Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

CCTA–QCA disagreement in calcified coronary lesions and the limited value of linear recalibration: a retrospective clustered method-comparison study

Yawei Wu
The Affiliated Hospital of Xuzhou Medical University
Xiaoyan Xu
The Affiliated Hospital of Xuzhou Medical University
Jin Chen
The Affiliated Hospital of Xuzhou Medical University
Zihang Sun
The Affiliated Hospital of Xuzhou Medical University
Chunfeng Hu
3325677068@qq.com
The Affiliated Hospital of Xuzhou Medical University

Research Article

Keywords: coronary computed tomography angiography, coronary calcification, quantitative coronary angiography, Bland–Altman analysis, regression calibration, internal validation

Posted Date: August 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10728777/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Objective

To quantify the continuous difference between coronary computed tomography angiography–derived diameter stenosis (CCTA-DS) and quantitative coronary angiography–derived diameter stenosis (QCA-DS) in calcified coronary lesions, examine lesion characteristics associated with this difference, and explore whether linear recalibration could reduce measurement error.

Methods

This single-center retrospective study screened 720 patients and included 450 patients with 862 anatomically matched target vessels. MLD and RVD were measured manually with QCA software in two crossed projections and averaged separately. QCA-DS was calculated as 100 × [1 − (mean MLD/mean RVD)]. The CCTA–QCA difference was defined as CCTA-DS minus QCA-DS. Agreement uncertainty was estimated by patient-cluster bootstrap, with patient-equal and random-intercept repeated-measures analyses. Exploratory recalibration was evaluated by 30 repetitions of patient-grouped 10-fold cross-validation.

Results

Mean CCTA-DS and QCA-DS were 68.91% ± 20.38% and 52.08% ± 20.51%, respectively. The mean CCTA–QCA difference was 16.83 percentage points (95% CI, 15.35 to 18.25), with global limits of agreement from − 26.47 to 60.12 percentage points. The random-intercept analysis gave similar limits. CCTA-DS was greater than QCA-DS in 79.2% of vessels. Mean differences were 20.31, 20.97, and 23.73 percentage points for target-vessel CACS ≥ 300, circumferential calcium grade 2, and diffuse lesions, respectively. In adjusted analysis, local calcium volume and diffuse morphology were associated with larger positive differences. Extended recalibration achieved an MAE of 15.25, an RMSE of 18.34, and an R^2 of 0.199. Errors were within ± 10 and ± 20 percentage points in 37.0% and 70.2% of vessels. The incremental benefit over CCTA-only recalibration was small and uncertain. At QCA-DS ≥ 50%, raw CCTA had a false-positive rate of 77.9%. The model did not reliably identify QCA-DS ≥ 70%.

Conclusions

In this selected cohort referred for invasive angiography, CCTA showed a positive average difference relative to QCA, but individual disagreement was substantial. Large local calcium volume and diffuse morphology were adjusted group-level warning features; high target-vessel CACS and extensive circumferential calcium were descriptive warning features. These findings support cautious

interpretation of the reported stenosis percentage. Linear recalibration should not be used for threshold-based clinical decision-making.

1 Introduction

Coronary computed tomography angiography (CCTA) has a high negative predictive value and is an important noninvasive anatomical test for suspected coronary artery disease [1–4]. CAD-RADS 2.0 continues to use the most severe luminal stenosis as a core classification component. It also considers plaque burden and ischemia-related modifiers when guiding subsequent management [2]. Therefore, the quantitative accuracy of CCTA stenosis measurements can directly affect referral for further testing.

Coronary calcification limits CCTA stenosis assessment. Partial-volume effects, beam hardening, and blooming artifacts can obscure the true lumen. They often cause CCTA to overestimate stenosis and reduce specificity and positive predictive value [5–10]. Earlier studies mainly reported diagnostic performance at binary thresholds of 50% or 70%, or stratified patients by total calcium score. Binary analyses lose information from continuous stenosis measurements. A patient-level total calcium burden may also fail to capture local target-lesion morphology. Kruk et al. [9] reported that calcium length, volume, arc, density, and lumen size may affect CCTA measurement error. More recent calcium-removal, spectral, and photon-counting CT studies further show that reducing calcium artifacts can improve stenosis quantification [11–13].

To our knowledge, relatively few studies have used continuous QCA-DS as the reference while accounting for multiple vessels within a patient and examining local calcification features. The primary objective was to assess continuous agreement between CCTA-DS and QCA-DS. The secondary objective was to examine lesion characteristics associated with the CCTA–QCA difference. The exploratory objective was to evaluate whether linear recalibration based on information available before CAG could reduce continuous measurement error.

2 Materials and methods

2.1 Study design and participants

This was a single-center retrospective agreement and model development study. The existing de-identified database included patients who underwent CCTA and coronary angiography (CAG) during the same admission at The Affiliated Hospital of Xuzhou Medical University between January 2023 and June 2024. Inclusion criteria were age ≥ 18 years, at least one calcified coronary plaque on CCTA, and CAG within one month after CCTA. The de-identified analytic extract retained confirmation of the within-one-month eligibility criterion but not the exact calendar dates or interval days. Exclusion criteria were incomplete clinical or imaging variables; anomalous coronary origin, course, or termination; complete target-vessel occlusion on CCTA; previous coronary artery bypass grafting, coronary stenting, pacemaker, implantable cardioverter-defibrillator, or prosthetic valve; and image quality insufficient for

reliable CCTA quantification or lesion matching. The source database contained 720 patients. We excluded 91 for incomplete data, 13 for coronary anomalies, 34 for complete target-vessel occlusion, 79 for previous revascularization or related implants, and 53 for poor image quality or unreliable matching. The complete-case analysis included 450 patients. Missing values were not imputed. The study flow is provided in Additional file 1.

The Ethics Committee of The Affiliated Hospital of Xuzhou Medical University approved the study in 2025 (XYFY2025-KL105-01) and waived informed consent for participation. Approval was obtained before study-specific data extraction and statistical analysis began. All analytic data were de-identified. Reporting followed STROBE for the observational study, TRIPOD + AI for model development and internal validation, and relevant STARD items for the secondary threshold analyses [17, 22, 23].

