HHS Public Access

Author manuscript

Eur Radiol. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

Eur Radiol. 2018 June ; 28(6): 2655–2664. doi:10.1007/s00330-017-5223-z.

Integrated prediction of lesion-specific ischemia from quantitative coronary CT Angiography using machine learning: a multicenter study

Damini Dey, PhD[a], **Sara Gaur**[b], **Kristian A Ovrehus**[b], **Piotr J Slomka**[c], **Julian Betancur**[c], **Markus Goeller**[a,d], **Michaela M Hell**[d], **Heidi Gransar**[c], **Daniel S Berman**[c], **Stephan Achenbach**[d], **Hans Erik Botker**[b], **Jesper Moller Jensen**[b], **Jens Flensted Lassen**[b], and **Bjarne Linde Norgaard**[b]

[a]Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA

[b]Department of Cardiology, Aarhus University Hospital, Aarhus, Denmark

[c]Departments of Imaging and Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA

[d]Department of Cardiology, Friedrich-Alexander Universitat Erlangen-Nurnberg, Erlangen, Germany

Abstract

Objectives—We aimed to investigate if lesion-specific ischemia by invasive fractional flow reserve (FFR) can be predicted by an integrated machine learning (ML) ischemia risk score from quantitative plaque measures from coronary Computed Tomography Angiography (CTA).

Address for correspondence: Dr. Damini Dey Associate Professor, Biomedical Imaging Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center Taper building, A238, 8700 Beverly Blvd, Los Angeles 90048 Phone: 310-423-1517; Damini.Dey@cshs.org.

Compliance with ethical standards:

Guarantor:

The scientific guarantor of this publication is Dr. Damini Dey (Associate Professor, Cedars-Sinai Medical Center).

Conflict of interest:

The authors of this manuscript declare relationships with the following companies: None.

Dr. Piotr Slomka, Dr. Daniel S Berman and Dr. Damini Dey received software royalties from Cedars-Sinai Medical Center and hold a patent.

Statistics and biometry:

One of the co-authors (Heidi Gransar, MS) is an experienced biostatistician and she provided her expertise for this study.

Informed consent:

Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval:

Institutional Review Board approval was obtained.

Study subjects or cohorts overlap:

This study is a new post hoc analysis comprising all 254 patients from the prospective, multicenter NXT trial (NCT01757678). The original study has been previously described (Norgaard et al J Am Coll Cardiology 2014, Reference 17). While the patient cohort is the same, there was no overlap with this study which evaluates objective machine learning integration of quantitative coronary CT Angiography to predict ischemia.

Methodology:

- prospective study
- post-hoc analysis
- multicentre study

Methods—In a multicenter trial of 254 patients, CTA and invasive coronary angiography were performed, with FFR in 484 vessels. CTA datasets were analyzed by semi-automated software to quantify stenosis and non-calcified (NCP), low-density NCP (LD-NCP, <30 HU), calcified and total plaque volumes, contrast density difference (CDD, maximum difference in luminal attenuation per unit area), and plaque length. ML integration included automated feature selection and model building from quantitative CTA with a boosted ensemble algorithm, and ten-fold stratified cross-validation.

Results—Eighty patients had ischemia by FFR (FFR≤0.80) in 100 vessels. Information gain for predicting ischemia was highest for CDD (0.172), followed by LD-NCP (0.125), NCP (0.097), and total plaque volumes (0.092). ML exhibited higher area-under-the-curve (0.84) than individual CTA measures, including stenosis (0.76), LD-NCP volume (0.77), total plaque volume (0.74), and pre-test likelihood of coronary artery disease (CAD) (0.63); p<0.006.

Conclusions—Integrated ML ischemia risk score improved the prediction of lesion-specific ischemia by invasive FFR, over stenosis, plaque measures and pre-test likelihood of CAD.

Keywords

Computed Tomography Angiography; atherosclerotic plaque; coronary stenosis; machine learning; ischemia

Introduction

Coronary Computed Tomography Angiography (CTA) is a noninvasive diagnostic test which allows direct noninvasive assessment of the coronary arteries[1–5]. In addition to luminal stenosis, CTA also permits noninvasive assessment of atherosclerotic plaque features as well as quantitative measurement of plaque burden [6–10].

Revascularization strategies guided by lesion-specific ischemia, assessed by invasive Fractional Flow Reserve (FFR), have been shown to improve patient outcomes[11; 12]. Selected plaque features from CTA, such as low-density non-calcified plaque (NCP) and total plaque burden, have been shown to noninvasively improve detection of lesion-specific ischemia[13; 14]. It is not known, however, whether several *correlated* quantitative plaque measures from CTA can be effectively combined to predict lesion-specific ischemia.

Machine learning is a field of computer science that uses computer algorithms to identify patterns in large datasets with a multitude of variables without making any prior assumptions. Accordingly, machine learning techniques have emerged as highly effective methods for prediction and intelligent decision-making in many areas of everyday living [15; 16]. Recently machine learning applied to medical imaging has shown to improve diagnostic accuracy [17; 18] and prognostic outcomes [19; 20]. The aim of this study was to investigate if lesion-specific ischemia by FFR can be effectively predicted by machine learning integration of quantitative plaque metrics measured from CTA.

Methods

Patients

This was a post hoc sub-study comprising all 254 patients from the prospective, multicenter NXT trial (NCT01757678), which has been previously described in detail [21]. This trial included patients suspected of stable coronary artery disease who underwent coronary CTA ≤ 60 days prior to clinically indicated, non-emergent invasive coronary angiography with fractional flow measurements. Exclusion criteria included prior stent implantation or coronary bypass surgery, contraindications to beta-blockers, nitrates or adenosine, suspicion of acute coronary syndrome, significant arrhythmia, and body mass index >35 kg/m^2 [21]. As described by Norgaard et al [21], the overall image quality in these 254 patients were good-to-excellent. All patients were included for quantitative plaque analysis. Age, gender and angina typicality for each patient were used to determine pre-test likelihood of coronary artery disease [22]. Institutional review committees approved the study protocol and all patients provided informed written consent.

