software. However, even in this streamlined setting, a full analysis of the 27,385 image sets in the whole SCAPIS CCTA cohort would require six primary readers and three senior readers

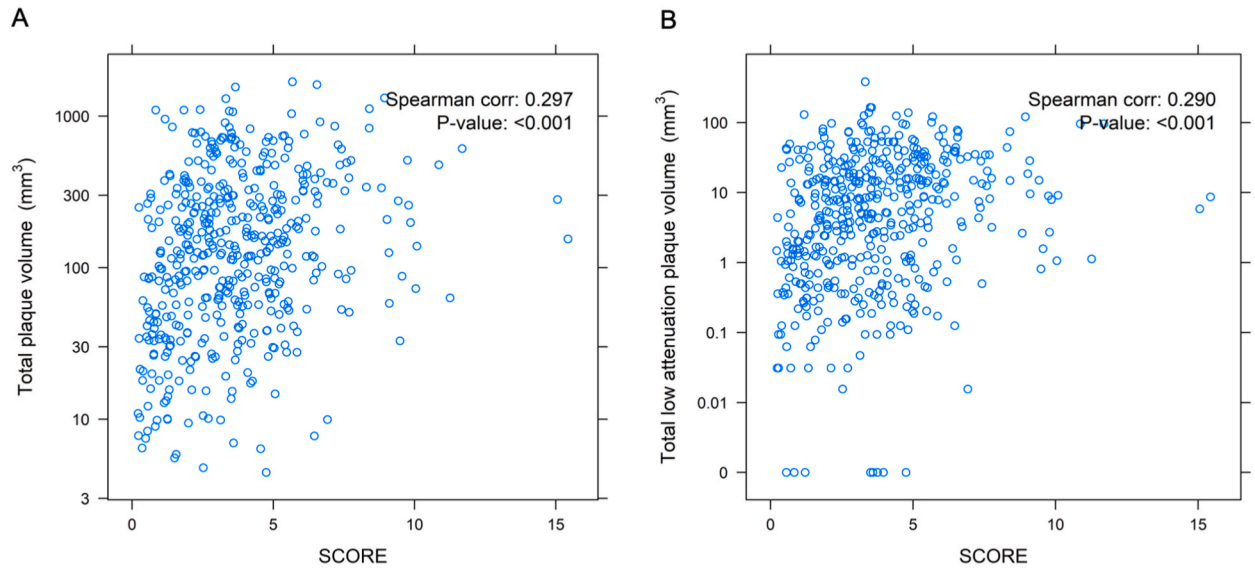

Fig. 3. Total plaque volume (A) and total low attenuation plaque volume (B) versus SCORE. Spearman correlation and test of Spearman correlation included.

Table 4
Plaque component volumes and plaque characteristics for the largest plaque in each subject by SCORE strata. Data presented as median (IQR).

	All subjects	SCORE 0-1 N = 93	SCORE 2 N = 79	SCORE 3 N = 90
Plaque volume (mm^3)	81.9 (33.8–182.6)	32.8 (19.3–79.1)	81.0 (40.9–146.6)	101.1 (46.7–220.4)
Low attenuation plaque (mm^3)	3.4 (0.7–16.3)	1.1 (0.3–3.3)	2.9 (0.7–9.7)	3.9 (0.9–17.4)
Fibrous fatty volume (mm^3)	19.5 (5.6–48.6)	5.7 (3.1–23.5)	15.7 (6.7–32.5)	23.9 (8.4–53.3)
Fibrous volume (mm^3)	25.5 (12.0–53.6)	13.0 (7.9–23.1)	25.6 (15.7–42.0)	31.4 (17.5–60.3)
Calcium volume (mm^3)	17.4 (6.0–44.2)	10.3 (4.1–23.0)	20.8 (7.9–50.5)	17.9 (6.3–57.2)
Diameter stenosis	0.2 (0.1–0.3)	0.1 (0–0.2)	0.2 (0.1–0.3)	0.2 (0.1–0.3)
Plaque burden	0.4 (0.3–0.5)	0.3 (0.3–0.4)	0.4 (0.3–0.5)	0.4 (0.3–0.5)
Remodelling index	1.0 (0.9–1.1)	1.0 (0.9–1.1)	1.0 (0.9–1.1)	1.0 (0.8–1.1)