2.2 CCTA acquisition and stenosis quantification

CCTA was performed with a third-generation dual-source CT system (SOMATOM Force; Siemens Healthineers, Forchheim, Germany) containing two 96-row detector arrays. The acquisition generated up to 192 reconstructed slices, with nominal z-axis coverage of 9.6 cm per rotation. Gantry rotation time was 0.25 s, and nominal single-sector temporal resolution was 66 ms. A fixed heart-rate cutoff was not required. Images were reconstructed with ADMIRE at strength 3 or 4. The local protocol used B30f for routine coronary reconstruction, B46f for high-resolution evaluation of calcified plaque, and B20f for soft tissue. Iodinated contrast medium containing 320–350 mg iodine/mL was injected at 4.5–5.5 mL/s. Bolus tracking was performed at the aortic root with a 100-HU threshold, and scanning began 5 s after triggering.

CCTA-DS and calcification measures were obtained semiautomatically with Shukun Coronary Analysis System v3.0 (Shukun Technology, Beijing, China). Two trained cardiovascular imagers jointly reviewed centerlines, curved multiplanar reconstructions, cross-sectional images, and lumen segmentation. They were blinded to CAG and QCA results. Segmentation errors caused by calcium blooming, motion, or step artifacts were corrected manually. Disagreements were resolved by consensus. Independent measurements and repeat readings were not stored. Therefore, interobserver and intraobserver intraclass correlation coefficients could not be calculated for CCTA-DS or calcification measures. CCTA-DS was the final software-reported diameter stenosis after manual correction of lumen segmentation.

When a vessel contained several calcified plaques, target-lesion selection considered both apparent CCTA stenosis severity and calcium burden. Because selection was based on CCTA appearance, its potential to favor extreme CCTA measurements is considered in the Limitations. Target-vessel coronary artery calcium score (CACS) was obtained from the non-contrast calcium-scoring scan by the Agatston method in Shukun Coronary Analysis System v3.0. The software first segmented the vessel. A physician then manually delineated or corrected the contour to separate the target-vessel score from the whole-heart score. Local target-lesion calcium volume was recorded in mm^3 using software segmentation with manual contour correction. We also recorded lesion type (focal or diffuse), circumferential calcium grade, and the software-reported mean calcium density at the most stenotic site. Circumferential grade

was 0 for an arc < 90°, 1 for 90°–180°, and 2 for > 180°. Target-vessel CACS strata were defined as low (< 100), intermediate (100–299), and high (≥ 300) for description in this cohort. These strata are not equivalent to patient-level Agatston risk categories.

2.3 Coronary angiography and QCA

All CAG examinations were performed by experienced cardiologists in a standard catheterization laboratory using GE Healthcare IGS 540 or Innova 3100 angiography systems. Iomeprol or iodixanol was used as the contrast agent. The left and right coronary arteries were examined sequentially in eight standard projections. Additional projections were obtained when needed for optimal vessel visualization. All images were stored digitally and transferred to an AW 4.7 workstation for post-processing and quantitative analysis.

QCA was performed on the AW 4.7 workstation. A 5-Fr angiographic catheter was used for calibration. For each calcified target segment, two crossed projections were selected that clearly showed the lesion and minimized foreshortening and overlap. Minimum lumen diameter (MLD) and reference vessel diameter (RVD) were measured manually with the QCA software in each projection. MLD and RVD were each averaged across the two projections. QCA-DS was calculated as 100 × [1 − (mean MLD/mean RVD)]. This study-specific two-view composite was used consistently and does not represent the DS from either individual projection. Numerical CCTA and QCA measurements were performed independently, and readers were blinded to the numerical results of the other modality. After numerical measurements were completed, both image sets were viewed together to match lesions by vessel segment, major branches, and adjacent anatomical landmarks. QCA was treated as the anatomical reference, but not as an error-free truth.

2.4 Outcomes, candidate factors, and model purpose

The primary agreement outcome was the paired difference between CCTA-DS and QCA-DS, defined as bias = CCTA-DS − QCA-DS. A positive value indicated that CCTA-DS was numerically greater than QCA-DS. The exploratory recalibration outcome was continuous QCA-DS. Candidate factors were selected on clinical grounds. They included CCTA-DS, $\log_2$(local calcium volume + 1), circumferential grade, mean calcium density at the most stenotic site, heart rate during CCTA, lesion type, and target vessel. CACS was highly correlated with calcium volume. It was therefore used mainly for stratified description and was not entered with local calcium volume in the exploratory model. Age, sex, body mass index, hypertension, diabetes, hyperlipidemia, smoking history, low-density lipoprotein cholesterol, and estimated glomerular filtration rate were used only in sensitivity analyses. Post-QCA variables, including MLD, RVD, and QCA lesion length, were not included as predictors in the recalibration model.

2.5 Statistical analysis

Continuous variables are reported as mean ± standard deviation or median (interquartile range), according to distribution. Categorical variables are reported as number and percentage. Global 95% limits of agreement were calculated as mean difference ± 1.96 SD, with repeated-measures extensions

for clustered observations [14, 15]. A random-intercept model fitted by restricted maximum likelihood partitioned between-patient and within-patient variance; repeated-measures limits were calculated as mean difference ± 1.96 × √(between-patient variance + within-patient variance). REML used a bounded two-parameter Nelder–Mead search on log-transformed variances (maximum 120 iterations; simplex tolerance 1×10^{-7}). Numerical stability was checked with nine dispersed starting values. An independent coordinate-wise likelihood search reproduced the variance estimates to three decimals. A patient-equal analysis gave each patient equal total weight. Patient-cluster bootstrap resampling with 2,000 replicates estimated 95% confidence intervals for agreement estimates and model coefficients. Heteroscedasticity was assessed by regression of absolute residuals on the method mean. Regression-based conditional limits used the fitted mean difference and the predicted mean absolute residual multiplied by √(π/2) to estimate the conditional SD. Stratified limits were also reported. Pearson correlation, Lin CCC [16], MAE, RMSE, and errors within ± 10 and ± 20 percentage points were calculated.

