Universidade de Lisboa

Faculdade de Medicina

Artificial Intelligence Applied to Invasive Coronary Angiography and Physiology

Miguel de Almeida Nobre de Menezes

Orientadores:

Prof. Doutor Fausto José da Conceição Alexandre Pinto

Prof. Doutor Arlindo Manuel Limede de Oliveira

Tese especialmente elaborada para obtenção do grau de Doutor em Medicina, Cardiologia.

2024

Universidade de Lisboa

Faculdade de Medicina

Artificial Intelligence Applied to Invasive Coronary Angiography and Physiology

Miguel de Almeida Nobre de Menezes

Orientadores:

Prof. Doutor Fausto José da Conceição Alexandre Pinto

Prof. Doutor Arlindo Manuel Limede de Oliveira

Tese especialmente elaborada para obtenção do grau de Doutor em Medicina, Cardiologia.

Júri:

Presidente: Doutor Luís Alberto da Cunha Mendes Pedro, Professor Catedrático e Vice-Presidente do Conselho Científico da Faculdade de Medicina da Universidade de Lisboa.

Vogais:

- PhD Mamas Andreas Mamas, Professor da Keele University (United Kingdom);
- Doutor Lino Manuel Martins Gonçalves, Professor Catedrático da Faculdade de Medicina da Universidade de Coimbra;
- Doutora Marta Ramos Tavares da Silva, Professora Auxiliar Convidada da Faculdade de Medicina da Universidade do Porto;
- Doutor Fausto José da Conceição Alexandre Pinto, Professor Catedrático da Faculdade de Medicina da Universidade de Lisboa (*Orientador*);
- Doutora Ana Maria Gomes de Almeida, Professora Associada com Agregação da Faculdade de Medicina da Universidade de Lisboa.

2024

A impressão desta tese foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 16 de Abril de 2024.

As opiniões expressas nesta publicação são da exclusiva responsabilidade do seu autor

ACKNOWLEDGMENTS

Acknowledgments are best delivered in the native language of the recipients and myself, as it better conveys my deep expression of gratitude to them all.

Ao Professor Fausto Pinto, por todo o apoio, encorajamento, supervisão e a ligação inter-institucional, sem a qual a implementação deste projecto, verdadeiramente translacional, não teria sido possível.

Ao Professor Arlindo Oliveira, por todo o apoio, disponibilidade, supervisão e operacionalização do lado técnico deste estudo, fundamental para a sua exequibilidade.

Ao Eng. João Lourenço Silva, que efectivamente executou o lado *engineering* de toda a investigação, incansavelmente constituindo *de facto* a sua outra metade.

Ao falecido e saudoso Dr. Pedro Canas da Silva, ex-Coordenador da Unidade de Cardiologia de Intervenção Joaquim Oliveira, e ao Dr. Eduardo Infante de Oliveira, que me inspiraram e acarinharam enquanto interno, levando-me a seguir a área de Cardiologia de Intervenção.

Às Professoras Manuela Fiúza, tutora do Internato, e Ana G. Almeida, tutora do Trabalho Final de Mestrado Integrado em Medicina, que desde cedo me inspiraram e motivaram o desenvolvimento do pendor científico.

À Dra. Beatriz Valente Silva, verdadeira *compagnon de route* de investigação, que personifica uma natural e mútua afinidade científica, cujos prolíficos resultados se encontram espelhados em numerosas publicações recentes.

Aos demais co-autores dos projectos aqui expostos, sem os quais não teria sido possível a sua execução.

Aos colegas da Unidade de Cardiologia de Intervenção Joaquim Oliveira e demais profissionais, que possibilitam a realização de tudo quanto fazemos quotidianamente.

À minha Família, passada e presente, a quem devo tudo quanto sou.

TABLE OF CONTENTS

INDEX OF FIGURES

Table of Contents

INDEX OF TABLES

ABBREVIATION LIST

2D	Two-Dimensional
3D	Three-Dimensional
ACC	American College of Cardiology
ACQ	Análise Coronária Quantitativa
ACS	Acute Coronary Syndrome
ADVISE	ADenosine Vasodilator Independent Stenosis Evaluation
AGI	Artificial General Intelligence
AHA	American Heart Association
AI	Artificial Intelligence
ALCAPA	Anomalous Left Coronary Artery from the Pulmonary Artery
AMI	Acute Myocardial Infarction
AV	Atrioventricular
BAI	Baseline Artificial Intelligence Model
BARI	Bypass Angioplasty Revascularization Investigation
BCIS	British Cardiovascular Intervention Society
BH	Baseline Human Model
BMS	Bare-Metal Stent
CABG	Coronary Artery Bypass Grafting
CAD	Coronary Artery Disease
CAG	Invasive Coronary Artery Angiography
CASS	Coronary Artery Surgery Study
CCS	Chronic Coronary Syndrome
CCTA	Coronary Computed Tomography Angiography
CFD	Computational Fluid Dynamics
CFR	Coronary Flow Reserve
CI	Confidence Interval
CMR	Cardiovascular Magnetic Resonance
CNN	Convolutional Neural Network
COURAGE	Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation

Table of Contents

CR	Cardiac Rehabilitation
CRG	Coronariografia Invasiva
CT	Computed Tomography
CTO	Chronic Total Occlusion
CVLPRIT	Complete versus Lesion-only Primary PCI trial
CX	Circumflex Artery
DAC	Doença Arterial Coronária
DANAMI	DANish trial in Acute Myocardial Infarction
DCS	Dice Similarity Coefficient
DE	Percentagem de estenose considerando o diâmetro normal
DEACQ	Percentagem de estenose considerando o diâmetro normal por Análise Coronária Quantitativa
DEEV	Percentagem de estenose considerando o diâmetro normal por Estimativa Visual
DEFER	Deferral of percutaneous coronary intervention
DEFINE-FLAIR	Distal Evaluation of Functional performance with Intravascular sensors to assess the Narrowing Effect - Functional Lesion Assessment of Intermediate Stenosis to Guide Revascularisation
DEFINE-PCI	Distal Evaluation of Functional performance with Intravascular sensors to assess the Narrowing Effect - Physiologic Assessment of Coronary Stenosis Following PCI
DES	Drug-Eluting Stent
DICOM	Digital Imaging and Communications in Medicine
DL	Deep Learning
DNN	Deep Neural Network
DS	Percentage Diameter Stenosis
DSC	Dice Similarity Coefficient
DSQCA	Percentage Diameter Stenosis by Quantitative Coronary Analysis
DSVE	Diameter Stenosis by Visual Estimation
EAI	Enhanced Artificial Intelligence Model
ECG	Electrocardiogram
EH	Enhanced Human Model

EXCEL	XIENCE versus Coronary Artery Bypass Surgery for Effectiveness of Left Main Revascularization
FAME	Fractional Flow Reserve Versus Angiography for Multivessel Evaluation
FAMOUS	Fractional flow reserve versus angiography in guiding management to optimize outcomes
FAST	Fast Assessment of STenosis severity
FAVOR	Functional Assessment by Various Flow Reconstructions
FFR	Fractional Flow Reserve
FFRCT	Fractional Flow Reserve Computed from CT
FIRE	Functional versus Culprit-only Revascularization in Elderly Patients with Myocardial Infarction and Multivessel Disease
FLOWER	Flow Evaluation to Guide Revascularization in Multivessel ST-Elevation Myocardial Infarction
FN	False Negative
FP	False Positive
FRAME - AMI	Fractional Flow Reserve vs. Angiography-Guided Strategy for Management of Non-Infarction Related Artery Stenosis in Patients with Acute Myocardial Infarction
FUTURE	FUnctional Testing Underlying coronary REvascularization
GDS	Generalized Dice Score
GPU	Graphics Processing Unit
GSS	Global Segmentation Score
GUIDE - DES	Quantitative Coronary Angiography versus Intravascular Ultrasound GUIDancE for Drug-Eluting Stent Implantation
HA	modelo Humano Aperfeiçoada
HB	modelo Humano Basal
IA	Inteligência Artificial
IAA	modelo de Inteligência Artificial Aperfeiçoado
IAB	Intra-Aortic Balloon
IC	Interventional Cardiology
ICD	Implantable Cardioverter-Defibrillator
ICP	Intervenção Coronária Percutânea