	SCORE 4 N = 79	SCORE 5 N = 76	SCORE≥6 N = 83	Spearman correlation	
				Correlation coefficient	P -value
Plaque volume (mm^3)	74.4 (33.3–211.5)	92.6 (41.8–175.9)	148.0 (68.9–257.7)	0.278	<0.001
Low attenuation plaque (mm^3)	4.2 (0.4–18.1)	8.0 (0.9–20.7)	9.6 (1.8–26.8)	0.257	<0.001
Fibrous fatty volume (mm^3)	16.6 (4.3–55.4)	26.6 (6.9–52.4)	35.2 (13.5–67.0)	0.271	<0.001
Fibrous volume (mm^3)	24.4 (11.5–50.8)	25.2 (13.0–55.6)	42.7 (20.2–71.5)	0.253	<0.001
Calcium volume (mm^3)	15.8 (5.8–39.7)	15.9 (7.2–43.5)	32.5 (9.7–67.4)	0.169	<0.001
Diameter stenosis	0.2 (0.1–0.4)	0.2 (0.1–0.3)	0.2 (0.1–0.4)	0.208	<0.001
Plaque burden	0.4 (0.3–0.5)	0.4 (0.3–0.5)	0.5 (0.4–0.5)	0.258	<0.001
Remodelling index	1.0 (0.9–1.1)	1.0 (0.9–1.1)	1.0 (0.9–1.0)	−0.096	0.037

working full time for two years approximately. This clearly shows the need of an automatized process for CCTA analysis.

Our cardiac image analysis laboratory showed good reproducibility for plaque detection with an agreement for plaque detection of 0.91. Plaque volumes were highly reproducible and improved further when removing the smallest plaques.

6.2. *Plaque characteristics*

The present study cohort is a risk-stratified selection from the SCAPIS population study resulting in a cohort with a higher cardiovascular risk compared to the average SCAPIS participant. The study subjects had a median total plaque volume around 130 mm^3 which is lower than many previously published studies enrolling patients with a symptomatic indication for CCTA [8,21] or in clinical trials [22]. This is in line with the fact that the SCAPIS cohort represents the early subclinical phase of the atherosclerotic disease process.

Specific plaque features signal high risk of future events irrespective of degree of stenosis. Current clinical guidelines suggest the

reporting of such HRP and identifying plaques that are two-feature positive [2,23]. In this study, without previous CAD, HRP features were quite common with 64% having one-feature positive plaques and 16% of all subjects having two-feature positive HRP. This is, interestingly, much higher than the frequency of two-feature positive patients observed in earlier studies reporting 4% in established or suspected CAD [5] and 7% in suspected chronic coronary syndrome [24]. The cohort used in the current study was stratified for increased risk. However, even if we calculate the anticipated frequency of two-feature positive plaques in the whole SCAPIS cohort (n = 27,385 performing CCTA) the frequency would be around 4%, which is higher than expected in a general population sample. It is likely that different cut-offs and criteria need to be applied when HRP are studied in populations without previous CAD. These new criteria can be developed when follow-up data is available for the SCAPIS cohort in the near future.

6.3. Correlations to SCORE

In this study, we found a positive correlation between SCORE and total plaque volume. The finding is concordant with previous studies on patients with chest pain that have shown positive correlation between atheroma volume and cardiovascular disease (myocardial infarction) [12,13]. Furthermore, we found a positive correlation between SCORE and LAP volume. The role of LAP as a predictor of cardiac events is well known in symptomatic patient cohorts [4,13,25], and in this study we show that LAP volume associates with other risk predictors of future CAD in a population sample free of previous disease.