Factors associated with the CCTA–QCA difference were examined in an exploratory multivariable linear model. The paired difference was the dependent variable. Predictors were method mean, $\log_2$(local calcium volume + 1), circumferential grade, calcium density per 100 HU, heart rate per 10 beats/min, lesion type, and target vessel. Because the difference and method mean share component measurements, coefficients were treated as conditional associations and not causal effects. Exploratory recalibration models predicted QCA-DS from CCTA-DS alone or from CCTA-DS plus pre-CAG factors. Circumferential grade was compared as a continuous trend and as categorical indicators. A post hoc piecewise-linear sensitivity model used knots at 30%, 50%, 70%, and 85% to span common clinical and upper-range stenosis values; these knots were not treated as prospectively prespecified. Interaction and clinical-variable models were also examined. QCA-DS ranged from 3.2% to 97.4%; no 0% or 100% endpoint was present, so the bounded logit model required no endpoint correction. Ordinary least-squares predictions were not truncated.

Exact calendar dates and interval days were not retained. Therefore, temporal validation and a single random split were not performed. Candidate factors were prespecified and limited in number, and all available participants were used for model development and internal validation, consistent with current guidance [18–21]. The random seed was 20250731. Internal validation used 30 repetitions of patient-grouped 10-fold cross-validation. For each vessel, the 30 out-of-fold predictions were averaged, and all reported model metrics and threshold tables used these fixed mean predictions. Calibration intercept and slope were estimated by ordinary least-squares regression of observed QCA-DS on the fixed mean out-of-fold prediction. Their intervals, and intervals for mean prediction difference, limits-of-agreement endpoints, and CCC, used 2,000 patient-cluster bootstrap samples of the fixed predictions without model refitting. Coefficient intervals used 2,000 patient-cluster bootstrap samples. Performance intervals for MAE, RMSE, and R^2 used 500 nested patient-cluster bootstrap samples; the full 30 × 10-fold procedure was repeated within each bootstrap sample, and all copies of an original patient remained in one fold. QCA-DS thresholds of 50% and 70% were secondary. Analyses used custom Python code with Python 3.12.13, NumPy 2.3.5, and pandas 3.0.1.

3 Results

3.1 Patients and lesion characteristics

The source database contained 720 patients. We excluded 270 patients across the five predefined categories. The complete-case analysis included 450 patients and 862 LAD, LCx, or RCA target vessels. One, two, and three vessels were contributed by 172, 144, and 134 patients. Mean age was 59.5 ± 9.3 years, and 295 patients (65.6%) were male. Previous coronary artery disease was present in 161 patients (35.8%), and 306 (68.0%) had symptoms or chest pain. Patient characteristics are shown in Table 1. Vessel characteristics and study flow are provided in Additional file 1.

Table 1
Baseline clinical characteristics of the study population (n = 450).

Variable	Value
Age, years	59.5 ± 9.3
Male	295 (65.6%)
BMI, kg/m²	24.3 ± 2.7
Hypertension	239 (53.1%)
Diabetes mellitus	106 (23.6%)
Hyperlipidemia	225 (50.0%)
Smoking history	115 (25.6%)
Heart rate during CCTA, beats/min	70.0 ± 8.2
LDL-C, mmol/L	2.29 ± 0.60
eGFR, mL/min/1.73 m²	88.3 ± 14.1
Previous coronary artery disease	161 (35.8%)
Symptoms or chest pain present	306 (68.0%)
Patients contributing 1/2/3 vessels	172 / 144 / 134

Data are mean ± standard deviation or n (%). BMI, body mass index; LDL-C, low-density lipoprotein cholesterol; eGFR, estimated glomerular filtration rate.

3.2 Agreement between CCTA and QCA

Mean CCTA-DS was 68.91% ± 20.38%, and mean QCA-DS was 52.08% ± 20.51%. CCTA-DS was greater than QCA-DS in 683 vessels (79.2%) and lower in 179 (20.8%). Pearson r was 0.416, and CCC was 0.311. Mean difference was 16.83 percentage points (95% CI, 15.35 to 18.25). The global limits of agreement

were − 26.47 (95% CI, − 29.15 to − 23.85) and 60.12 (95% CI, 57.42 to 62.49). Errors were within ± 10 percentage points in 246 vessels (28.5%) and within ± 20 percentage points in 447 (51.9%). Patient-equal mean difference was 16.27, with limits from − 26.53 to 59.08. The random-intercept mean difference was 16.82, with limits from − 26.47 to 60.12 and an ICC of 0.002. All nine dispersed REML starting-value checks converged to the same variance solution. Figure 1 shows a representative matched calcified lesion, and Fig. 2 shows the wide individual variability.

In the unadjusted proportional-difference regression, each 10-percentage-point increase in method mean was associated with a − 0.09-percentage-point change in difference (95% CI, − 0.96 to 0.77). The absolute residual decreased by 1.87 percentage points per 10-percentage-point increase (95% CI, − 2.35 to − 1.39), indicating heteroscedasticity. Regression-based conditional limits narrowed as method mean increased: −39.86 to 74.04 at a mean of 30%, − 30.84 to 64.68 at 50%, − 21.83 to 55.32 at 70%, and − 12.82 to 45.96 at 90%. These post hoc conditional estimates and stratified limits are shown in Additional file 1. Global limits are retained as an overall summary, not constant individual error bounds.

3.3 Calcification-burden strata and exploratory difference model

Mean bias was − 0.11, 16.34, and 20.31 percentage points in the predefined low, intermediate, and high target-vessel calcium-burden groups (CACS < 100, 100–299, and ≥ 300). Mean bias was 14.22, 17.15, and 20.97 percentage points for circumferential calcium grades 0, 1, and 2, respectively. Bias was greater in diffuse than in focal lesions (23.73 vs 15.44 percentage points) (Additional file 1). These patterns identify settings that warrant greater caution when interpreting the numerical CCTA stenosis estimate, but they are descriptive and not validated individual prediction rules.