IDEAL	Iberian-Dutch-English
IFR	Instantaneous wave-Free Ratio
IQ	Interquartile
IQR	Interquartile Range
ISCHEMIA	International Study of Comparative Health Effectiveness with Medical and Invasive Approaches
IVUS	Intravascular Ultrasound
LAD	Left Anterior Descending artery
LAO	Left Anterior Oblique
LCA	Left Coronary Artery
LIMA	Left Internal Mammary Artery
MI	Myocardial Infarction
ML	Machine Learning
MLD	Minimum Lumen Diameter
MRI	Magnetic Resonance Imaging
NLP	Natural Language Processing
NN	Neural Network
NOBLE	Nordic–Baltic–British Left Main Revascularisation Study
NPV	Negative Predictive Value
NSTEMI	Non-ST Elevation Myocardial Infarction
OCT	Optical Coherence Tomography
PA	Pulmonary Artery
PCI	Percutaneous Coronary Intervention
PDA	Patent Ductus Arteriosus
PET	Positron Emission Tomography
PLA	Postero-Lateral Artery
PPG	Photoplethysmography
PPV	Positive Predictive Value
PROMISE	Prospective Multicenter Imaging Study for Evaluation of Chest Pain
PTCA	Percutaneous Transluminal Coronary Angioplasty
QCA	Quantitative Coronary Angiography
QFR	Quantitative Flow Ratio

RAO	Right Anterior Oblique
RCA	Right Coronary Artery
ReVEAL-iFR	Radiographic imaging Validation and EvALuation for Angio-iFR
REVIVED	Revascularization for Ischemic Ventricular Dysfunction
RFR	Resting Full-Cycle Ration
RIMA	Right Internal Mammary Artery
RVD	Reference Vessel Diameter
SCAI	Society for Cardiovascular Angiography and Interventions
SCC	Síndrome Coronário Crónico
SD	Standard Deviation
SDG	Score de Dice Generalizado
SDS	Score de Dice de Similaridade
SPSS	Statistical Package for the Social Sciences
SSG	Score de Segmentação Global
ST	Segmento ST
STEMI	ST-Elevation Myocardial Infarction
STICH	Surgical Treatment for IsChemic Heart failure
SVG	Saphenous Vein Graft
SWEDEHEART	Swedish Web-system for Enhancement and Development of Evidence-based care in Heart disease Evaluated According to Recommended Therapies
SYNTAX	Synergy Between PCI with Taxus and Cardiac Surgery
TACTICS	Treat Angina with Aggrastat and Determine Cost of Therapy with an Invasive or Conservative Strategy
TAXUS	Third generation pAcliTaxel-elUting coronary Stent
TIMI	Thrombolysis In Myocardial Infarction
TN	True Negative
TP	True Positive
USA	United States of America
VERDICT	Very EaRly vs Deferred Invasive evaluation using Computerized Tomography
VN	Verdadeiro Negativo

Table of Contents

VP	Verdadeiro Positivo
VPN	Valor Preditivo Negativo
VPP	Valor Preditivo Positivo
WHO	World Health Organization

ABSTRACT

INTRODUCTION

Coronary artery disease (CAD) is the dominant cause of cardiac pathology and remains the most relevant cause of morbidity and mortality worldwide. The invasive management of CAD, with revascularization by Coronary Artery Bypass Graft (CABG) surgery or Percutaneous Coronary Intervention (PCI), was made possible by the development of invasive coronary angiography (CAG) in the 1950's and 60's.

A fundamental step in interpreting CAG is the estimating of lesion severity, as the decision to proceed with revascularization hinges upon it. Notwithstanding, this critical step has hardly changed for over 50 years. Indeed, visual estimation of percentage diameter stenosis (DS) remains the cornerstone of severity assessment, but is prone to operator dependency, resulting in potential over (or even under) estimation of severity, as multiple studies spanning five decades have shown.

The development of Quantitative Coronary Analysis (QCA) enabled the objective and reproducible assessment of lesion severity but has been understudied and underused.

To address these limitations, coronary physiology based methods were developed. Two main indexes are used: Fractional Flow Reserve (FFR) and instantaneous wave-free ratio (iFR). The former has been studied directly in outcomes studies, whereas the latter was mostly studied as a surrogate of FFR. Nonetheless, no single index is perfect, and iFR may be more suitable in certain scenarios, such tandem lesions assessment. Multiple clinical trials have confirmed the usefulness and superiority of physiology vs angiography-alone in improving revascularization decisions and hence clinical outcomes. However, their usefulness in more complex clinical scenarios, such as multivessel disease and acute coronary syndrome, has been challenged. The increased risk of iatrogeny in such contexts, together with the physiological limitations imposed by the acute setting, likely contributed to these findings. Even in the chronic coronary syndromes setting, coronary physiology remains underused.

The non-invasive derivation of coronary physiology from CAG images alone may partially overcome these limitations. Such software systems have been developed, but are semi-automatic and their reliability may be hampered by operator heterogeneity and inexperience. Furthermore, available evidence to support their use is still insufficient and almost all studies have focused solely on the derivation of FFR, not iFR. Lastly, the contribution of AI technology

in improving or surpassing these systems has seldom been explored, as most methods have used computational fluid dynamics (CFD) and three-dimensional (3D) QCA to derive physiology.

OBJECTIVES

This dissertation aims to fulfil part of the pitfalls and gaps in evidence pertaining to the interpretation of coronary angiography, exploring the potential role of AI as the primary tool for the task. The specific objectives are:

1. To create fully automatic AI models capable of accurate CAG segmentation and develop a scoring system capable of assessing their quality from a clinical perspective, in addition to those of conventional metrics.
2. To validate both the AI models and said score in a multicentric external dataset.
3. To assess the impact of AI-based segmentation models in the assessment of coronary lesions severity, namely its impact in inter-operator heterogeneity and overestimation tendency.
4. To develop AI-based validation models capable of deriving instantaneous Wave-Free Ratio (iFR) data from CAG images alone in fully non-invasive fashion.

METHODS

The following research projects were conducted to address the established objectives:

1. A retrospective single center study of randomly selected patients undergoing CAG. Per incidence, an ideal frame was segmented, forming a baseline human dataset (BH), used for training a baseline AI model (BAI). An enhanced human segmentation (EH) was created by combining the best of both. An enhanced AI model (EAI) was trained using the EH. Results were assessed by expert physicians using eleven weighted criteria, combined into a Global Segmentation Score (GSS- 0 – 100 points). The Generalized Dice Score (GDS) and Dice Similarity Coefficient (DSC) were also used for assessing the performance of AI models.

2. A retrospective observational study with four centres was conducted. Consecutive patients undergoing CAG and percutaneous coronary intervention or invasive physiology assessment over a one-month period were selected. A single frame was selected from images containing a lesion with a 50-99% DS by visual estimation ($\%DS_{VE}$). DS by Quantitative Coronary Analysis ($\%DS_{QCA}$) was performed with a validated software. Images were then segmented by the AI model. Lesion diameters, area overlap [based on true positive (TP) and true negative (TN) pixels] and the global segmentation score (GSS – 0 -100 points) - previously developed and published - were measured.
3. $\%DS_{QCA}$ was performed in the validation dataset of project 2. Operators were asked to estimate $\%DS_{VE}$ of lesions in angiography vs AI-segmented images in separate sessions and differences were assessed using angiography $\%DS_{QCA}$ as reference.
4. A retrospective single-centre observational study of consecutive patients undergoing invasive iFR. We developed AI models capable of binarily classifying target lesions as positive (iFR $\leq$ 0,89) or negative (iFR $>$ 0,89) based on a single 2D-frame. The predictions were then compared to the true invasive measurements.

RESULTS

1. 1664 processed images were generated. GSS for BH, EH, BAI and EAI were 96,9 +/- 5,7; 98,9 +/- 3,1; 86,1 +/- 10,1 and 90 +/- 7,6, respectively (95% CI, $p < 0,001$ for both paired and global differences). The GDS for the BAI and EAI was 0,9234 ± 0,0361 and 0,9348 ± 0,0284, respectively. The DSC for the coronary tree was 0,8904 ± 0,0464 and 0,9134 ± 0,0410 for the BAI and EAI, respectively. The EAI outperformed the BAI in all coronary segmentation tasks but performed less well in some catheter segmentation tasks.
2. 123 regions of interest from 117 images across 90 patients were included. There were no significant differences between lesion diameter, percentage diameter stenosis and distal border diameter between the original/segmented images. There was a statistically significant albeit minor difference [0,19 mm (0,09 – 0,28)] regarding proximal border diameter. Overlap accuracy ((TP+TN)/(TP+TN+FP+FN)), sensitivity (TP / (TP + FN)) and Dice Score (2TP / (2TP + FN + FP)) between original/segmented images was

99,9%, 95,1% and 94,8%, respectively. The GSS was 92 (87 – 96), similar to the previously obtained value in the training dataset.