Invasive coronary angiography and fractional flow reserve measurements

Angiography and FFR were performed according to standard practice [21]. The FFR pressure-wire was positioned a minimum distance of 20 mm distal to vessel stenosis; hyperaemia was induced by intravenous adenosine. Lesion-specific ischemia was defined by FFR ≤ 0.80.

Coronary computed tomography acquisition

Coronary CTA was performed using single or dual-source CT scanners with at least 64 detector rows in accordance with societal guidelines[23]. Oral and/or intravenous beta-blockade was administered targeting a heart rate < 60 beats/min and sublingual nitrates were administered in all patients [21]. Both prospective and retrospective gating were used for scan acquisitions. Data acquisition was performed with 100-kV tube voltage in patients weighing ≤ 70 kg and 120-kV otherwise. The mean DLP was (214.3 ± 157.1) mGy.cm for prospective and (1021.4 ± 500.0) mGy.cm for retrospective acquisitions.

Plaque analysis from coronary CT Angiography

Plaque analysis in coronary segments with distal lumen diameter ≥ 2 mm was performed retrospectively with semi-automated software (AutoPlaque research software, Cedars-Sinai Medical Center, Los Angeles, CA, USA), by one of two experienced readers (both Level 3 certified in cardiac CT, with 5 years or longer experience) blinded to patient characteristics, coronary CTA readings, and FFR[24; 25]. For each patient, the Segmental Involvement Score was defined as the total number of coronary segments with any coronary plaque. Scan-specific attenuation thresholds were computed for non-calcified and calcified plaque and lumen as previously described [24; 25]. Low-density non-calcified plaque was defined as non-calcified plaque below a preset low-density threshold (30 HU) [26]. The plaque volumes and burden (defined as plaque volume x 100%/vessel volume) for each plaque component was calculated. Quantitative percent stenosis was calculated by dividing the narrowest lumen diameter by the mean of two normal non-diseased reference points.

Aggregate plaque volume was computed as (total plaque volume x 100%/vessel volume). Remodeling index was determined as the ratio of maximum vessel area to that at the proximal normal reference point[7]. Plaque length (in mm) was the length of the diseased vessel segments. Contrast density difference over the lesion was computed as follows: the luminal contrast density, defined as mean luminal attenuation per unit area was computed automatically over 1 mm cross-sections of the arterial segment. The contrast density difference was defined as the maximum percent difference in contrast densities, relative to the proximal reference cross-section (with no disease)[27]. For each artery, maximum diameter stenosis, continuous remodeling index and contrast density difference values were reported along with plaque measurements. The CTA images were evaluated by one expert reader with and without plaque overlay, using transverse and MPR views [25; 28]. If needed, minor edits were made using the automated vessel wall correction options [24; 25]. There was good interobserver reproducibility of plaque characteristics between the 2 expert readers in this study as assessed by Bland-Altman analysis in a consecutive selection of 10% of patients, as reported previously by Gaur et al [14] [Supplementary material, Figure S1]. Figure 1 shows an example of plaque analysis.

Machine learning integration of plaque features

Quantitative CTA plaque measures, along with age, gender and number of risk factors, were imported to the Waikato Environment for Knowledge Analysis (WEKA) environment where all the steps for machine learning analysis were performed[29]. Figure 2 shows an overview of the methods for machine learning integration. Machine learning integration involved automated feature selection by information gain ranking, model building with a boosted ensemble algorithm, and ten-fold stratified cross-validation for the entire method [19]. These methods have been effectively used in our other studies with cardiac imaging data [19; 30].

Feature selection—Feature selection was performed on all patient and quantitative CTA measures using a method known as ‘information gain attribute ranking’[31] [19], using stratified ten-fold cross-validation. Continuous variables were discretized by minimum description length method using information entropy minimization, as described in Fayyad and Irani 1993, with default options in the Weka environment. The number of bins were automatically determined by the minimum distance length, as the minimum number of bins to uniquely specify the variable [32]. Information gain is defined as a measure of the effectiveness of a covariate in classifying an outcome. It is measured as the amount by which the entropy of the outcome decreases, which reflects the additional information provided by the covariate. The resulting information gain ranged as continuous values from 0 to 1. Only attributes resulting in information gain >0.001 were subsequently used in machine learning.

Machine learning integration—Prediction of lesion-specific ischemia was performed by an ensemble classification approach (“boosting”), employing the iterative, additive LogitBoost algorithm using decision stumps (single-node decision trees) for each feature-selected parameter[29; 33]. This method, an example of supervised ensemble learning [34] has already been shown to be efficient in predicting revascularization following ischemia testing [30] and for prediction of all-cause mortality following coronary CTA in a multicenter international registry [19]. The principle behind machine learning ensemble

boosting is that a set of weak base classifiers can be combined to create a single strong classifier by iteratively and automatically adjusting their appropriate weighting according to misclassifications. A series of base classifier predictions and an updated weighting distribution are produced with each iteration. These predictions are then combined by weighted majority voting to derive an overall classifier or risk score ranging from 0 to 1.

Cross-validation—The entire machine learning process (feature selection and LogitBoost) was conducted using stratified ten-fold cross-validation, which is currently the preferred technique in data mining[19; 35], for robust performance. Briefly, the dataset was first randomly divided into 10 equally sized subsamples, each with the same number of events. In ten-fold cross validation, training was done in $9/10^{th}$ of the data and tested on the remaining unseen $1/10^{th}$ of the data. Machine learning was thus consistently evaluated on previously unseen data; the evaluation was not performed on the same data as model building. Ten-fold cross validation has been shown to have smaller bias for discriminant analysis than traditional split-sample approach (test and validation)[35]. In our implementation of the 10-fold cross-validation, the arteries were also always grouped at the patient level, thus in each-fold of the cross-validation the data in the validation and the training sets strictly belonged to different patients. Such cross-validation is standard in machine learning [19] and is done to obtain a more accurate, unbiased estimate of the diagnostic performance of the model, and further, to mirror test of the application in everyday clinical practice, where the method is tested in new patients only.