In the exploratory multivariable difference model, each 10-percentage-point increase in method mean was associated with a − 0.93-percentage-point change in difference (95% CI, − 1.82 to − 0.10), compared with − 0.09 in the unadjusted analysis. Variance inflation factors ranged from 1.01 to 1.91 and were 1.23 for method mean, which did not indicate problematic multicollinearity. A one-unit increase in $\log_2$(local calcium volume + 1), representing a doubling of volume + 1, was associated with a 4.18-percentage-point larger positive difference (95% CI, 2.45 to 5.93). Diffuse lesions showed a 6.25-percentage-point larger positive difference than focal lesions (95% CI, 2.41 to 10.25). Confidence intervals for circumferential grade, calcium density, and heart rate included zero. Clinical-variable adjustment did not materially improve predictive performance. Local calcium volume and diffuse morphology were therefore the clearest adjusted warning features.

Table 2
Exploratory multivariable factors associated with the CCTA–QCA difference.

Variable	Coefficient, pp	Bootstrap 95% CI
Intercept	−13.11	−33.10 to 8.08
Mean stenosis of CCTA and QCA (per 10 pp)	−0.93	−1.82 to − 0.10
$\log_2$(local calcium volume + 1)	4.18	2.45 to 5.93
Circumferential grade (per grade)	0.31	−2.04 to 2.51
Mean calcium density (per 100 HU)	2.14	−0.73 to 5.02
Heart rate (per 10 beats/min)	−0.01	−1.75 to 1.63
Diffuse (vs focal)	6.25	2.41 to 10.25
LCx (vs LAD)	2.38	−1.56 to 5.84
RCA (vs LAD)	3.34	−0.29 to 6.94

The dependent variable is CCTA-DS − QCA-DS. Positive coefficients indicate a larger positive CCTA–QCA difference. The model is exploratory because the dependent variable and method mean share component measurements. LAD is the reference vessel, and focal lesion is the reference lesion type. pp, percentage points.

3.4 Exploratory regression recalibration and internal validation

The final calibration model retained all predefined variables without selection by univariable P values. The equation was fitted in the full dataset. LCx, RCA, and diffuse lesion are binary indicators, and circumferential grade is coded as 0, 1, or 2. Ordinary least-squares predictions were not truncated.

Predicted QCA-DS = 28.124 + 0.377 × CCTA-DS + 0.421 × $\log_2$(local calcium volume + 1)

− 1.013 × circumferential grade + 1.413 × (calcium density/100) − 0.528 × (heart rate/10)

− 3.275 × diffuse lesion − 5.800 × LCx − 9.252 × RCA

The bias model was conditional on method mean, whereas the calibration model was conditional on CCTA-DS. Their coefficients answer different questions and should not be compared directly. The full-data equation was reproduced from the raw workbook. Coefficients are rounded to three decimals for display; full-precision coefficients were used for all predictions. In the extended equation, the 95% bootstrap interval excluded zero for CCTA-DS, LCx, and RCA, but included zero for the other added imaging and examination factors. The maximum variance inflation factor was 2.03 in the clinical sensitivity model. Extended-model residual skewness was − 0.07 and excess kurtosis was − 0.54. Twenty-four observations exceeded the 2p/n leverage threshold, and 37 exceeded the 4/n Cook-distance

threshold; the maximum Cook distance was 0.020. Coefficient-level interpretation therefore remains cautious. Full-precision coefficients and diagnostics are in Additional file 1.

After averaging the 30 out-of-fold predictions for each vessel, the extended model had an MAE of 15.25 percentage points (nested-bootstrap 95% CI, 14.70 to 16.05), an RMSE of 18.34 (95% CI, 17.70 to 19.20), and an R^2 of 0.199 (95% CI, 0.145 to 0.232). Calibration intercept was 2.06, calibration slope was 0.961, and mean prediction difference was − 0.01. Errors were within ± 10 percentage points in 319 vessels (37.0%) and within ± 20 in 605 (70.2%). Out-of-fold predictions ranged from 22.35% to 70.37%. Full-data predictions ranged from 22.12% to 70.00%.

The CCTA-only model had an MAE of 15.55. Adding the predefined factors reduced MAE by 0.30 percentage points (nested-bootstrap 95% CI, − 0.03 to 0.56), reduced RMSE by 0.33 (95% CI, − 0.01 to 0.51), and increased R^2 by 0.029 (95% CI, − 0.001 to 0.044). All three intervals included no incremental benefit. Thus, most improvement came from recalibrating CCTA-DS itself, and the added factors had a small and uncertain contribution. The post hoc piecewise model had an MAE of 15.09, only 0.16 lower than the linear extended model.

Table 3
Continuous performance of raw CCTA and cross-validated exploratory recalibration models.

Method	Metric	Estimate	95% CI
Raw CCTA	MAE	22.33	21.14 to 23.42
	RMSE	27.76	26.49 to 28.91
	R^2	−0.834	−1.028 to − 0.652
	Calibration intercept	23.21	19.18 to 27.33
	Calibration slope	0.419	0.360 to 0.478
	Mean prediction difference	16.83	15.35 to 18.25
	95% limits of agreement	−26.47 to 60.12	Lower LoA: −29.15 to − 23.85; upper LoA: 57.42 to 62.49
	CCC	0.311	0.265 to 0.360
CCTA-only linear calibration	MAE	15.55	14.86 to 16.31
	RMSE	18.68	17.91 to 19.45
	R^2	0.170	0.128 to 0.212
	Calibration intercept	0.60	−6.73 to 8.02
	Calibration slope	0.989	0.848 to 1.131
	Mean prediction difference	0.00	−1.31 to 1.22
	95% limits of agreement	−36.63 to 36.63	Lower LoA: −38.48 to − 34.80; upper LoA: 34.44 to 38.53
	CCC	0.293	0.251 to 0.333
CCTA plus predefined factors	MAE	15.25	14.70 to 16.05
	RMSE	18.34	17.70 to 19.20
	R^2	0.199	0.145 to 0.232
	Calibration intercept	2.06	−4.07 to 8.19
	Calibration slope	0.961	0.844 to 1.075
	Mean prediction difference	−0.01	−1.32 to 1.22
	95% limits of agreement	−35.99 to 35.96	Lower LoA: −37.79 to − 34.14; upper LoA: 33.80 to 37.93

Method	Metric	Estimate	95% CI
	CCC	0.341	0.301 to 0.382

Raw CCTA does not involve model fitting. Raw-data intervals used 2,000 patient-cluster bootstrap samples. For fitted models, intervals for MAE, RMSE, and R^2 used 500 nested patient-cluster bootstrap samples that repeated the full cross-validation. Intervals for calibration intercept, calibration slope, mean prediction difference, limits of agreement, and CCC used 2,000 patient-cluster bootstrap samples applied to the fixed mean out-of-fold predictions without model refitting. Calibration intercept and slope were obtained by regressing observed QCA-DS on the fixed mean out-of-fold prediction. For limits of agreement, the interval column reports separate intervals for the lower and upper endpoints. MAE, RMSE, and mean prediction difference are in percentage points. Predictions were not truncated.