3. $\%DS_{VE}$ was significantly higher in both the angiography (77% +-/ 20% vs 56% +-/ 13%, p < 0,001) and segmentation groups (59% +-/ 20% vs 56% +-/ 13%, p < 0,001), with a much smaller absolute difference in the latter. For lesions with $\%DS_{QCA}$ of 50 - 70% (60% +/- 5%), an even higher discrepancy was found (Angiography: 83% +/- 13% vs 60% +/- 5% p < 0,001; Segmentation: 63% +/- 15% vs 60% +/- 5%, p < 0,001). Similar, but less pronounced, findings were observed for $\%DS_{QCA}$ < 50% lesions, but not $\%DS_{QCA}$ > 70% lesions. Agreement between $\%DS_{QCA}$ / $\%DS_{VE}$ across $\%DS_{QCA}$ strata (< 50%, 50 - 70%, > 70%) was approximately twice in the segmentation group (60,4% vs 30,1%; p < 0,001). $\%DS_{VE}$ inter-operator differences were smaller with segmentation.
4. 250 iFR measurements were included. Three models were developed. Model 3 had the best overall performance: accuracy, NPV (NPV), PPV (PPV), sensitivity and specificity were 69%, 88%, 44%, 74% and 67%, respectively. Performance differed per target vessel. For the LAD, model 3 had the highest accuracy (66%), while model 2 the highest NPV (86%) and sensitivity (91%). PPV was always low/modest. Model 1 had the highest specificity (68%). For the right coronary artery, model 1's accuracy was 86%, NPV 97% and specificity 87%, but all models had low PPV (maximum 25%) and low/modest sensitivity (maximum 60%). For the circumflex, model 1 performed best: accuracy, NPV, PPV, sensitivity and specificity were 69%, 96%, 24%, 80% and 68%, respectively.

CONCLUSIONS

We successfully developed an AI model for CAG segmentation and validated it with data from four centers, across a wide range of stenosis severity, target vessel, operators and equipment. The models' performance was extensively assessed by three different methods, ensuring that segmentation is accurate as a whole, in stenosed regions, and also based on clinical criteria as defined by experts (i.e. Interventional Cardiologists).

Visualization of AI automatic segmentation of CAG images rendered its interpretation more reproducible across operators, significantly reducing the tendency severity overestimation. Our immediate and fully automatic segmentation model, which requires no

input but the image itself, therefore has the potential of reducing unwarranted revascularization, potentially increase the use of physiology and/or imaging and hopefully have a meaningful impact on clinical outcomes.

We successfully developed, for the first time, AI models capable of binary iFR lesion classification using a single bidimensional frame. Despite the modest accuracy of all models, the high NPV is clinically relevant, as a negative result effectively enables concluding the procedure without further testing, invasive or non-invasive. Because 60 to 70% of measurements in physiology are usually negative, this finding particularly relevant, opening a totally new patient pathway. This is indeed the first time the derivation of iFR from CAG images has been reported, namely using AI as the primary tool for the task.

These findings therefore highlight that automatic digital mapping systems of CAG images are likely warranted, as they render their interpretation more objective, paving the way for future research and potentially improving outcomes and reducing costs, so important in the current health care management.

Keywords

Artificial Intelligence; Coronary Angiography; Coronary Artery Disease; Percutaneous Coronary Intervention; Coronary Physiology.

RESUMO

INTRODUÇÃO

A doença das artérias coronárias (DAC) representa ainda a principal causa de patologia cardíaca, morbidade e mortalidade a nível global. O manejo invasivo da DAC, através da cirurgia de revascularização coronária ou intervenção coronária percutânea, foi viabilizado pelo desenvolvimento da angiografia coronária invasiva (CRG) nas décadas de 1950 e 60.

A avaliação da gravidade das lesões em CRG é fundamental, pois aí assenta a decisão de revascularização, mas pouco mudou em mais de 50 anos. A estimativa visual da percentagem de estenose considerando o diâmetro normal (DE) permanece, ainda hoje, a metodologia padrão. Porém, múltiplos estudos, ao longo de 5 décadas, mostraram que esta abordagem é propensa a variabilidade inter-operadores e a sobre (ou mesmo sub)-estimativa.

O desenvolvimento da Análise Coronária Quantitativa (ACQ) possibilitou a avaliação objetiva e reprodutível da gravidade da lesão, mas a sua base de evidência é reduzida, sendo a técnica subutilizada.

Para ultrapassar estas limitações desenvolveram-se novos métodos baseados no conhecimento da fisiologia coronária, utilizando-se predominantemente dois índices: *Fractional Flow Reserve* (FFR) e *instantaneous wave-free ratio* (iFR). O primeiro foi directamente estudado em ensaios de *outcomes*/desfechos, enquanto o segundo foi principalmente estudado em comparação com o FFR. Contudo, nenhum é perfeito e o iFR poderá ser, por vezes, mais adequado, como na avaliação de lesões em *tandem*. Múltiplos ensaios clínicos confirmaram a utilidade/superioridade da fisiologia *versus* angiografia na tomada de decisões de revascularização, com impacto em *outcomes*/desfechos. Porém, em cenários mais complexos, como doença multivaso e síndrome coronária aguda, a vantagem destas técnicas não está inteiramente estabelecida, provavelmente devido ao acrescido risco de iatrogenia, juntamente com as limitações de fisiologia impostas pelo contexto agudo. Adicionalmente, mesmo no síndrome coronário crónico, a fisiologia coronária é subutilizada.

A derivação não invasiva da fisiologia coronária a partir de imagens de CRG poderá superar parte destas limitações. Vários sistemas foram desenvolvidos, mas são semi-automáticos e a fiabilidade pode ser prejudicada pela variabilidade inter-operador e/ou inexperiência. A evidência disponível é ainda insuficiente e os estudos concentram-se quase exclusivamente na derivação do FFR - não do iFR. Adicionalmente, a contribuição da

Resumo

Inteligência Artificial (IA)para melhorar (ou superar) estes sistemas raramente foi explorada, já que a dinâmica de fluidos computacional e a AQC tridimensional têm sido o principal método utilizado nestes sistemas.

OBJETIVOS

Esta dissertação visa contribuir para preencher parte das lacunas e limitações da interpretação contemporânea da CRG, explorando o enorme potencial da IA, aqui utilizada como ferramenta primária. Para tal, definiram-se os seguintes objectivos:

1. Criar modelos de IA totalmente automáticos para segmentação precisa de imagens de CRG, desenvolvendo um sistema de *scoring* de avaliação de qualidade numa perspectiva clínica, não somente por métricas convencionais.
2. Proceder a validação externa multicêntrica dos modelos de IA e do sistema de *scoring*
3. Avaliar o impacto dos modelos de segmentação baseados em IA na apreciação de lesões pelos operadores, aferindo o seu impacto na variabilidade inter-operador e tendência para sobre-estimativa
4. Desenvolver modelos de IA capazes de derivar não invasivamente dados de iFR, partindo somente de imagens de CRG

MÉTODOS

Para alcançar os objectivos supra-expostos, desenvolveram-se os seguintes projectos:

1. Um estudo observacional retrospectivo unicêntrico de doentes submetidos a CRG, selecionados aleatoriamente. Por cada incidência, segmentou-se um *frame* ideal, formando uma segmentação humana basal (HB), usada para treinar um modelo de IA basal (IAB). Da combinação de ambos formou-se uma segmentação humana aperfeiçoada (HA), utilizada para treinar um modelo de IA aperfeiçoado (IAA). Os resultados foram aferidos com 11 critérios balanceados definidos por peritos, combinados num Score de Segmentação Global (SSG – 0 – 100 pontos). Aplicaram-se o Score de Dice Generalizado (SDG) e Score de Dice de Similaridade (SDS) aos modelos de IA.