The machine learning ischemia risk score is computed directly by concatenating the probability estimates obtained in these 10-folds from unseen data and corresponded to the whole cohort. This ischemia risk score (with continuous values from 0 to 1) was saved, and further statistical analyses were performed using Analyse-it software (Analyse-it, Leeds, UK). Using the same data folds, comparison of Receiver Operator Characteristic (ROC) area under the curve (AUC) was used to compare the predictive performance of machine learning, logistic regression, plaque measures and pre-test likelihood of coronary artery disease, based on the method of DeLong et al [36]. For the ischemia risk score, the accuracy was determined at a threshold at which Youden's J-statistic ($J = sensitivity + specificity - 1$) was the highest.

Statistical Analysis

Statistical analyses were performed with STATA software (version 11 StataCorp LP) and Analyse-it software (Analyse-it, Leeds, UK). For all continuous variables, the Shapiro-Wilk test was first used to assess normality. All continuous variables were described as mean ± standard deviation, or median and interquartile range (IQR). Wilcoxon rank-sum test or two-sample t-tests were used to compare groups regarding continuous variables, while Pearson Chi-square or Fisher exact test were used to compare groups regarding categorical variables. All tests were two-sided. We also evaluated the Integrated Discrimination Improvement (IDI) index, which complements AUC and evaluates the improvement in model discrimination over a set of existing predictors. IDI was used since established risk categories are not available for the integrated ischemia risk score or quantitative plaque features [37]. IDI analyses were performed with a SAS 9.2 software module. We also

compared the machine learning ischemia risk score with conventional statistical logistic regression performed using a *Stata* version 11 module, using the same data, 10-fold cross-validation and the same data folds. A p-value of <0.05 was considered statistically significant.

Results

Median patient age was 64 years (range 31–84 years) and 64% was male, with mean body mass index 26±3 kg/m^2. The heart rate during CTA was 63±10 beats/min (range 37–110 beats/min). The demographic and clinical risk factors of the patient population are listed in Table 1. Eighty out of 254 patients (31%) had lesion-specific ischemia, in a total of 100 vessels. Except a higher prevalence of male gender (p<0.001), we identified no difference in any patient characteristics in the impaired versus normal FFR group. In total, 2758 coronary artery segments were analyzed in 254 patients. Figure 1 shows a case example from our study for a patient with impaired FFR, in the right coronary artery.

Information gain ranking

Figure 2 shows an overview of the machine learning integration methods. Figure 3 shows the results of information gain ranking with 10-fold cross-validation. Among quantitative CTA measures, CDD exhibited the highest information gain (0.172). Among quantitative plaque measures, LD-NCP volume had the highest information gain (0.125), followed by NCP volume (0.097) and plaque length (0.092); information gain for CP volume and CP burden were <0.001. In terms of information gain, the contribution of total plaque length and total plaque volume was similar (0.094 and 0.092). Table 2 shows the quantitative plaque measures in ischemic and non-ischemic vessels. Ischemic vessels were marked by greater stenosis, CDD, plaque volumes (in particular low-density NCP and NCP plaque volumes), as well as greater plaque length. While CP volume reached significance in Table 2 (p=0.01), there was significant overlap for inter-quartile range for non-ischemic (0–25.95 mm^3) and ischemic vessels (2.1–46.2 mm^3). The AUC for CP volume for the prediction of ischemia was 0.58 (95% CI: 0.52–0.64), which was significantly lower than all other quantitative CTA plaque or stenosis measures. In multivariable logistic regression with low-density NCP volume, NCP volume and stenosis, CP volume was not significant (p=0.332).

Integrated ischemia risk score by machine learning

Figure 4 shows the ROC curves for plaque measures and the integrated machine learning score from combining quantitative CTA measures (after 10-fold cross-validation). As seen in Figure 4, ML had the highest area-under-the-curve (AUC): 0.84(95% CI:0.79–0.88). There was a significant increase in AUC for the integrated score when compared to individual quantitative CTA metrics, for example, LD-NCP volume [0.77 (95% CI:0.71–0.82), p=0.002], NCP volume [0.75 (95% CI:0.69–0.80), p=0.0001] and total plaque volume [0.74 (95% CI:0.69–0.79), p<0.0001]. The integrated score also had higher AUC than total plaque burden and LD-NCP plaque burden 0.61 (95% CI:0.55–0.67), 0.69 (95% CI:0.65–0.75), p<0.0001, as well as plaque Segment Involvement Score [0.65 (95% CI:0.59–071), p<0.0001]. Notably, AUC for LD-NCP plaque volume was significantly higher than for total plaque volume (p=0.01). The integrated score also had significantly higher AUC than

quantitative stenosis or minimum luminal area [0.84 (95% CI:0.79–0.88) vs 0.76 (95% CI: 0.72–0.82) for quantitative stenosis, p=0.002 (Figure 5); and 0.73 for minimum luminal area (95% CI:0.68–0.78), p<0.0001]. Figure 5 shows that quantitative stenosis had significantly higher AUC than pre-test likelihood [0.63(95% CI:0.57-–.70)], and the machine learning score, in turn, had a higher AUC than quantitative stenosis (p=0.005) or pretest-likelihood (p<0.0001). The processing time of machine learning (automated feature selection and integration of quantitative coronary CTA with 10-fold cross-validation) was less than 30 seconds on a standard windows laptop computer with a i7-46000 CPU @ 2.1 GHZ processor.

At a threshold of 0.24, determined by Youden's J-statistic, the accuracy, sensitivity, and specificity were 80%, 73% and 80% respectively, for discrimination of lesion-specific ischemia. On a per patient basis, AUC for the integrated score was 0.82 (95% CI:0.76–0.88). We compared the ischemia risk score to conventional logistic regression using 10-fold cross-validation and the same data folds; the ischemia risk score had significantly higher AUC than conventional logistic regression [0.84 (95% CI:0.79–0.88) vs 0.78(95% CI:0.72–0.83), p=0.02].