6.4. Limitations

In this study, we have oversampled cases with higher SCORE values, which leads to an overrepresentation of male sex. In the SCAPIS cohort, about 50% of all cases have a SCORE value between 0 and 1, i.e., they belong to the lowest stratum. On the other hand, strata 5 and 6 only represents 2.7% each of the SCAPIS cohort. The oversampling of higher SCORE values was performed in order to get a sufficient number of large and more advanced plaques for deep learning model training and to be able to evaluate the annotation workflow and perform analysis of plaque correlations with SCORE.

In order to lower the diversity in the dataset, we only included 100 kV examinations. Another limitation of this study is the lack of uniform definitions of HRP features in the literature. We chose to use easily quantifiable HRP features, that are the most commonly used in the literature and can be derived from semi-quantitative plaque analysis [5,13]. The threshold for low attenuation plaque used in our study was derived in the ROMICAT II study [13], but other definitions have been suggested [10,25] and the optimal threshold might depend on scanner settings and patient population. We used the same cut-off for outward remodelling as in the CAD-RADS definition [23]. Qualitative variables, such as napkin ring sign and spotty calcifications, are more subjective and were not evaluated in this study.

7. Conclusions

In conclusion, our cardiac image analysis laboratory has generated a large annotated CCTA image dataset with high reproducibility and expected correlation to known risk factors which indicates high quality of the data set. The stratified selection of cases resulted in a high prevalence of advanced disease and high risk plaques, which makes the data well suited as training and validation data for a fully automatic analysis tool based on deep learning. The analysis process was time consuming and required large resources, which supports the need of a fully automatic model better suited for analysis of very large datasets such as the SCAPIS cohort.

Author contribution statement

Erika Fagman, Pieter Kitslaar, Michael Kercsik: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Jennifer Alvén, Johan Westerbergh, Ola Hjelmgren: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Kerstin Cederlund, Olov Duvernoy, Hanna Markstad, Ellen Ostenfeld: Conceived and designed the experiments; Analyzed and interpreted the data.

Jan Engvall: Analyzed and interpreted the data.

Isabel Gonçalves: Conceived and designed the experiments.

Göran Bergström: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

Göran Bergström was supported by Hjärtlungfonden {20180324}, Vetenskapsrådet {2019-01140}.

Jennifer Alvén was supported by VINNOVA {254002109}.

Data availability statement

The authors do not have permission to share data.

Sources of funding

This work was made possible thanks to support given by VINNOVA to the "SCAPIS AI platform". SCAPIS is supported by the Swedish Heart and Lung Foundation, the Knut and Alice Wallenberg Foundation, the Swedish Research Council and VINNOVA. The SCAPIS pilot study also received funding from the Sahlgrenska Academy at Gothenburg University and strategic grants from ALF/LUA in Western Sweden. We are grateful to all the participants in this study. A special thanks all test personnel at the SCAPIS test-centre in Göteborg. Work by GB was supported by Heart and Lung foundation (20180324), the Swedish Research Council (2019-01140) and LUA/ALF: ALFGBG-718851.

Declaration of competing interest

Pieter Kitslaar is Employee of Medis Medical Imaging Systems BV. The other authors report no conflicts.

Acknowledgements

The authors wish to thank Wen-Chih Chao, Maria Elisson, Marit Johannesson and Niklas Lundqvist for performing the annotation work in the cardiac image analysis laboratory and Ebba Bergman, Elizabeth Norseng and Nasser Hosseini for data handling and data administration. We further wish to thank Analytic Imaging Diagnostic Arena (AIDA) for technical support, and Medis medical imaging for software development and support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e16058.