2. Um estudo observacional retrospectivo multicêntrico (quatro centros). Selecionaram-se doentes consecutivos submetidos a ICP ou fisiologia invasiva durante um período de um mês. Por caso, selecionou-se um *frame* ideal onde constasse uma lesão com DE de 50-99% por estimativa visual (%DEEV). Procedeu-se a aferição de DE por ACQ ($\%DE_{ACQ}$) com um software validado. As imagens foram posteriormente segmentadas pelo modelo de IA. Foram analisados os diâmetros da lesão, área de sobreposição [baseada em pixels positivos verdadeiros (VP) e negativos verdadeiros (VN)] e Score de Segmentação Global (SSG – 0 – 100 pontos).
3. A $\%DE_{ACQ}$ foi medida no conjunto de dados de validação do projeto 2. Os operadores estimaram as lesões através %DEEV em imagens de angiografia *versus* imagens segmentadas por IA, em sessões separadas, avaliando as diferenças, tomando a $\%DE_{ACQ}$ como referência.
4. Um estudo observacional retrospectivo unicêntrico de doentes consecutivos submetidos a medição de iFR. Desenvolveram-se modelos de IA capazes de classificar binariamente lesões-alvo como positivas (iFR $\leq$ 0,89) ou negativas (iFR > 0,89), comparando as previsões dos modelos com medições reais.

RESULTADOS

1. Geraram-se 1664 imagens processadas. Os SSG para a HB, HA, IAB e IAA foram 96,9 +/-5,7; 98,9 +/- 3,1; 86,1 +/- 10,1 e 90 +/- 7,6, respectivamente (IC 95%, $p < 0,001$ - diferenças globais e emparelhadas). O SDG para o IAB e IAA foi 0,9234 ± 0,0361 e 0,9348 ± 0,0284, respectivamente. O SDS foi 0,8904 ± 0,0464 e 0,9134 ± 0,0410 para o IAB e IAA, respectivamente. O IAA exibiu superior desempenho ao IAB para todas as tarefas de segmentação coronária, mas não para todas as de cateter.
2. Incluiram-se 123 regiões de interesse de 117 imagens (90 doentes). Não se observaram diferenças significativas entre o diâmetro da lesão, o $\%DE_{ACQ}$ e o diâmetro do bordo distal entre imagens originais e segmentadas. Verificou-se uma ligeira diferença estatisticamente significativa [0,19 mm (0,09 – 0,28)] nas medições do bordo proximal. A intersecção sobre sobreposição ((TP+TN)/(TP+TN+FP+FN)), sensibilidade (TP / (TP + FN)) e Dice Score (2TP / (2TP + FN + FP)) foram de 99,9%, 95,1% e 94,8%, respectivamente. O SSG foi de 92 (87 – 96), semelhante ao valor anteriormente obtido nos dados de treino.

3. O %DEEV foi significativamente maior no grupo de angiografia (77% +-/ 20% vs 56% +-/ 13%, p < 0,001) vs segmentação (59% +-/ 20% vs 56% +-/ 13%, p < 0,001), com uma diferença absoluta de %DEEV muito inferior no segundo. Para lesões com $\%DE_{ACQ}$ de 50 a 70% (60% +/- 5%), observou-se uma discrepância ainda maior (Angiografia: 83% +/- 13% vs 60% +/- 5% p < 0,001; Segmentação: 63% +/- 15% vs 60% +/- 5%, p < 0,001). Observaram-se resultados semelhantes, mas menos pronunciados, para lesões com $\%DE_{ACQ}$ < 50%, mas não para lesões com $\%DE_{ACQ}$ > 70%. A concordância entre $\%DE_{ACQ}$ / %DEEV entre estratos de %DSQCA (< 50%, 50 - 70%, > 70%) foi aproximadamente o dobro no grupo de segmentação (60,4% vs 30,1%; p < 0,001). As diferenças inter-operadores de %DEEV foram menores com segmentação vs angiografia.
4. Incluiram-se 250 medições de iFR e desenvolveram-se três modelos. O modelo 3 teve o melhor desempenho geral: a acuidade, VPN (Valor Preditivo Negativo), VPP (Valor Preditivo Positivo), sensibilidade e especificidade foram 69%, 88%, 44%, 74% e 67%, respectivamente. Houve diferenças consoante o vaso-alvo. Para a descendente anterior, o modelo 3 exibiu a maior acuidade (66%), enquanto no modelo 2 se observou o melhor VPN (86%) e sensibilidade (91%). O VPP foi sempre reduzido. O modelo 1 exibiu a melhor especificidade (68%). Para a artéria coronária direita, observou-se uma acuidade de 86%, VPN 97% e especificidade 87% com o modelo 1, pese embora todos os modelos tenham exibido um VPP reduzido (máximo de 25%) e um grau de sensibilidade apenas moderado (máximo de 60%). Para a artéria circunflexa, o modelo 1 exibiu o melhor desempenho: a acuidade, VPN, VPP, sensibilidade e especificidade foram 69%, 96%, 24%, 80% e 68%, respectivamente.

CONCLUSÃO

Desenvolveu-se um modelo de IA capaz de segmentação de CRG, validado em quatro centros, com vasta abrangência de gravidade de estenose das artérias coronárias, vaso-alvo, operadores e equipamentos. O desempenho foi extensamente avaliado por três métodos diferentes, demonstrando-se boa acuidade globalmente, em regiões de estenose, e ainda com base em critérios definidos por peritos (i.e. Cardiologistas de Intervenção).

A segmentação de CRG por IA tornou a sua interpretação mais reprodutível, reduzindo significativamente a tendência para a sobre-estimativa da severidade de lesões. O nosso modelo de segmentação imediata, totalmente automática, não carecendo de nenhum *input* para além da imagem, poderá potencialmente reduzir a revascularização desnecessária, influenciando os operadores a recorrer a fisiologia e/ou imagem, com potencial impacto clínico.

Pela primeira vez, desenvolveram-se modelos de IA capazes de classificar binariamente lesões com base em iFR, usando um único *frame* bidimensional. Não obstante a modesta acuidade de todos os modelos, o elevado VPN observado é muito relevante clinicamente, e permitiria o término do procedimento sem testes adicionais, invasivos ou não. Dado que 60 a 70% das medições que usam fisiologia são geralmente negativas, este achado é particularmente relevante. Esta é, pois, a primeira vez que a derivação do iFR a partir de imagens de CRG é reportada, usando IA como ferramenta principal.

Estes achados sugerem que o desenvolvimento de sistemas automáticos de mapeamento digital de imagens de CRG se justifica, podendo tornar a sua interpretação mais objetiva, abrindo caminho para investigação futura, com expectável melhoria de desempenho, resultados clínicos e redução de custos.

Palavras-Chave

Inteligência Artificial; Coronariografia; Doença Arterial Coronária; Intervenção Coronária Percutânea; Fisiologia Coronária.

CHAPTER 1: BACKGROUND

1. Invasive assessment and management of coronary artery disease: the seminal role of coronary angiography from its origins to today

1.1. Coronary artery disease: a persistent burden

Coronary artery disease (CAD) is most commonly caused by atherosclerotic disease and broadly encompasses two clinical presentations: acute and chronic coronary syndromes. The World Health Organization (WHO) data places CAD as the greatest cause of mortality and loss of Disability Adjusted Life Years in the world[1]. CAD is the most frequent cause of cardiovascular mortality[2]. In Europe, CAD is also the leading cause of cardiovascular morbidity and mortality. The prevalence of cardiovascular disease is approximately 113 million (with a small decline in the past 30 years), with an annual incidence of 12,7 million. Currently, cancer is rivaling cardiovascular disease in many countries in Europe as the leading cause of death[3].

Preventive and therapeutic approaches to CAD encompass a healthy lifestyle, pharmacotherapy and revascularization. The routine availability of the latter was made possible due to several pioneering steps in the history of Medicine. To proceed with percutaneous coronary revascularization, one must establish percutaneous access, catheterize the patient, perform angiography of the coronary arteries and, finally, proceed with coronary intervention. Multiple seminal steps in history were key for todays' developments.

1.2. The development of invasive Coronary Angiography

Werner Forssmann, a German Physician, aimed to develop a technique for injecting drugs directly in the heart with a minimally invasive method. He performed the first cardiac catheterization on himself in 1929, by inserting a urethral catheter (he went on to specialize in Urology) in his antecubital vein and advancing it until his right atrium[4], having co-won the Nobel Prize for Physiology or Medicine in 1956.