Discrimination improvement for ischemia—Table 3 shows the discrimination improvement for the integrated ischemia risk score compared to CTA plaque measures. As listed in Table 3, we found significant improvement in IDI over CDD (0.13) as well as individual CTA plaque metrics (0.16 for LD-NCP volume, 0.20 for NCP volume and 0.22 for total plaque volume), p<0.0001 for all.

Table 4 shows the Spearman rank correlation for the quantitative plaque measures with continuous FFR values. CDD, quantitative stenosis, plaque volumes and plaque length showed significant moderate negative correlation with continuous FFR, and minimum luminal area showed significant positive correlation.

Discussion

In this multicenter study, our results show that objective combination of clinical data and quantitative CTA plaque measures by machine learning improves the prediction of hemodynamic significance of lesions as measured by invasive FFR. We show that the performance of the integrated machine learning score was superior to that of visually assessed CTA metrics as well as individual quantitative measures. Our results also indicate that prediction of lesion-specific ischemia by the ischemia risk score was significantly higher compared to a standard statistical logistic regression using the same data. Thus, our results suggest that state-of-the-art data mining approaches can potentially outperform conventional statistical methods. It has been shown that plaque assessment from CTA adds significantly to stenosis assessment for the identification of lesion-specific ischemia[38; 39]. To correctly identify the quantitative CTA measures which best predict lesion-specific ischemia, however, is a challenge, since individual quantitative plaque measures correlate strongly, and there is often significant interaction between indvidual metrics. The complexity of risk stratification is further increased since all the clinical factors affecting risk need to be considered. In this study, we show that machine learning integration is able to

overcome these challenges. Our results further show that quantitative stenosis adds incremental value to standard pre-test likelihood of coronary artery disease for the prediction of ischemia; and further, the integrated machine learning score adds to quantitative stenosis and individual plaque measures, and pretest likelihood.

Among individual quantitative measures, we also found that CDD, a relatively new CTA metric, had the highest information gain for lesion-specific ischemia and showed significant negative correlation with FFR. This is concordant with the previous findings of Hell et al, who showed improved discrimination of lesion-specific ischemia by CDD over measured transluminal attenuation gradient in a completely separate cohort of 59 consecutive patients [27]. CDD, by definition, is the normalized difference in lumen attenuation per unit area over a segment, and is closely related to both minimum luminal area and luminal attenuation gradient; a higher CDD implies a lower minimum luminal area; thus it includes the contribution of both quantitative measures. Among plaque measures, we found LD-NCP volume was the highest ranked feature, followed by NCP volume; both showed significant negative correlation with FFR. This finding is in line with previous studies investigating the relationship of adverse plaque features and FFR [38; 39].

Several other investigators have examined the relationship between quantitative plaque characteristics and lesion-specific ischemia [27; 38; 39]. Ko et al have also developed the ASLA score combining several pre-defined plaque features and the APPROACH score into a composite score to predict ischemia[40]. The performance of noninvasive FFR, derived from standard CTA based on computational fluid dynamics-based methods from CTA (FFR_{CT}), have been reported to be superior to CTA anatomical interpretation in prospective multicenter studies [21; 41; 42]. Onsite FFR_{CT} using a reduced-order computational fluid dynamics computation on a separate workstation has also been reported [43]. Recent studies have also investigated the performance of an onsite machine-learning based technique (research software from *Siemens Heathineers*) to identify lesion-specific ischemia by invasive FFR. The machine-learning method was trained on a large database of synthetically generated coronary models and found to be equivalent to an onsite computational fluid dynamics-based algorithm, with shorter execution times [44]. CT perfusion has also been shown to effectively predict the hemodynamic significance of coronary stenoses[45]. CT perfusion requires an additional stress CT scan, with attendant radiation exposure and need for pharmacologic stress.

Our study describes a simpler approach, which automatically ranks and then combines coronary plaque measures on a standard personal computer from the resting CTA scan to predict the functional significance of coronary stenoses, provided vessel-based plaque measurement is available. Our study also demonstrates the contribution of a new measure, CDD a normalized measure of luminal contrast kinetics from quantitative CTA to identify lesion-specific ischemia, in a multicenter study. In practice, such a ischemia risk score may help to objectively stratify patients as low or high risk following CTA, and indicate which patient may benefit from further noninvasive ischemia testing, noninvasive FFR, or FFR assessment.

In the current study we have only been able to discuss the feasibility and performance of machine learning, but not its prospective practical implementation. In the near future, we envisage machine learning working in the background of standard plaque analysis software, gathering the variables automatically and allowing on-the-fly risk score computation. This principle is already utilized daily by several applications. For example, the personalized advertisements that appear in real-time whilst web-browsing are all based on the passive collection of variables and their seamless input into machine learning algorithms. Despite this study limitation, however, an integrated machine learning ischemia risk score, as we show in this study, can be derived automatically following plaque analysis, from standard rest CTA images on a standard personal computer without additional imaging, radiation exposure or cost. Such objective integration of coronary plaque features directly measured from the CTA scan of an individual patient can predict the hemodynamic significance of coronary stenosis from standard CTA scans – either onsite or in the cloud, increasing the accuracy of standard CTA for the prediction of lesion-specific ischemia.

We used automated feature selection, which distinguishes itself from traditional approaches by making no prior assumptions about the data. Our results further suggest that machine learning approach may outperform conventional statistical integration of the same data. Any given machine learning algorithm requires only minimal input during model-building, and none after that. This feature of machine learning is important, since new data can be easily incorporated to continuously update and optimize prediction. In the future, we would expect big data and machine learning algorithms to improve diagnostic accuracy and help the imaging physician find the right answer for the patients, whose “lives and medical histories shape the algorithms” [46].

There were several limitations in our study. Although 10-fold cross-validation was used for validation, our study lacked an external validation cohort. Plaque findings were not confirmed by invasive intravascular ultrasound; however, plaque measurement by coronary CTA has been previously shown to strongly correlate with intravascular ultrasound[25].