3.5 Clinical thresholds and sensitivity analyses

At QCA-DS ≥ 50%, prevalence was 472/862 (54.8%). Raw CCTA yielded 446 true positives, 26 false negatives, 304 false positives, and 86 true negatives; sensitivity and specificity were 94.5% and 22.1%. The corresponding false-positive rate was 77.9%. The extended model yielded 380, 92, 189, and 201, respectively; sensitivity and specificity were 80.5% and 51.5%. At QCA-DS ≥ 70%, prevalence was 245/862 (28.4%). Raw CCTA yielded 211 true positives, 34 false negatives, 335 false positives, and 282 true negatives. The extended model yielded no true positive, 245 false negatives, one false positive, and 616 true negatives. It should not be used as a 70% threshold classifier.

Sensitivity analyses supported the limited value of added factors. Patient-equal mean difference was 16.27. With equal patient weighting during model fitting and the same repeated patient-grouped cross-validation, extended-model MAE, RMSE, and R^2 were 15.27, 18.35, and 0.199. Categorical coding of circumferential grade gave 15.27, 18.37, and 0.197. Adding clinical variables gave 15.32, 18.45, and 0.190. Replacing local calcium volume with CPV gave 15.25, 18.35, and 0.199. The bounded logit model gave 15.21, 18.39, and 0.195, with predictions from 21.74% to 72.27%. None materially outperformed the main model. In a post hoc influence analysis, exclusion and refitting after removal of the 9 and 18 vessels with the largest Cook distances changed out-of-fold MAE by + 0.00 and − 0.01 percentage points and RMSE by − 0.01 and − 0.02, respectively, when compared on the same retained vessels. Mean absolute changes in predictions were 0.47 and 0.75 percentage points. Predictive performance was stable, although several secondary coefficients changed, supporting cautious coefficient interpretation.

4 Discussion

4.1 Principal findings

This study had four main findings. First, CCTA-DS showed a positive average difference of about 17 percentage points relative to QCA-DS in calcified coronary lesions, with wide and heteroscedastic individual differences. Second, mean differences increased across descriptive target-vessel CACS and

circumferential calcium strata. Local calcium volume and diffuse morphology were associated with larger positive differences in the exploratory multivariable model. Third, linear recalibration reduced MAE and RMSE, but calcification and examination factors added little beyond CCTA-only recalibration. Fourth, model sensitivity was 0% at QCA-DS ≥ 70%. Improvement in continuous error did not translate into severe-stenosis classification.

4.2 Calcification-related CCTA–QCA differences

The CORE-64 study and several meta-analyses showed that increasing calcification burden reduces the specificity of CCTA for obstructive disease, whereas sensitivity often remains high [4–6]. Kwan et al. [8] reported frequent positive CCTA–QCA differences and many false-positive classifications among patients with CACS > 1000. Our findings agree with that direction and quantify the difference on a continuous scale. CCTA-DS was numerically greater than QCA-DS in more than three-quarters of vessels, but any positive paired difference was counted and individual differences were wide. Continuous analysis shows that the same binary result of ≥ 50% can conceal very different error magnitudes.

Proportional difference and heteroscedasticity are distinct. The unadjusted regression showed no clear linear proportional difference, but absolute residuals decreased as mean stenosis increased. The more negative method-mean coefficient after adjustment likely reflects conditioning on calcium burden and morphology. Because variance inflation factors were low, this shift was not attributed to severe multicollinearity and was not interpreted causally. Patient-equal and random-intercept analyses gave nearly the same global limits as the vessel-level analysis because the difference ICC was only 0.002. However, conditional and stratified limits varied across the measurement range. Global limits therefore remain an overall summary and not a constant prediction interval for every stenosis level.

The associations of calcium volume and diffuse disease with bias are technically and anatomically plausible. Greater calcium volume increases the proportion of the lumen cross-section affected by partial-volume and blooming artifacts. Diffuse disease also makes it harder to define proximal, narrowest, and distal reference boundaries. Kruk et al. [9] likewise found that calcium length, volume, cross-sectional area, and smaller lumen size affect CCTA error. Circumferential grade showed a clear unadjusted trend, but its adjusted confidence interval included zero after accounting for calcium volume and lesion type. These measures may share information about total calcification burden.

CACS in this study was a target-vessel score, not a patient-level total score. It was closer to the target vessel than a whole-heart score but remained distinct from a plaque-level measure. Earlier work suggests that regional calcium distribution and morphology are more closely related to stenosis in a specific vessel than is a global score [7, 9]. We therefore used target-vessel CACS for predefined descriptive strata and avoided whole-heart Agatston risk labels. Local target-lesion calcium volume was preferred in the multivariable model to reduce mismatch of information level and collinearity with CACS. The CPV replacement analysis produced nearly identical results because the two volume measures were highly correlated and carried redundant information.

4.3 Limited value of exploratory regression recalibration

Recalibration reduced average prediction error but did not restore adequate individual agreement. CCTA-only recalibration produced most of the gain. The additional variables reduced MAE by only 0.30 percentage points, and the nested-bootstrap interval for this increment included zero. Model R^2 remained about 0.20, CCC improved only modestly, and recalibrated limits remained about ± 36 percentage points. The available variables were insufficient for precise prediction of QCA-DS in an individual vessel. Coefficient intervals were wide, and the equation should not be interpreted as a stable causal model.