References

[1] A. Timmis, P. Vardas, N. Townsend, et al., European Society of Cardiology: cardiovascular disease statistics 2021, Eur Heart J Qual Care Clin Outcomes (2022), https://doi.org/10.1093/ehjqcco/qcac014.
[2] L.J. Shaw, R. Blankstein, J.J. Bax, et al., Society of cardiovascular computed tomography/north American society of cardiovascular imaging - expert consensus document on coronary CT imaging of atherosclerotic plaque, J. Cardiovasc. Comput. Tomogr. 15 (2) (2021) 93–109, https://doi.org/10.1016/j.jcct.2020.11.002.
[3] M.J. Budoff, S. Lakshmanan, P.P. Toth, et al., Cardiac CT angiography in current practice: an American society for preventive cardiology clinical practice statement(.), Am. J. Prev. Cardiol. 9 (2022), 100318, https://doi.org/10.1016/j.ajpc.2022.100318.
[4] M.C. Williams, J. Kwiecinski, M. Doris, et al., Low-attenuation noncalcified plaque on coronary computed tomography angiography predicts myocardial infarction: results from the multicenter SCOT-HEART trial (Scottish computed tomography of the HEART), Circulation 141 (18) (2020) 1452–1462, https://doi.org/10.1161/CIRCULATIONAHA.119.044720.
[5] S. Motoyama, M. Sarai, H. Harigaya, et al., Computed tomographic angiography characteristics of atherosclerotic plaques subsequently resulting in acute coronary syndrome, J. Am. Coll. Cardiol. 54 (1) (2009) 49–57, https://doi.org/10.1016/j.jacc.2009.02.068.
[6] S. Motoyama, H. Ito, M. Sarai, et al., Plaque characterization by coronary computed tomography angiography and the likelihood of acute coronary events in mid-term follow-up, J. Am. Coll. Cardiol. 66 (4) (2015) 337–346, https://doi.org/10.1016/j.jacc.2015.05.069.
[7] M.A. de Graaf, A. Broersen, W. Ahmed, et al., Feasibility of an automated quantitative computed tomography angiography-derived risk score for risk stratification of patients with suspected coronary artery disease, Am. J. Cardiol. 113 (12) (2014) 1947–1955, https://doi.org/10.1016/j.amjcard.2014.03.034 (In eng).
[8] M. Ferencik, C.L. Schlett, B.B. Ghoshhajra, et al., A computed tomography-based coronary lesion score to predict acute coronary syndrome among patients with acute chest pain and significant coronary stenosis on coronary computed tomographic angiogram, Am. J. Cardiol. 110 (2) (2012) 183–189, https://doi.org/10.1016/j.amjcard.2012.02.066.
[9] M. Ferencik, T. Mayrhofer, S.B. Puchner, et al., Computed tomography-based high-risk coronary plaque score to predict acute coronary syndrome among patients with acute chest pain–Results from the ROMICAT II trial, J. Cardiovasc. Comput. Tomogr. 9 (6) (2015) 538–545, https://doi.org/10.1016/j.jcct.2015.07.003.
[10] S.B. Puchner, T. Liu, T. Mayrhofer, et al., High-risk plaque detected on coronary CT angiography predicts acute coronary syndromes independent of significant stenosis in acute chest pain: results from the ROMICAT-II trial, J. Am. Coll. Cardiol. 64 (7) (2014) 684–692, https://doi.org/10.1016/j.jacc.2014.05.039 (In eng).
[11] S. Deseive, R. Straub, M. Kupke, et al., Quantification of coronary low-attenuation plaque volume for long-term prediction of cardiac events and reclassification of patients, J. Cardiovasc. Comput. Tomogr. 12 (2) (2018) 118–124, https://doi.org/10.1016/j.jcct.2018.01.002.
[12] S. Kishi, T.A. Magalhaes, R.J. Cerci, et al., Total coronary atherosclerotic plaque burden assessment by CT angiography for detecting obstructive coronary artery disease associated with myocardial perfusion abnormalities, J. Cardiovasc. Comput. Tomogr. 10 (2) (2016) 121–127, https://doi.org/10.1016/j.jcct.2016.01.005 (In eng).
[13] D.O. Bittner, T. Mayrhofer, S.B. Puchner, et al., Coronary computed tomography angiography-Specific definitions of high-risk plaque features improve detection of acute coronary syndrome, Circ. Cardiovasc. Imaging 11 (8) (2018), e007657, https://doi.org/10.1161/CIRCIMAGING.118.007657.
[14] G. Bergstrom, M. Persson, M. Adiels, et al., Prevalence of subclinical coronary artery atherosclerosis in the general population, Circulation 144 (12) (2021) 916–929, https://doi.org/10.1161/CIRCULATIONAHA.121.055340.
[15] M.F. Piepoli, A.W. Hoes, S. Agewall, et al., European Guidelines on cardiovascular disease prevention in clinical practice, 2016, Eur. Heart J. 37 (29) (2016) 2315–2381, https://doi.org/10.1093/eurheartj/ehw106.
[16] G. Bergstrom, G. Berglund, A. Blomberg, et al., The Swedish CArdioPulmonary BioImage Study: objectives and design, J. Intern. Med. 278 (6) (2015) 645–659, https://doi.org/10.1111/joim.12384.
[17] A.S. Agatston, W.R. Janowitz, F.J. Hildner, N.R. Zusmer, M. Viamonte Jr., R. Detrano, Quantification of coronary artery calcium using ultrafast computed tomography, J. Am. Coll. Cardiol. 15 (4) (1990) 827–832, https://doi.org/10.1016/0735-1097(90)90282-t.