Conclusions

The integrated ischemia risk score based on machine learning integration of quantitative plaque measures from CTA improved the prediction of lesion-specific ischemia by invasive FFR, over stenosis, plaque measures and pre-test likelihood of coronary artery disease.

Acknowledgments

Funding:

This study has received funding by National Heart Lung Blood Institute (NIH/NHLBI grant R01HL133616) and the Bundesministerium für Bildung und Forschung (01EX1012B, Spitzencluster Medical Valley)

List of abbreviations and acronyms

CTA	CT Angiography
FFR	Fractional Flow Reserve

DLP	Dose-length Product
HU	Hounsfield Unit
CDD	Contrast Density Difference
NCP	Noncalcified plaque
LD-NCP	Low-density noncalcified plaque
CP	Calcified plaque

References

1. Budoff MJ, Dowe D, Jollis JG, et al. Diagnostic Performance of 64-Multidetector Row Coronary Computed Tomographic Angiography for Evaluation of Coronary Artery Stenosis in Individuals Without Known Coronary Artery Disease: Results From the Prospective Multicenter ACCURACY (Assessment by Coronary Computed Tomographic Angiography of Individuals Undergoing Invasive Coronary Angiography) Trial. Journal of the American College of Cardiology. 2008; 52:1724–1732. [PubMed: 19007693]
2. Hausleiter J, Meyer T, Hadamitzky M, et al. Non-invasive coronary computed tomographic angiography for patients with suspected coronary artery disease: the Coronary Angiography by Computed Tomography with the Use of a Submillimeter resolution (CACTUS) trial. European Heart Journal. 2007; 28:3034–3041. [PubMed: 17540851]
3. Achenbach S, Ropers U, Kuettner A, et al. Randomized Comparison of 64-Slice Single- and Dual-Source Computed Tomography Coronary Angiography for the Detection of Coronary Artery Disease. J Am Coll Cardiol Cardiovascular Imaging. 2008; 1:177–186.
4. Miller JM, Rochitte CE, Dewey M, et al. Diagnostic Performance of Coronary Angiography by 64-Row CT. N Engl J Med. 2008; 359:2324–2336. [PubMed: 19038879]
5. Meijboom WB, van Mieghem CAG, Mollet NR, et al. 64-Slice Computed Tomography Coronary Angiography in Patients With High, Intermediate, or Low Pretest Probability of Significant Coronary Artery Disease. J Am Coll Cardiol. 2007; 50:1469–1475. [PubMed: 17919567]
6. Achenbach S, Moselewski F, Ropers D, et al. Detection of calcified and noncalcified coronary atherosclerotic plaque by contrast-enhanced, submillimeter multidetector spiral computed tomography: a segment-based comparison with intravascular ultrasound. Circulation. 2004; 109:14–17. [PubMed: 14691045]
7. Achenbach S, Ropers D, Hoffmann U, et al. Assessment of coronary remodeling in stenotic and nonstenotic coronary atherosclerotic lesions by multidetector spiral computed tomography. J Am Coll Cardiol. 2004; 43:842–847. [PubMed: 14998627]
8. Leber AW, Becker A, Knez A, et al. Accuracy of 64-slice computed tomography to classify and quantify plaque volumes in the proximal coronary system: a comparative study using intravascular ultrasound. J Am Coll Cardiol. 2006; 47:672–677. [PubMed: 16458154]
9. Leber AW, Knez A, Becker A, et al. Accuracy of multidetector spiral computed tomography in identifying and differentiating the composition of coronary atherosclerotic plaques: a comparative study with intracoronary ultrasound. J Am Coll Cardiol. 2004; 43:1241–1247. [PubMed: 15063437]
10. Pundziute G, Schuijf JD, Jukema JW, et al. Evaluation of plaque characteristics in acute coronary syndromes: non-invasive assessment with multi-slice computed tomography and invasive evaluation with intravascular ultrasound radiofrequency data analysis. European Heart Journal. 2008; 29:2373–2381. [PubMed: 18682447]
11. Tonino PA, De Bruyne B, Pijls NH, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. N Engl J Med. 2009; 360:213–224. [PubMed: 19144937]
12. Pijls NH, Fearon WF, Tonino PA, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention in patients with multivessel coronary artery disease: 2-year follow-up of the FAME (Fractional Flow Reserve Versus Angiography for Multivessel Evaluation) study. J Am Coll Cardiol. 2010; 56:177–184. [PubMed: 20537493]