In the secondary QCA-DS ≥ 70% analysis, the model produced only one positive prediction, which was false positive. Ordinary least squares estimates a conditional mean. When CCTA and QCA are only moderately correlated and QCA values are widely dispersed at a given CCTA value, extreme predictions are compressed toward the mean. The model is therefore intended for continuous measurement calibration, not for 70% threshold classification. A dedicated probability model would be required for that purpose and would need external validation.

Candidate variables were defined before model fitting. Post-angiography variables such as QCA MLD, RVD, and lesion length were not used as predictors. Resampling and cross-validation kept all vessels from the same patient together. The pooling rule, random seed, coefficient bootstrap, and nested validation are now reported. These steps reduce leakage and optimism but do not remove the limitations of single-center retrospective data or missing repeatability information.

4.4 Clinical implications

When CCTA-DS is intermediate or severe in a calcified lesion, the raw percentage should be interpreted cautiously, especially when local calcium volume is large, target-vessel CACS is ≥ 300, circumferential calcium exceeds 180°, or the lesion is diffuse. The CACS and circumferential findings were descriptive group-level signals. Local calcium volume and diffuse morphology remained associated with larger positive differences after adjustment. None is a validated individual prediction rule, and a direct false-positive probability model was not fitted. A lower recalibrated value cannot exclude true severe disease. When anatomical severity or intervention planning must be clarified, guideline-based functional testing, CAG, and intravascular imaging should be integrated as needed. Calcium subtraction, spectral CT, and photon-counting CT may reduce calcification artifacts [11–13]. Improving image formation may be a more promising strategy than post hoc recalibration and warrants direct comparative evaluation.

4.5 Limitations

First, this was a single-center retrospective complete-case study. It included only patients who underwent both CCTA and CAG. Referral bias and spectrum enrichment are therefore likely, and the findings may not represent a general CCTA population.

Second, 270 patients were excluded because of incomplete data or imaging, coronary anomalies, complete target-vessel occlusion, previous bypass surgery or stenting and related implants, poor image quality, or unreliable matching. Exclusion of patients with incomplete data or poor image quality may

have caused complete-case selection bias. Baseline characteristics were not compared between included and excluded patients, so the direction and magnitude of this bias cannot be quantified.

Third, exact calendar dates and CCTA-to-CAG interval days were not retained, which prevented temporal and interval-stratified validation. The scanner model, temporal resolution, reconstruction method, kernels, contrast flow, and trigger settings are now reported. However, patient-level tube-voltage, tube-current, gating-mode, and radiation-dose fields were unavailable for protocol-stratified bias analysis. Use of one center, one scanner platform, and one analysis software version limits generalizability.

Fourth, standardized image-quality scores were not recorded. The operational definitions retained for focal versus diffuse morphology and software-reported mean calcium density were insufficient for independent reproduction. CCTA-DS and local calcification measures came from the same software-assisted, manually corrected workflow and may share method error. CCTA-available lumen dimensions, lesion location, bifurcation involvement, and other image-quality features were not available for residual-confounding assessment. Two physicians jointly reviewed CCTA post-processing and resolved disagreements, but independent readings and repeat measurements were not stored. Interobserver and intraobserver repeatability could not be estimated for CCTA-DS or calcification measures.

Fifth, target-lesion selection considered both apparent CCTA stenosis severity and calcium burden. The rule did not define how discordant rankings by these two criteria were resolved. Selection based partly on the largest apparent CCTA stenosis may preferentially retain extreme CCTA values and amplify the positive average difference through extreme-value selection and regression to the mean. The retained study dataset did not permit lesion-level reanalysis using an alternative selection rule. Completely occluded vessels and images that could not be analyzed reliably were excluded, which may underestimate the risk of non-diagnostic lesions in practice.

Sixth, QCA is a two-dimensional projection-based anatomical reference. The study-specific QCA-DS combined mean MLD and mean RVD from two crossed projections and therefore did not represent the DS from either single view. Projection selection, reference-segment definition, and manual boundary identification may affect the reference measurement. Independent core-laboratory repeat quantification was not performed, and independent or repeat QCA measurements were not stored. QCA interobserver and intraobserver repeatability could not be estimated.

Seventh, patient-equal and random-intercept agreement analyses were completed and gave results similar to the global vessel-level analysis. The between-patient variance estimate was small. The regression-based conditional limits were post hoc and depended on linear models for mean difference and absolute residuals plus a normal-error scaling assumption. They require confirmation in an external cohort and should not be interpreted as validated individual prediction intervals.

Eighth, individual prediction error remained large. The study did not include invasive fractional flow reserve, clinical outcomes, or treatment effects. The exploratory model cannot determine physiological significance or directly guide intervention.

5 Conclusions

In patients with calcified coronary lesions who were referred for invasive angiography, CCTA-DS showed a positive average difference of approximately 17 percentage points relative to QCA-DS, although individual differences were wide and varied across the measurement range. Mean differences were larger across descriptive target-vessel CACS strata, whereas local calcium volume and diffuse lesion morphology were associated with larger positive differences in the exploratory multivariable model. Linear recalibration reduced average continuous error, but the incremental contribution of calcification and examination variables was small, and individual prediction remained imprecise. These findings support cautious interpretation of numerical CCTA stenosis estimates, particularly in lesions with large local calcium volume or diffuse morphology. High target-vessel CACS and extensive circumferential calcium should be treated as descriptive warning features, and the proposed equation should not be used for clinical threshold classification.

Declarations

Ethics approval and consent to participate:

The study complied with the Declaration of Helsinki. The Ethics Committee of The Affiliated Hospital of Xuzhou Medical University approved the study in 2025 (XYFY2025-KL105-01). Approval was obtained before study-specific data extraction and statistical analysis began. The requirement for informed consent to participate was waived.

Consent for publication:

The patient shown in Fig. 1 provided written informed consent for publication of the de-identified CCTA and CAG images.

Competing interests:

The authors declare that they have no competing interests.

Funding:

This study received no specific grant or other dedicated funding.

Author Contribution

Contributions are reported using the CRediT taxonomy.

Yawei Wu: conceptualization, methodology, formal analysis, and writing—original draft.

Xiaoyan Xu: data curation, visualization, and investigation.