Sven Seldinger, a Swedish Radiologist, is well known for the invention of the Seldinger technique in 1953, which consists of advancing a wire into a vessel through a needle, using the wire as a guide for placing a catheter, thereby effectively establishing a

truly percutaneous access[5]. Prior to his achievement, surgical cutdown of an access was a necessary requirement for introducing catheters large enough for angiography.

The visualization of cardiovascular structures by opacification with contrast and simultaneous use of X-ray owes a great deal to several Portuguese Physicians, who were truly pioneers of the field. Egas Moniz and Reynaldo dos Santos performed the first carotid angiography in 1927[6]. The latter, together with Augusto Lamas and José Pereira Caldas inaugurated aortography and peripheral arterial angiography in 1929[7]. Egas Moniz and Lopo de Carvalho visualized the pulmonary circulation in 1931[6]. João Cid dos Santos, the son of Reynaldo dos Santos, performed the first in human phlebography in 1938[6].

Coronary angiography was also pioneered by a Portuguese Physician. Eduardo Coelho, among many other notable achievements, performed the first non-selective *in vivo* coronary angiography in 1952 at the Santa Marta Hospital in Lisbon[8], despite seldom receiving the recognition he deserves, as he is rarely cited[6,8,9]. J.W. Hurst, of the Emory School of Medicine, wrote[8]: "Sones should be credited for developing the technique and moving on to his clinical usage. The volume of work he did in Cleveland led to its general acceptance (1958). E. Coelho should be credited with originating the technique."

Frank Mason Sones Jr performed the first selective coronary angiography in vivo in 1958 at the Cleveland Clinic in Ohio, USA. As is often the case in the History of Medicine and Science, engagement of a catheter in the ostium of the right coronary artery happened accidentally as, after a ventriculography, he was retracting the catheter into the aorta to perform an aortography (Figure 1). As asystole occurred, probably due to occlusion of the artery due to the caliper of the catheter, 30 mL of contrast were directly injected into the coronary artery. After catheter removal, flow ensued, and the patient recovered. Sones further developed the technique as he contributed to the development of the C-arm, enabling lateral and cranio-caudal views, essential for properly appreciating coronary artery anatomy and disease. His foundational work was widely praised from the lately 1950's onwards, and the idea of visualizing the coronary arteries with selective contrast injection with multiple incidences remains fundamentally unchanged to this day. The one issue where Sones' technique has completely changed was how access was established, as he used the brachial artery via surgical arteriotomy, a rudimentary approach by today's standards[10].

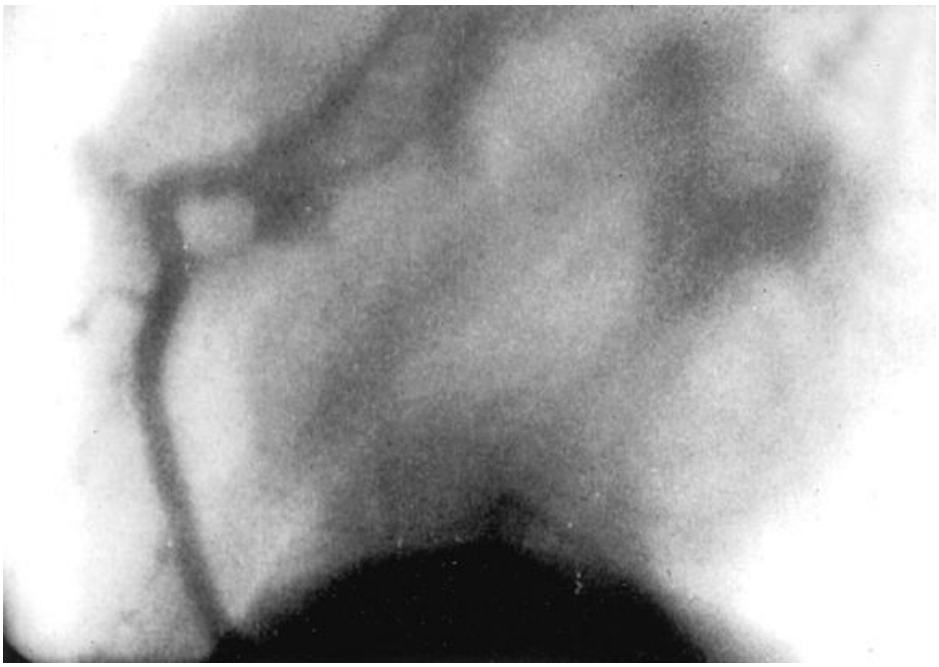


Figure 1: The first image of a selective coronary angiography by Sones in 1958 [10]**.**

Melvin Judkins trained with Sones and realized that his technique had two major pitfalls. On one hand, Sones access route was not "truly" percutaneous, as mentioned. On the other, no specific catheters with dedicated curves for facilitating access to the coronary ostia were available. Judkins developed Teflon coated J wires placed distal to the catheters for securely advancing them, while also developing catheters specifically curved for the aortic anatomy[11]. To this day, the Judkins technique – both the catheters named after him, as well as use of J-wire guidewires distal to the catheter, remain the cornerstone of a safe, easy and quick coronary angiography, which can be performed by in less than 10 minutes by experienced hands. Despite having trained in Lund (Sweden) with Stig Radner, who pioneered aortic aortography from the radial artery in 1958[12], Judkins used the transfemoral percutaneous approach, and the dedicated curves of the Judkins catheters clearly reflect this. Indeed, while his catheters can also be easily used via the radial approach, femoral use of these catheter curves is truly seamless, as even fellows with little experience are able to perform a coronary angiography under supervision. Cordis began commercializing the Judkins catheters in 1968, and today every medical company offers the Judkins catheters in their portfolio.

1.3. Revascularization: a revolution in the treatment of coronary artery disease

The availability of CAG spawned the development of revascularization therapies. First, surgical procedures were developed. Direct anastomosis of left internal mammary artery (LIMA) into the left ventricle wall was invented by Arthur Vineberg in 1946[13], prior to the advent of CAG images. But the first anastomosis to a coronary was performed in 1962 by Robert Goetz in New York, who anastomosed the right internal mammary artery (RIMA) to the RCA[14]. In the following years, saphenous venous grafts (SVG) were developed as well[15,16] and in 1968 George Green performed the first LIMA to LAD anastomosis[17], which since then became the gold standard for CABG.

The development of CAG also led to the emergence of percutaneous intervention. Once minimally invasive selective coronary access became established, it was only a matter of time until percutaneous intervention of the coronaries themselves would ensue. Charles Dotter performed the first transluminal angioplasty. In 1963, he inadvertently recanalized an occluded right iliac artery with a diagnostic catheter, only to find it remained patent after catheter removal. After partnering with William Cook, a dilatation kit was developed, and the first intentional angioplasty was performed in a femoral artery in 1964[18]. Despite his success, his approach was quite controversial at the time. His findings led to the concept of "the Dotter effect", whereby plaques were found to be compressible by dilatation, a concept which still underlies much of the percutaneous vascular intervention today. Dotter did not, however, perform coronary interventions, nor did he apply balloon catheters, despite having worked on its concept.

Eberhard Zeitler went on to further develop Dotter's technique in Germany during the late 1960's and early 1970's and gained significant recognition in the medical community[19]. Later on, with Zeitler as mentor, Andreas Grüntzig developed a single-lumen catheter balloon in 1974 and a double lumen device in 1975 (to enable distal perfusion while inflating), which were used for peripheral artery disease. After several experiments in dogs, then 38-year-old Grüntzig successfully performed the first transluminal percutaneous coronary angioplasty (PTCA) in Zurich, Switzerland, in the left anterior descending artery of an also 38-year old man, using a double lumen catheter balloon (Figure 2). After presenting his first results at a poster session, he published the results of the first 50 cases[20]. Due to this seminal step, he became known as the father of Interventional Cardiology, as he effectively gave birth to the immensity that is the field today.