13. Park HB, Heo R, o Hartaigh B, et al. Atherosclerotic plaque characteristics by CT angiography identify coronary lesions that cause ischemia: a direct comparison to fractional flow reserve. JACC Cardiovasc Imaging. 2015; 8:1–10. [PubMed: 25592691]
14. Gaur S, Øvrehus KA, Dey D, et al. Coronary plaque quantification and fractional flow reserve by coronary computed tomography angiography identify ischaemia-causing lesions. European Heart Journal. 2016; doi: 10.1093/eurheartj/ehv690
15. Deo RC. Machine Learning in Medicine. Circulation. 2015; 132:1920–1930. [PubMed: 26572668]
16. Waljee AK, Higgins PD. Machine learning in medicine: a primer for physicians. Am J Gastroenterol. 2010; 105:1224–1226. [PubMed: 20523307]
17. Arsanjani R, Xu Y, Dey D, et al. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. J Nucl Cardiol. 2013; 20:553–562. [PubMed: 23703378]
18. Kang D, Slomka PJ, Nakazato R, et al. Automated knowledge-based detection of nonobstructive and obstructive arterial lesions from coronary CT angiography. Med Phys. 2013; 40:041912. [PubMed: 23556906]
19. Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J. 2017; 38:500–507. [PubMed: 27252451]
20. Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. Circ Res. 2017; 9:311312.
21. Norgaard BL, Leipsic J, Gaur S, et al. Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease: the NXT trial (Analysis of Coronary Blood Flow Using CT Angiography: Next Steps). J Am Coll Cardiol. 2014; 63:1145–1155. [PubMed: 24486266]
22. Montalescot G, Sechtem U, Achenbach S, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J. 2013; 34:2949–3003. [PubMed: 23996286]
23. Leipsic J, Abbara S, Achenbach S, et al. SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee. J Cardiovasc Comput Tomogr. 2014; 8:342–358. [PubMed: 25301040]
24. Dey D, Cheng VY, Slomka PJ, Nakazato R, Ramesh A, Gurudevan S, Germano Berman DS. Automated Three-dimensional Quantification of Non-calcified and Calcified Coronary Plaque from Coronary CT Angiography. Journal of Cardiovascular Computed Tomography. 2009:372–382. [PubMed: 20083056]
25. Dey D, Schepis T, Marwan M, Slomka PJ, Berman DS, Achenbach S. Automated Three-dimensional Quantification of Non-calcified Coronary Plaque from Coronary CT Angiography: comparison with Intravascular Ultrasound Radiology. 2010; 257:516–522.
26. Motoyama S, Sarai M, Harigaya H, et al. Computed Tomographic Angiography Characteristics of Atherosclerotic Plaques Subsequently Resulting in Acute Coronary Syndrome. Journal of the American College of Cardiology. 2009; 54:49–57. [PubMed: 19555840]
27. Hell MM, Dey D, Marwan M, Achenbach S, Schmid J, Schuhbaeck A. Non-invasive Prediction of Hemodynamically Significant Coronary Artery Stenoses by Contrast Density Difference in Coronary CT Angiography. European Journal of Radiology. 2015; 84:1502–1508. [PubMed: 26001435]
28. Schuhbaeck A, Dey D, Otaki Y, et al. Interscan Reproducibility of Quantitative Coronary Plaque Volume and Composition from CT Coronary Angiography using an automated method. Journal of the American College of Cardiology. 2013; 61doi: 10.1016/S0735-1097(1013)61040-61042
29. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. SIGKDD Explor Newsl. 2009; 11:10–18.
30. Arsanjani R, Dey D, Khachatryan T, et al. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. J Nucl Cardiol. 2015; 22:877–884. [PubMed: 25480110]
31. Hall MA, Holmes G. Benchmarking attribute selection techniques for discrete class data mining. IEEE Transactions on Knowledge and Data Engineering. 2003; 15:1437–1447.

32. Fayyad, UM., Irani Keki, B. Multi-interval discretization of continuous valued attributes for classification learning. Thirteenth International Joint Conference on Artificial Intelligence; 1993. p. 1022-1027.
33. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting. Annals of Statistics. 2000; 28:337–407.
34. Witten, IH., Frank, E., Hall, MA. Data Mining: Practical Machine Learning Tools and Techniques. 3. Morgan Kaufmann Publishers; Burlington, MA: 2011.
35. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinformatics. 2005; 21:3301–3307. [PubMed: 15905277]
36. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics. 2007; 44:837–845.
37. Pencina MJ, D'Agostino RB, Pencina KM, Janssens ACJW, Greenland P. Interpreting Incremental Value of Markers Added to Risk Prediction Models. American Journal of Epidemiology. 2012; 176:473–481. [PubMed: 22875755]
38. Park H-B, Heo R, ó Hartaigh B, et al. Atherosclerotic Plaque Characteristics by CT Angiography Identify Coronary Lesions That Cause IschemiaA Direct Comparison to Fractional Flow Reserve. JACC: Cardiovascular Imaging. 2015; 8:1–10. [PubMed: 25592691]
39. Diaz Zamudio M, Dey D, Schuhbaeck A, Nakazato R, Slomka PJ, Bermand DS, Achenbach S, Min JK, Doh JH, Koo BK. Automated quantitative plaque burden from coronary CT Angiography noninvasively predicts hemodynamic significance by fractional flow reserve in intermediate coronary lesions. Radiology. 2015 Aug; 276(272):408–215. http://dx.doi.org/10.1148/radiol.2015141648. [PubMed: 25897475]
40. Ko BS, Wong DT, Cameron JD, et al. The ASLA Score: A CT Angiographic Index to Predict Functionally Significant Coronary Stenoses in Lesions with Intermediate Severity-Diagnostic Accuracy. Radiology. 2015; 276:91–101. [PubMed: 25710278]
41. Min JK, Leipsic J, Pencina MJ, Berman DS, Koo BK, Mieghem CV, Erglis A, Lin F, Dunning AM, Apruzzese P, Budoff MJ, Cole JH, Jaffer FA, Leon MB, Malpeso J, Mancini GBJ, Park SJ, Schwartz RS, Shaw LS, Mauri L. Diagnostic accuracy of fractional flow reserve from anatomic CT angiography. Jama. 2012; 308:1237–1245. [PubMed: 22922562]
42. Koo BK, Erglis A, Doh JH, et al. Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms. Results from the prospective multicenter DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) study. J Am Coll Cardiol. 2011; 58:1989–1997. [PubMed: 22032711]
43. Coenen A, Lubbers MM, Kurata A, et al. Fractional flow reserve computed from noninvasive CT angiography data: diagnostic performance of an on-site clinician-operated computational fluid dynamics algorithm. Radiology. 2015; 274:674–683. [PubMed: 25322342]
44. Itu L, Rapaka S, Passerini T, et al. A machine-learning approach for computation of fractional flow reserve from coronary computed tomography. J Appl Physiol (1985). 2016; 121:42–52. [PubMed: 27079692]
45. Ko BS, Cameron JD, Leung M, et al. Combined CT coronary angiography and stress myocardial perfusion imaging for hemodynamically significant stenoses in patients with suspected coronary artery disease: a comparison with fractional flow reserve. JACC Cardiovasc Imaging. 2012; 5:1097–1111. [PubMed: 23153909]
46. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016; 375:1216–1219. [PubMed: 27682033]

Key points

- Integrated ischemia risk score improved prediction of ischemia over quantitative plaque measures
- Integrated ischemia risk score showed higher prediction of ischemia than standard approach
- Contrast density difference had the highest information gain to identify lesion-specific ischemia