[18] M.A. de Graaf, A. Broersen, P.H. Kitslaar, et al., Automatic quantification and characterization of coronary atherosclerosis with computed tomography coronary angiography: cross-correlation with intravascular ultrasound virtual histology, Int. J. Cardiovasc. Imag. 29 (5) (2013) 1177–1190, https://doi.org/10.1007/s10554-013-0194-x.
[19] H. Brodoefel, A. Reimann, M. Heuschmid, et al., Characterization of coronary atherosclerosis by dual-source computed tomography and HU-based color mapping: a pilot study, Eur. Radiol. 18 (11) (2008) 2466–2474, https://doi.org/10.1007/s00330-008-1019-5.
[20] R.C. Cury, S. Abbara, S. Achenbach, et al., Coronary artery disease - reporting and data system (CAD-RADS): an expert consensus document of SCCT, ACR and NASCI: endorsed by the ACC, JACC Cardiovasc. Imaging 9 (9) (2016) 1099–1113, https://doi.org/10.1016/j.jcmg.2016.05.005.
[21] K. Vaidya, C. Arnott, G.J. Martinez, et al., Colchicine therapy and plaque Stabilization in patients with acute coronary syndrome: a CT coronary angiography study, JACC Cardiovasc. Imaging 11 (2 Pt 2) (2018) 305–316, https://doi.org/10.1016/j.jcmg.2017.08.013.
[22] M.J. Budoff, S.S. Ellenberg, C.E. Lewis, et al., Testosterone treatment and coronary artery plaque volume in older men with low testosterone, JAMA, J. Am. Med. Assoc. 317 (7) (2017) 708–716, https://doi.org/10.1001/jama.2016.21043 (In eng).
[23] R.C. Cury, S. Abbara, S. Achenbach, et al., CAD-RADS(TM) coronary artery disease - reporting and data system. An expert consensus document of the society of cardiovascular computed tomography (SCCT), the American college of radiology (ACR) and the north American society for cardiovascular imaging (NASCI). Endorsed by the American college of cardiology, J. Cardiovasc. Comput. Tomogr. 10 (4) (2016) 269–281, https://doi.org/10.1016/j.jcct.2016.04.005.
[24] E. Conte, A. Annoni, G. Pontone, et al., Evaluation of coronary plaque characteristics with coronary computed tomography angiography in patients with non-obstructive coronary artery disease: a long-term follow-up study, Eur. Heart J. Cardiovasc. Imaging 18 (10) (2017) 1170–1178, https://doi.org/10.1093/ehjci/jew200.
[25] D. Andreini, M. Magnoni, E. Conte, et al., Coronary plaque features on CTA can identify patients at increased risk of cardiovascular events, JACC Cardiovasc. Imaging 13 (8) (2020) 1704–1717, https://doi.org/10.1016/j.jcmg.2019.06.019.