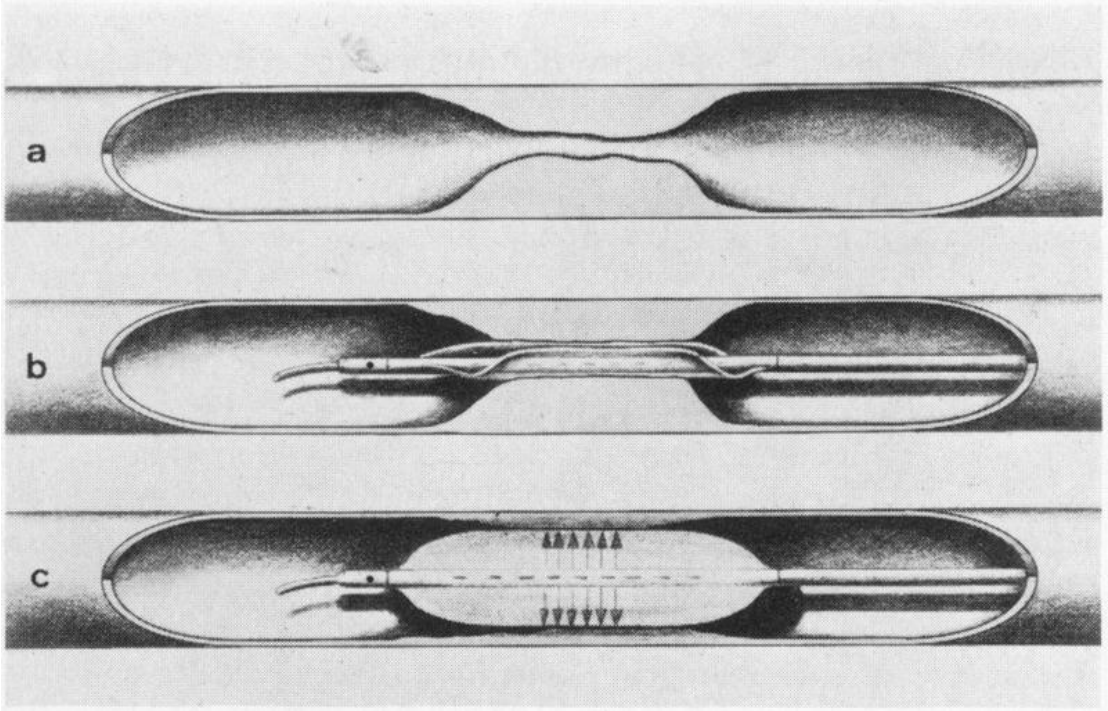


Figure 2: schematic depiction of Grüntzig's double lumen angioplasty balloon. From Grüntzig *et al*[20].

A true revolution in percutaneous coronary revascularization techniques ensued, which Grüntzig, tragically dead after a plane accident in 1985, would not come to witness. In the early 1980's, the use of guiding catheters, together with the development of coronary guidewires (Grüntzig's balloon was advanced directly into the coronaries) made device delivery possible into virtually any coronary segment - John Simpson played a key role in their development[21]. The creation of the monorail systems by Tassilo Bonzel enabled rapid and easy exchange of coronary devices in 1986[22]. The development of first bare-metal stents (the self-expanding Wasllstent[23] and the balloon-expandable Palmaz-Schatz[24]) enabled the treatment and prevention of iatrogenic dissection, abrupt vessel closure and recoil. Its' benefits were established in a major randomized trial[25], but were hampered by restenosis. At the turn of the century, J. Eduardo Sousa presented the first results regarding the use of drug-eluting stents with an antiproliferative drug (sirolimus) in a series of 30 patients[26]. Two major clinical trials further confirmed the advantage of these new stent types[27], but the devices' very low rate of restenosis were limited by the occurrence of late-stent thrombosis. This limitation was overcome in the last 10-15 years with many second-generation DES with very thin struts and polymers. Numerous trials[28,29 30,31 32,33], have rendered the DES the default strategy for all PCI, effectively relegating BMS towards history, as clearly reflected in current guidelines[34].

Lastly, the development of percutaneous coronary therapeutic techniques also enabled the development of new diagnostic tools. Intra-coronary imaging, by means of Intravascular Ultrasound (IVUS)[35] or Optical Coherence Tomography (OCT)[36] emerged in the 1980's and 1990's, respectively. Ever since, a large body of evidence has confirmed PCI clinical outcomes can be improved with both IVUS[37–39] and OCT[39–42] with a potential impact on mortality, while also adding complementary data regarding the insights provided by CAG alone.

Invasive coronary physiology was another breakthrough in the diagnostic assessment of coronary artery disease and will be the subject of a dedicated subchapter.

Today, revascularization plays a vital role in the management of coronary artery disease. In the USA, over 600000 PCI and 200000 CABG procedures are performed annualy[43,44]. In Europe, these numbers account for approximately 5000 invasive diagnostic and 2500 PCI per million people[45]. In Portugal, approximately 14000 PCI procedures are performed every year[46].

In acute coronary syndromes, a routine invasive approach has been demonstrated to improve outcomes. In non-ST elevation myocardial infarction (NSTEMI), landmark trials such as the TACTICS-TIMI[47], TIMACS[48]or VERDICT[49] have clearly shown that a routine invasive strategy should be employed, ideally in the first 24 hours for many cases. In ST elevation myocardial infarction (STEMI), primary PCI has been the default strategy for approximately 20 years now, since the publication of one of the most important trials in the field -DANAMI-2[50]. More recent trials have also shown that a complete revascularization strategy should be pursued, for both STEMI[51] and (with less evidence) NSTEMI[52,53] patients. And indeed, since the introduction of such techniques over the last 30 years, the mortality of acute coronary syndromes, especially short-term (where the relative reduction is approximately 80%), has dropped dramatically[51,54].

In chronic coronary syndromes, the issue of revascularization remains controversial. While the vast majority of trials have shown symptomatic improvement even in the era of contemporary anti-anginal drugs[49,55], a great deal of debate is commonplace in the Cardiology community. The COURAGE[56] trial and ISCHEMIA[49] trial were formally negative trials. However, even in the latter, revascularization by either CABG or PCI reduced the occurrence of spontaneous myocardial infarction by 33%. The FAME-2 trial[57] also compared PCI (guided by physiology) to medical therapy and a similar reduction in the rate of spontaneous myocardial infarction was noted as well.

Thus, whatever pitfalls remain, it is clear that revascularization is here to stay[34,58]. The pioneers of coronary angiography made it all possible, as without a clear delineation of coronary anatomy and lesions, both reliable diagnosis and treatment would not be possible.

2. Interpreting the coronary angiography: a persistent conundrum

Currently, CAG images are acquired in specialized catheterization laboratories, where dedicated X-ray equipment, digital imaging systems and emergency support are available. CAG can be undertaken electively if chronic coronary syndrome is suspected, or urgently/emergently in the case of acute coronary syndromes. Despite its age and simplicity as compared to more recent techniques (such as CCTA or CMR), it remains the cornerstone of the diagnostic management of coronary artery disease, as no decisions regarding revascularization are undertaken without this prior fundamental step.

The purpose of a CAG is to (1) characterize the coronary anatomy and (2), if present, coronary artery disease, as enabled by X-ray fluoroscopy and contrast luminography. Further invasive or non-invasive testing, as well as revascularization, may ensue, either *ad hoc* or electively.

2.1. Coronary Anatomy Characterization

2.1.1. Coronary anatomy and classification systems

The left coronary artery arises from the left sinus of Valsalva, consisting of a short main trunk bifurcating into the left anterior descending artery (LAD) and circumflex (Cx). The LAD is often classified per its length into the uncommon type I (only runs two thirds of the distance between the base and the apex) and the common types II (reaches the apex) and III (runs over the apex). Collateral diagonal branches supply the antero-lateral wall, while septal branches the anterior part of the interventricular septum. The Cx travels the left auriculo-ventricular sulcus, emitting obtuse marginal branches for the lateral and posterior walls[59].

The right coronary artery (RCA) usually arises from the right sinus of Valsalva, traverses the right auricular ventricular sulcus, giving off marginal branches that supply the right ventricle, while also supplying the right ventricular outflow track (conus branch), the sinus node and much of the atrio-ventricular node. If dominant (85% of cases), it bifurcates into a crux, issuing the posterior-lateral (PLA) branch and the posterior descending artery (PDA), irrigating part of the posterior-lateral walls and the inferior septum, respectively. If the Cx is dominant, the PDA arises from this vessel, and the RCA is usually a vessel small in both caliper and length[59].