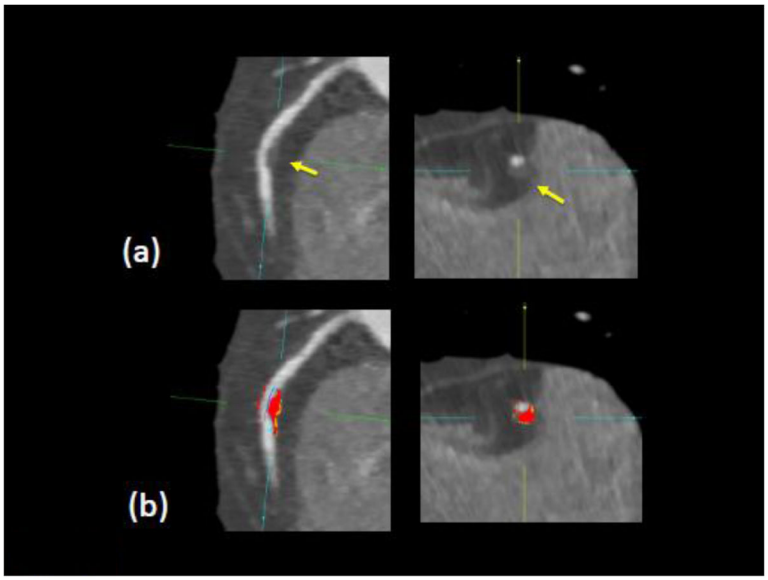

Figure 1.

Example of plaque quantification for RCA lesion in one patient in our study. Arrows show noncalcified plaque in longitudinal and cross sectional views. Red and orange shows measured NCP voxels (orange overlay indicates LD-NCP component). Measured FFR was 0.80 in this vessel.

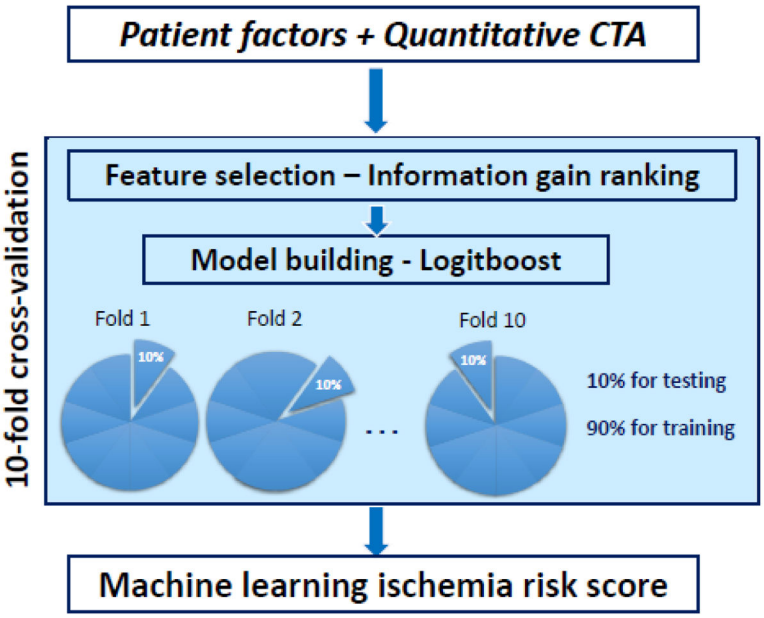

Figure 2.
Overview of the machine learning integration method

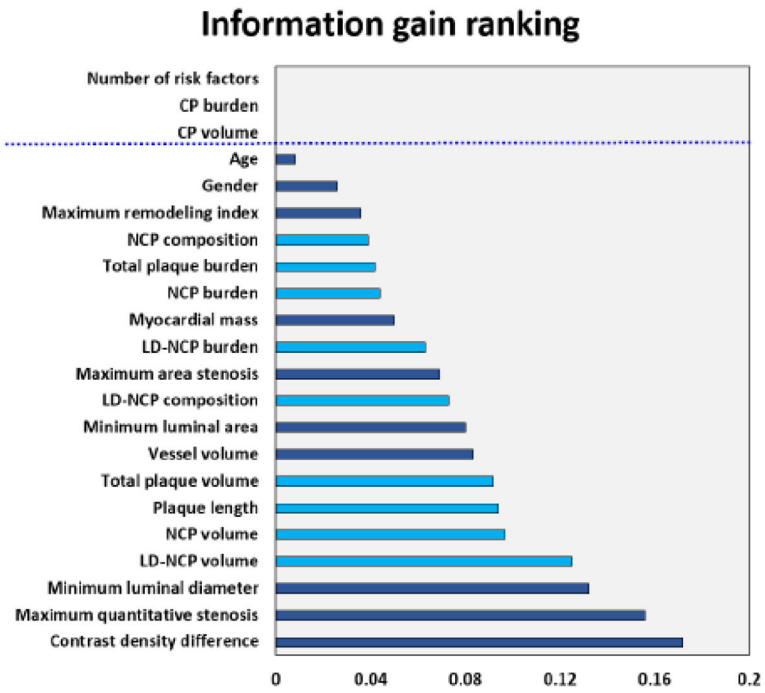

Figure 3.

Figure showing the information gain for age, gender and quantitative CTA measures. Measures directly related to plaque volume are in light blue and the remaining measures in dark blue. Variables with information gain > 0.001 were used in machine learning. Contrast density difference was associated with the highest information gain among all the plaque parameters.

