e-Cardiology/Digital Health – Artificial Intelligence, Deep Learning

Fully automated plaque quantification with human-level performance and validation in large-scale cardiac CT cohort

I. Sobirov[1], A. Bakleh[2], K. Chan[1], E. Wahome[1], C. Xie[1], T. Halborg[1], A. Antonopoulos[1], S. Fry[2], M. Siddique[2], H. Ahmed[2], S. Neubauer[1], K. Channon[1], J. Deanfield[3], C. Antoniades[1], P. Tomlins[2]

[1]University of Oxford, Oxford, United Kingdom of Great Britain & Northern Ireland
[2]Caristo Diagnostics, Oxford, United Kingdom of Great Britain & Northern Ireland
[3]University College of London, Institute of Cardiovascular Science, London, United Kingdom of Great Britain & Northern Ireland
On behalf of the ORFAN Consortium

Funding Acknowledgements: Type of funding sources: Public Institution(s). Main funding source(s): University of Oxford, British Heart Foundation

Background: Coronary artery disease is the leading cause of morbidity and mortality worldwide, with atherosclerotic plaque burden recognised as a critical biomarker for cardiovascular risk. Although calcium scoring is widely used, it provides only partial information, and its manual nature limits scalability. Large-scale cardiac CT angiography (CCTA) registries linked with long-term outcomes offer a unique opportunity for population-level risk stratification. However, the absence of robust, fully automated tools for comprehensive plaque quantification—including both calcified plaque (CP) and non-calcified plaque (NCP)—continues to impede clinical translation.
Purpose: We aimed to develop and validate a fully automated AI pipeline for vessel and lumen segmentation, plaque region identification, and quantification of both CP and NCP. We evaluated its agreement with expert assessment and its correlation with conventional calcium scoring.
Methods: We employed an active learning framework to train and refine the vessel and lumen segmentation model, leveraging iterative feedback from clinical experts as illustrated in Figure 1 (a). A total of 1,200 patients from the ORFAN and NHS-Pilot studies were used across all training tasks, with non-overlapping subsets allocated for vessel/lumen segmentation and for quantification of CP and NCP. Plaque region identification and CP/NCP segmentation models were trained to delineate plaque within the anatomical "sandwich" between lumen and vessel walls (See Figure 1(b) for a sample). Model outputs were compared against expert annotations. Validation included assessment of the correlation between AI-derived CP/NCP burden and manual calcium scoring using the correlation coefficient at the patient level.
Results: The lumen and vessel segmentation models performed exceptionally well, attaining an overall agreement of 0.95 Dice similarity with expert annotators in an external cohort through model refinement via an active learning process. Initially, the plaque quantification method was compared to clinical annotations of burden in 104 cases, resulting in a Pearson's correlation of 0.93 for NCP and 0.98 for CP regions. Following this, the method was externally validated in a multicentre cohort of 19463 patients from ORFAN to thoroughly confirm its performance against human-calcium scoring, achieving a Spearman's correlation coefficient of 0.81 (p-value<0.01) for CP and 0.66 (p-value<0.001) for NCP, as shown in Figure 1 (b).
Conclusion: This study demonstrates the feasibility of fully automating the quantification of calcified and non-calcified coronary plaque using artificial intelligence. Manual assessment of plaque burden in large-scale cohorts such as ORFAN is impractical; AI offers a scalable alternative that enables population-level risk stratification. Given the strong association between plaque burden and cardiovascular risk, this work lays the foundation for improved disease prediction and management in large and diverse patient populations.

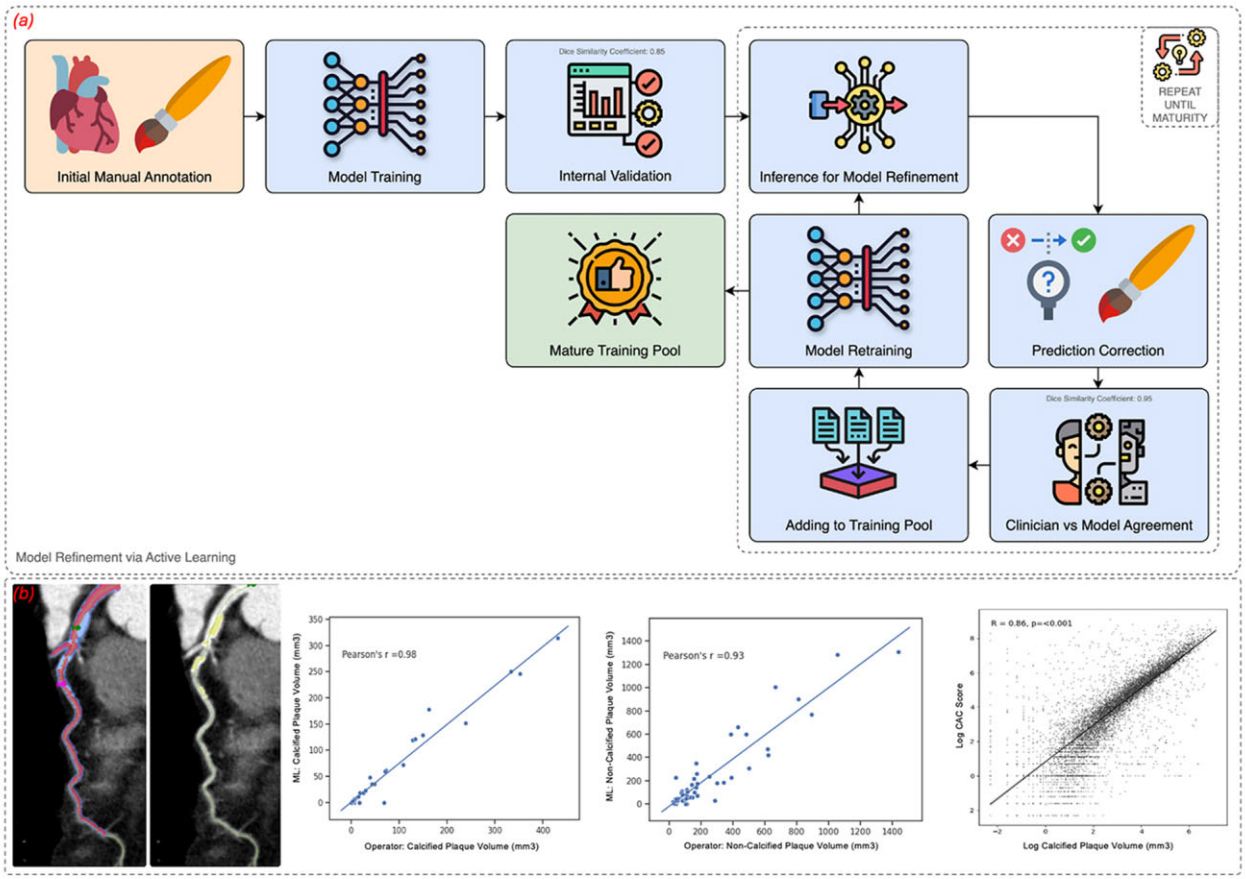

Figure 1. (a) shows the 'Model Refinement via Active Learning' framework. Within this framework, we iteratively improve the model's performance by feeding the model with experts' feedback to increase the training pool. (b) shows a sample of vessel/lumen segmentation, and CP/NCP correlation with human annotation, along with the correlation plot for calcium score with CP.