Background

The coronary arteries are often divided into segments. An initial division of the American Heart Association (AHA) encompassed 14 segments[60] (Figure 3). Later on, the Coronary Artery Surgery Study (CASS) classification divided the coronaries into 27 segments[61], while a further two (Figure 4, Table 1) were defined in the Bypass Angioplasty Revascularization Investigators (BARI)[62]. The latter classification was also formally adopted by the American College of Cardiology (ACC), AHA and the Society for Cardiac Angiography and Interventions (SCAI)[63].

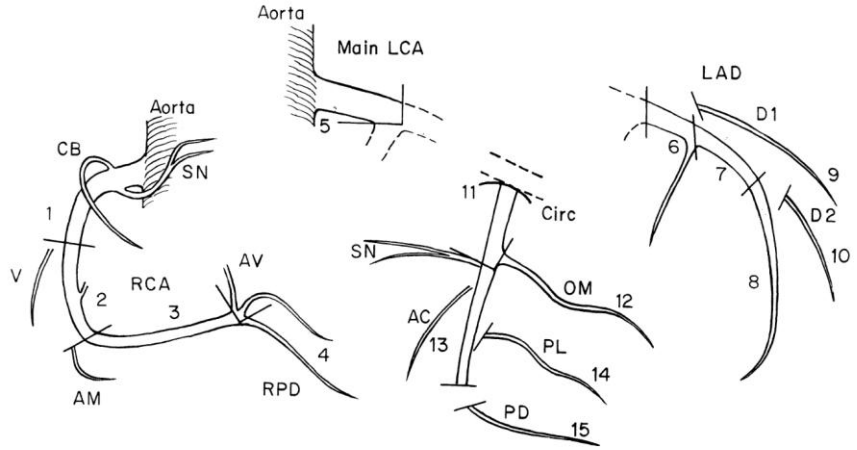


Figure 3: the American Heart Association coronary classification with 14 segments (1975) [60]**.**

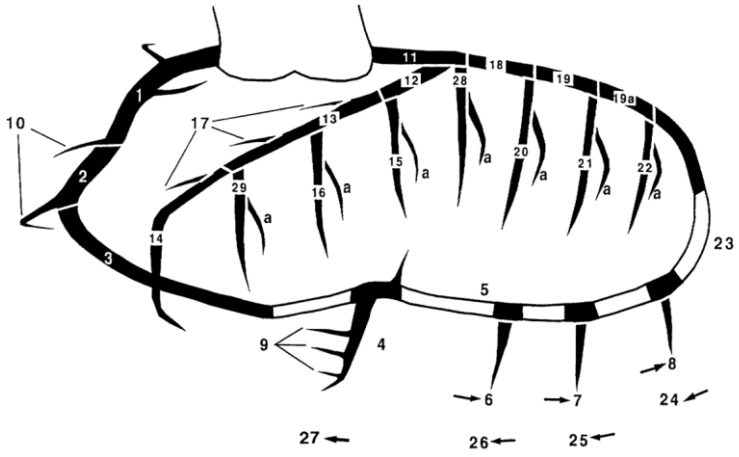


Figure 4: schematic of the 29 coronary segments classification from the BARI trial [62]**.**

Table 1: the BARI 29 segment coronary classification [62]**.** AV: atrioventricular; LAD: left anterior descending.

Segment	Map Location	Segment	Map Location	Segment	Map Location
1	Proximal right coronary artery conduit segment	13	Mid-LAD artery segment	21	Second obtuse marginal branch segment
2	Mid-right coronary artery conduit segment	14	Distal LAD artery segment	21a	Lateral second obtuse marginal branch segment
3	Distal right coronary artery conduit segment	15	First diagonal branch segment	22	Third obtuse marginal branch segment
4	Right posterior descending artery segment	15a	Lateral first diagonal branch segment	22a	Lateral third obtuse marginal branch segment
5	Right posterior atrioventricular segment	16	Second diagonal branch segment	23	Circumflex artery AV groove continuation segment
6	First right posterolateral segment	16a	Lateral second diagonal branch segment	24	First left posterolateral branch segment
7	Second right posterolateral segment	17	LAD septal perforator segments	25	Second left posterolateral branch segment
8	Third right posterolateral segment	18	Proximal circumflex artery segment	26	Third posterolateral descending artery segment
9	Posterior descending septal perforators segment	19	Mid-circumflex artery segment	27	Left posterolateral descending artery segment
10	Acute marginal segment(s)	19a	Distal circumflex artery segment	28	Ramus intermedius segment
11	Left main coronary artery segment	20	First obtuse marginal branch segment	28a	Lateral ramus intermedius segment
12	Proximal LAD artery segment	20a	Lateral first obtuse marginal branch segment	29	Third diagonal branch segment
-	-	-	-	29a	Lateral third diagonal branch segment

In addition to substantial “normal” variability, multiple pathological congenital coronary abnormalities exist, particularly regarding their origin and anatomical course. Briefly, there may be a single coronary artery originating from either the right or left coronary sinus; anomalous origin of the left main coronary artery from the right aortic sinus, which may then take an anterior, posterior or inter-arterial course with regards to the pulmonary artery (an intra-septal course is also possible); anomalous origin of the left circumflex from the RCA or the right coronary sinus, passing behind the coronary sinus; anomalous origin of the RCA from the left coronary sinus; and rarely, the anomalous origin of the LCA from the pulmonary artery (ALCAPA).

Lastly, coronary fistulas and intra-coronary bridging may sometimes also be appreciated during CAG.

2.1.2. Optimal angiographic views

To fully appreciate the coronary arteries, multiple orthogonal projections are necessary. Despite minor variations across sites, usually a minimum of four projections for the left coronary artery and three projections for the right coronary artery are undertaken. These projections have been long established and are described in detail in specialty textbooks[59,64].

For the LCA, a left caudal view (left 60°-70°, caudal 25°-40°, the so called "spider" view because of its resemblance with this arthropod) enables clear visualization of the left main and the emergence of the left anterior descending (LAD) artery and the circumflex (Cx). The proximal and mid-segments of the latter are also reasonably well individualized in this projection.

A right caudal view (right 10º – 20º, caudal 30º – 40º) is complementary for the appreciation of the remainder of the circumflex artery and obtuse marginal branches, which are properly elongated in this view. Some proximal LAD lesions are well visualized in this projection as well.

A right cranial view (right 10º – 20º, cranial 30º – 40º) is of particular use for an elongated view of the LAD, usually enabling a full appreciation of its anatomy, particularly for the mid and distal segments. At times, a more rightward (~30º) projection may be of use. Lastly, the left cranial view (30°-45° left, 25°-40°cranial) also elongates the mid-distal segments of the LAD, while often enabling clear appreciation of the ostial and proximal segments of diagonal branches. In cases of left dominance, the crux and posterior descending artery are clearly demarcated in cranial projections as well. Furthermore, the relationship between the left main ostium and the sinus of Valsalva may also be clearly visible in cranial projections, which is of particular importance for left main ostial lesions and interventions. Lastly, cranial and caudal postero-anterior (PA) projections may be of use sometimes.

For the RCA, a simple left anterior oblique view (30º – 40º) is usually ideal for cannulation and clearly displays the middle segment of the artery. A cranial angulation (~30º) elongates the full vessel, clearly separating the crux and its bifurcation (i.e. the postero-lateral and posterior descending branches). At times, a postero-anterior cranial angulation separates the crux bifurcation even better. Lastly, a simple right anterior oblique view elongates the posterior descending artery and its septal branches, while clearly demarcating marginal branches emerging from the main vessel. Sometimes, full appreciation of ostial lesions requires a "spider"-like projection as well.

2.2. Assessment of Coronary Artery Disease with Invasive Coronary Angiography

When characterizing coronary artery disease as viewed in CAG, coronary flow is immediately appreciated. In the setting of myocardial infarction, especially STEMI, flow is often classified as per the Thrombosis in Myocardial Infarction (TIMI) classification[65] (Table 2). In the setting of chronic coronary syndromes and/or non-culprit lesions, the absence of flow is only encountered in chronic total occlusions. In such cases, flow originating from collaterals is usually described as per the Rentrop classification: grade 0 implies no filling; grade 1, small side branches filling; grade 2, partial epicardial filling of the occluded artery; grade 3, complete epicardial filling of the coronary artery[66].

Table 2: Thrombosis in Myocardial Infarction (TIMI) classification [65].