Jin Chen: software, visualization, and investigation.

Zihang Sun: investigation and data curation.

Chunfeng Hu: supervision, project administration, and writing—review and editing.

All authors interpreted the results, critically revised the manuscript, and approved the final version.

Acknowledgements:

Not applicable.

Data Availability

De-identified participant-level data are not publicly available because of patient privacy, ethics approval, and institutional clinical-data requirements. Controlled access may be discussed with the corresponding author for scientifically justified non-commercial requests, subject to approval by the institutional ethics committee and data custodian.

References

1. Narula J, Chandrashekhar Y, Ahmadi A, et al. SCCT 2021 Expert Consensus Document on Coronary Computed Tomographic Angiography. J Cardiovasc Comput Tomogr. 2021;15(3):192–217. 10.1016/j.jcct.2020.11.001.
2. Cury RC, Leipsic J, Abbara S, et al. Radiol Cardiothorac Imaging. 2022;4(5):e220183. 10.1148/ryct.220183. CAD-RADS 2.0-2022 Coronary Artery Disease-Reporting and Data System.
3. Miller JM, Rochitte CE, Dewey M, et al. Diagnostic performance of coronary angiography by 64-row CT. N Engl J Med. 2008;359(22):2324–36. 10.1056/NEJMoa0806576.
4. Arbab-Zadeh A, Miller JM, Rochitte CE, et al. Diagnostic accuracy of computed tomography coronary angiography according to pre-test probability and severity of coronary calcification: the CORE-64 study. J Am Coll Cardiol. 2012;59(4):379–87. 10.1016/j.jacc.2011.06.079.
5. Abdulla J, Pedersen KS, Budoff M, Kofoed KF. Influence of coronary calcification on the diagnostic accuracy of 64-slice CT coronary angiography: a systematic review and meta-analysis. Int J Cardiovasc Imaging. 2012;28(4):943–53. 10.1007/s10554-011-9902-6.
6. den Dekker MAM, de Smet K, de Bock GH, et al. Diagnostic performance of coronary CT angiography for stenosis detection according to calcium score: systematic review and meta-analysis. Eur Radiol. 2012;22(12):2688–98. 10.1007/s00330-012-2551-x.

7. Qi L, Tang LJ, Xu Y, et al. The diagnostic performance of coronary CT angiography for the assessment of coronary stenosis in calcified plaque. PLoS ONE. 2016;11(5):e0154852. 10.1371/journal.pone.0154852.
8. Kwan AC, Gransar H, Tzolos E, et al. The accuracy of coronary CT angiography in patients with coronary calcium score above 1000 Agatston units: comparison with quantitative coronary angiography. J Cardiovasc Comput Tomogr. 2021;15(5):412–8. 10.1016/j.jcct.2021.03.007.
9. Kruk M, Noll D, Achenbach S, et al. Impact of coronary artery calcium characteristics on accuracy of CT angiography. JACC Cardiovasc Imaging. 2014;7(1):49–58. 10.1016/j.jcmg.2013.07.013.
10. Shmilovich H, Cheng VY, Tamarappoo BK, et al. Optimizing image contrast display improves quantitative stenosis measurement in heavily calcified coronary arterial segments on coronary CT angiography. Acad Radiol. 2014;21(6):797–804. 10.1016/j.acra.2014.02.016.
11. Qiao J, Li S, Yang H, et al. Subtraction improves the accuracy of coronary CT angiography in patients with severe calcifications in identifying moderate and severe stenosis: a multicenter study. Acad Radiol. 2023;30(12):2801–10. 10.1016/j.acra.2022.11.033.
12. Mergen V, Rusek S, Civaia F, et al. Virtual calcium removal in calcified coronary arteries with photon-counting detector CT-first in-vivo experience. Front Cardiovasc Med. 2024;11:1367463. 10.3389/fcvm.2024.1367463.
13. Ran Q, Chen D, Zhang H, Zhang L. When can coronary computed tomography angiography in patients with calcified plaque be accurate? Front Cardiovasc Med. 2025;12:1570517. 10.3389/fcvm.2025.1570517.
14. Parker RA, Scott C, Inácio V, Stevens NT. Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners. BMC Med Res Methodol. 2020;20:154. 10.1186/s12874-020-01022-x.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307–10. 10.1016/S0140-6736(86)90837-8.
16. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989;45(1):255–68. 10.2307/2532051.
17. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. 10.1136/bmj-2023-078378.
18. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ. 2024;384:e074819. 10.1136/bmj-2023-074819.
19. Efthimiou O, Seo M, Chalkou K, et al. Developing clinical prediction models: a step-by-step guide. BMJ. 2024;386:e078276. 10.1136/bmj-2023-078276.
20. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368:m441. 10.1136/bmj.m441.
21. Steyerberg EW. Validation in prediction research: the waste by data splitting. J Clin Epidemiol. 2018;103:131–3. 10.1016/j.jclinepi.2018.07.010.

22. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. 10.1371/journal.pmed.0040296.
23. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527. 10.1136/bmj.h5527.