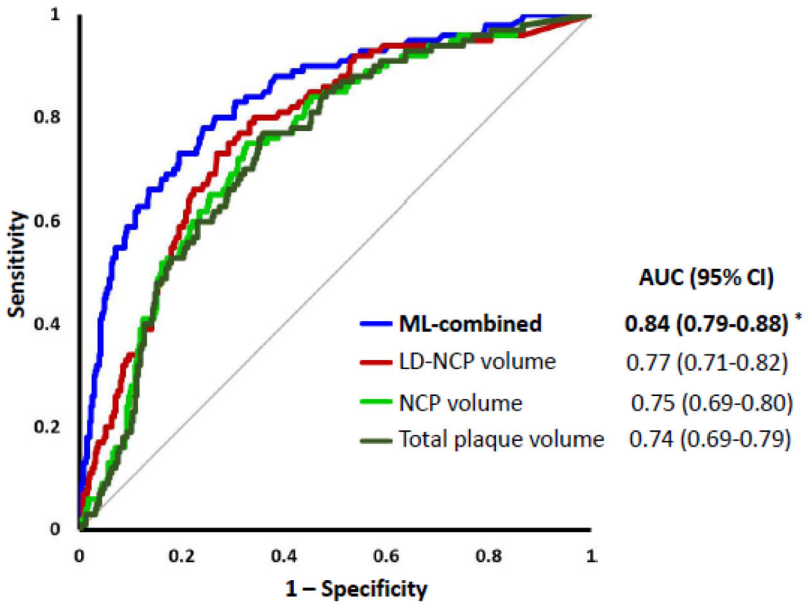

Figure 4.

Prediction of lesion-specific ischemia by the integrated ischemia risk score by machine learning (ML-combined) and quantitative plaque volumes (LD-NCP, NCP and total plaque volume). ML-combined had a significantly higher AUC compared to individual quantitative CTA plaque volumes (*p<0.003 compared to LD-NCP, NCP and total plaque volume). AUC for LD-NCP plaque volume was significantly higher than for total plaque volume (p=0.01)

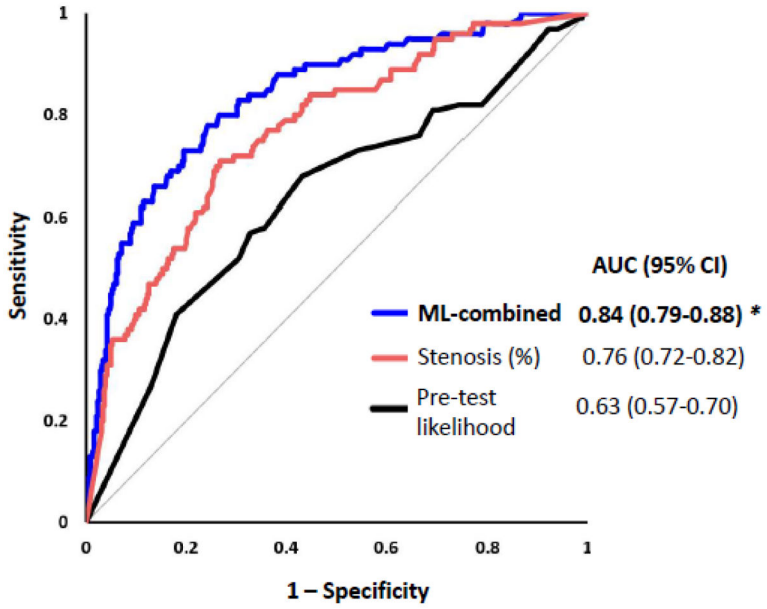

Figure 5.

Prediction of lesion-specific ischemia by the integrated ischemia risk score by machine learning (ML-combined), quantitative stenosis and pre-test likelihood of coronary artery disease. Quantitative stenosis had significantly higher AUC than pre-test likelihood (p=0.0005). ML-combined(*) had a significantly higher AUC than quantitative stenosis (p=0.005) or pretest-likelihood of coronary artery disease (p<0.0001).

Table 1

Patient Characteristics (254 patients)

Age (years)	64±10
Male (%)	162 (64)
Risk factors	
Hypertension (%)	174 (69)
Hyperlipidemia (%)	200 (79)
Smoker (%)	46 (18)
Diabetes (%)	58 (23)
Family history (%)	79 (31)

Table 2

Quantitative CTA plaque measures in ischemic and non-ischemic vessels (median and interquartile ranges)

	Ischemic (FFR≤ 0.8)	**Non-ischemic (FFR>0.8)**	**p-value**
Diameter stenosis (%)	67.6 (55.1 – 93.9)	47.4 (31.2 – 59.0)	**<0.0001** *
Plaque volumes (mm^3)			
Low-density NCP	46.7 (28.2 – 68.4)	13.8 (3.7 – 32.7)	**<0.0001** *
NCP	273.3 (159.9 – 346.3)	96.8 (40.3 – 213.3)	**<0.0001** *
CP	10.6 (2.1 – 46.2)	6.5 (0 – 25.95)	**0.01** *
Total Plaque	296.3 (173.7 – 390.1)	112.5 (47.1 – 243.6)	**<0.0001** *
Plaque length (mm)	42.9 (24.7 – 58.4)	19.6 (8.8 – 35.5)	**<0.0001** *
Contrast density difference (%)	33.1 (31.8 – 48.9)	14.3 (7.4 – 23.1)	**<0.0001** *

* significant

Table 3

IDI for the integrated machine learning score over plaque measures

Plaque measure	IDI (95% CI)	p-value
Contrast density difference	0.13 (0.08–0.18)	**<0.0001** *
Low-density non-calcified plaque volume	0.16 (0.10–0.21)	**<0.0001** *
Non-calcified plaque volume	0.20 (0.14–0.25)	**<0.0001** *
Total plaque volume	0.22 (0.16–0.27)	**<0.0001** *

* significant

Table 4

Spearman rank correlations for quantitative plaque measures with FFR

Quantitative plaque measures	Correlation with FFR	p-value
Low-density non-calcified plaque volume	−0.50	**<0.0001** *
Non-calcified plaque volume	−0.44	**<0.0001** *
Total plaque volume	−0.44	**<0.0001** *
Quantitative stenosis	−0.48	**<0.0001** *
Minimum luminal area	0.38	**<0.0001** *
Contrast density difference	−0.51	**<0.0001** *
Plaque length	−0.44	**<0.0001** *

* significant