Figures

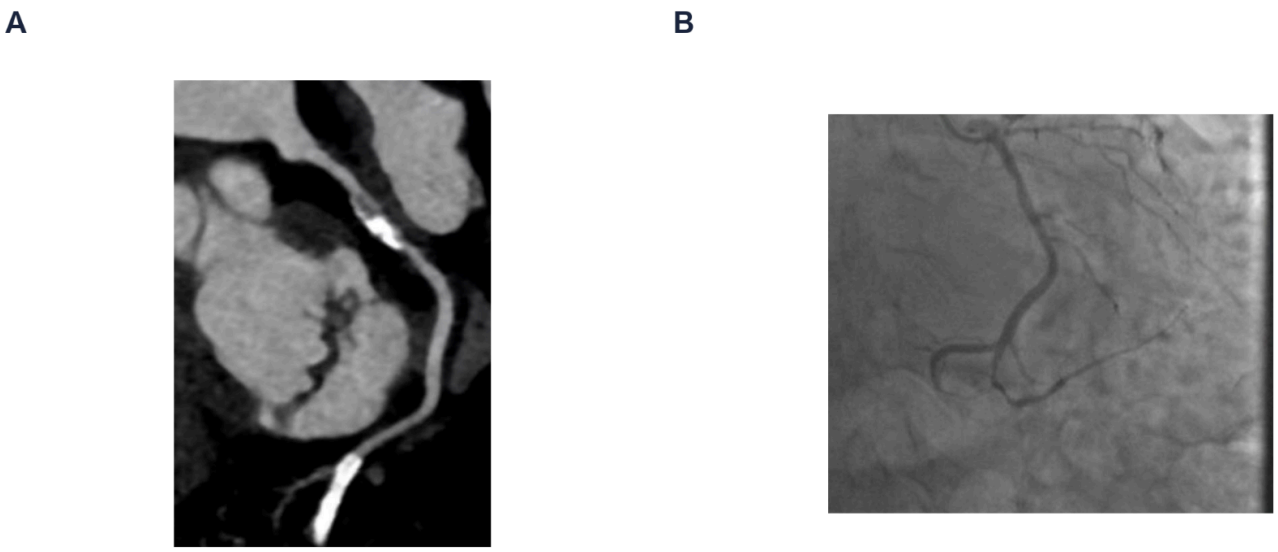


Figure 1

Representative calcified coronary stenosis. A, Curved multiplanar CCTA image of a proximal left circumflex lesion at the most stenotic site. CCTA-DS was 71.0%; the arrow marks the narrowest lumen. B, CAG image of the same lesion at the matched anatomical location. Mean MLD and mean RVD across the two crossed projections were 2.2 mm and 3.2 mm, respectively, and QCA-DS was 31.1%; the arrow marks the target lesion. CCTA-DS was 39.9 percentage points greater than QCA-DS. MLD and RVD are displayed to one decimal; recalculation of QCA-DS from these displayed values may therefore differ slightly because of rounding.

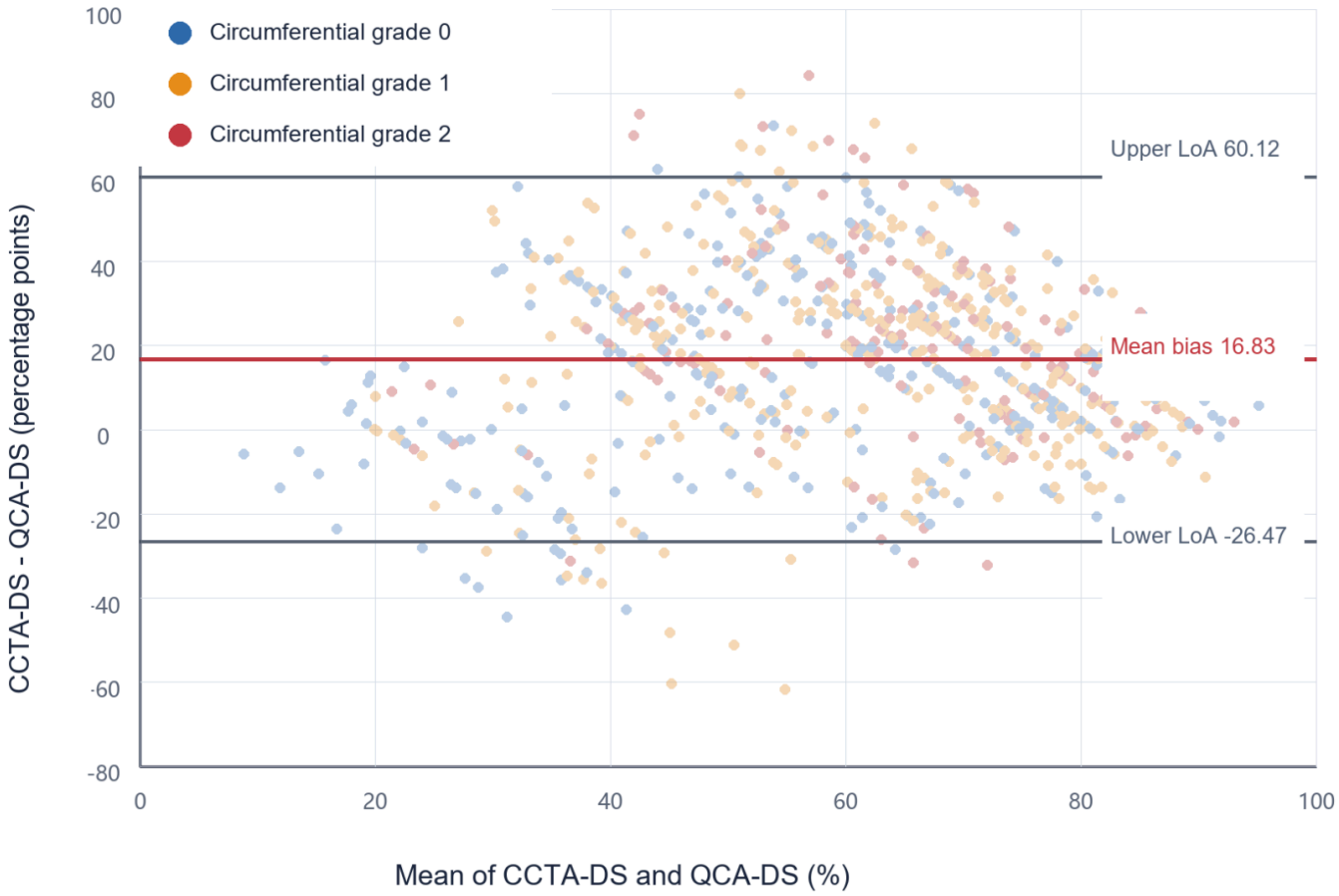


Figure 2

Bland–Altman analysis of CCTA and QCA. The vertical axis is CCTA-DS − QCA-DS, and the horizontal axis is the mean of the two methods. The red line shows the mean difference, and the gray lines show descriptive global 95% limits of agreement. Color indicates circumferential calcium grade. Error variance was not constant across the measurement range; constant limits should therefore not be interpreted as conditional individual error bounds. Patient-cluster bootstrap confidence intervals were estimated but are not displayed.