Grade	Description
0	(no perfusion) There is no antegrade flow beyond the point of occlusion.
1	(penetration without perfusion) The contrast material passes beyond the area of obstruction but "hangs up" and fails to opacify the entire coronary bed distal to the obstruction for the duration of the cineangiographic filming sequence.
2	(partial perfusion) The contrast material passes across the obstruction and opacifies the coronary bed distal to the obstruction. However, the rate of entry of contrast material into the vessel distal to the obstruction or its rate of clearance from the distal bed (or both) are perceptibly slower than its entry into or clearance from comparable areas not perfused by the previously occluded vessel — e.g., the opposite coronary artery or the coronary bed proximal to the obstruction.
3	(complete perfusion) Antegrade flow into the bed distal to the obstruction occurs as promptly as antegrade flow into the bed proximal to the obstruction, and clearance of contrast material from the involved bed as rapid as clearance from an uninvolved bed in the same vessel or the opposite artery.

In addition to assessing flow, operators identify and characterize each individual lesion and then consider their aggregate. When doing so, three major steps should be undertaken: (1) classify lesion severity/significance; (2) characterize lesion morphology and complexity; (3) combine both findings to fully appreciate the scope of coronary artery disease in a given patient.

Establishing the severity/significance of a lesion or set of lesions has several goals. First, from an interventionalist perspective, it fundamentally leads to the decision whether revascularization should be considered or is unnecessary. Furthermore, regardless of revascularization, the presence of significant lesions is associated with an increased risk of cardiovascular events and thus a worse prognosis, rendering optimal medical therapy for the prevention of cardiovascular events and, if present, symptom (i.e. angina or equivalent) relief, mandatory[67,68].

Background

The morphology and complexity of lesions is, once again from an interventionalist perspective, primarily relevant for considering the risks and probability of success of revascularization techniques.

The combination of both of the above is key in deciding the modality of revascularization (i.e. CABG, PCI or hybrid). Furthermore, in the presence of impaired left ventricle systolic function, the full appreciation of coronary artery disease may partially or completely explain such a finding, while also significantly influence the decision to implant cardioverter-defibrillator (ICD) devices, which are of particular use in ischemic dilated cardiomyopathy vs non-ischemic dilated cardiomyopathy[69,70].

Multiple methods and scores are often used for attaining the goal of fully characterizing coronary lesions in a patient based on CAG images. The following sections outline the most relevant and commonly employed ones.

2.2.1. Assessment of coronary artery disease severity

In both clinical practice and many clinical trials, the severity of a lesion is primarily assessed as the percentage diameter stenosis (DS). The operator assesses the diameter reduction in a stenosed region, as compared with a normal reference segment, by visual estimation. Because no exact measurement is actually undertaken, this method is prone to operator heterogeneity. Indeed, prior to the development of an objective measurable method of estimating stenosis severity, significant intra-observer and inter-observer variability had already been reported in relatively small studies from the 1970's, even when using expert panels' assessment as reference[71,72].

This limitation led to the development of Quantitative Coronary Analysis (QCA). This method consists of measuring the actual diameters of a vessel, effectively calculating the severity of a stenosis. An ideal end-diastolic frame, where perfusion is maximal and coronary opacification usually best, is used. The diameters of the vessel are calculated based on the distance between the centerline and the lumen contour. The true (rather than lesion-luminal) vessel diameter can also be estimated, thereby enabling the calculation of the reference diameter as well. This can be done by either averaging the diameters of adjacent normal regions, considering their largest diameter, or using a more complex interpolation method with regression methods (a straight tapering line is obtained)[73,74]. Calibration can be performed either by using a specific part of the image as reference (such as the catheter caliper), or, in the digital era, automatically, given that image scale is present in the metadata of current imaging systems.

Conducting such measurements prior to the routine availability of computers or digital methods was, however, a significant limitation in the early decades of CAG. However, the first publication regarding this method dates back to the 1970's, demonstrating its feasibility[75]. The development of QCA using the above-mentioned methods saw significant expansion and improvement in further decades, particularly from the 1980's and 1990's onwards. Enhancements in image quality and digital systems has led to the creation of QCA software that is available both online and offline, is highly reproducible and can provide measurements for diameter values with less than 1 mm, which is of relevance given the small diameter of the coronary arteries[73,74,76–78]. Furthermore, current software has been developed to address more complex regions, such as bifurcations[77,79–82], as well as provide 3D reconstructions of the coronary artery[79,80,83,84], which may be of particular relevance in eccentric lesions.

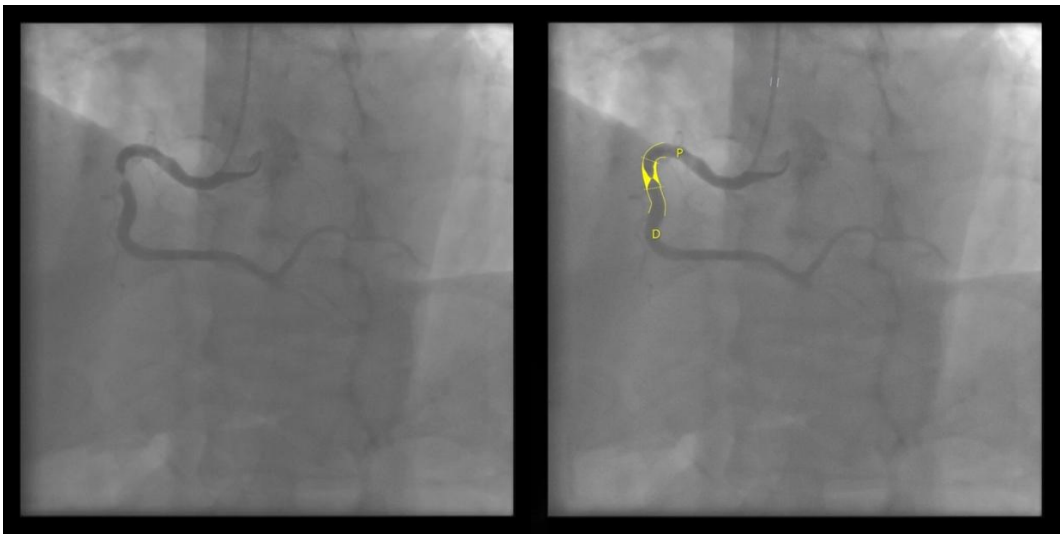


Figure 5: an example of modern day QCA software in a right coronary artery, with semi-automatic measurements. Calibration in blue (catheter). Targeted region in yellow. From: Unidade de Cardiologia de Intervenção Joaquim Oliveira. QCA: Quantitative Coronary Angiography.

Current QCA calculations are fully automatic and operators do not need concern themselves with software technicalities. However, the process of actually obtaining measurements itself is semi-automatic, as operators still have to provide some level of manual input regarding the vessel pathway, while also sometimes adjusting the vessel contour manually. Nevertheless, current software can be easily and quickly used (Figure 5). Whatever the process, by comparing the estimated reference vessel diameter (RVD) with the luminal diameter, minimal lumen diameter (MLD) and DS can be calculated as: DS = (1 – MLD) / RVD.

Background

The appearance of QCA provided an objective means of quantifying DS, thereby enabling the possibility of clearly assessing operators' performance in their visual estimation of DS. Since then, multiple studies, ranging from the 1990's to relatively recent ones[85–90], across different hospitals and countries have consistently shown two major findings: there is significant heterogeneity in operator's visual assessment of DS; operators tend to overestimate the severity of lesions.

In 1990, Beauman *et al* conducted such an assessment with just 13 lesions. They found that despite the reproducibility of QCA, visual interpretations of DS were quite inaccurate (R = 0,78; SD 14,5%), as lesion severities of just 50% DS as measured by QCA were estimated visually by operators ranging from 30 to 95%. Importantly, this did not improve with operator experience[87]. More than a decade later, in 2002, Fischer *et al* found low rates of agreement (55%) across 83 lesions and 3 operators[85].

Much more recently, a large multicentric Chinese study was undertaken, with 1548 lesions, all of which were treated by PCI. This is of particular relevance, because it focused on lesions deemed as significant by operators, where disagreement might be anticipated to be lower, as stenosis would be visually more impressive. The assessment of the significance of lesions was carried out by visual estimation in approximately 90% of cases and further analysis (physiology, IVUS or imaging) in the remainder, further emphasizing how DS estimation by visual estimation remains the cornerstone of revascularization decisions all over the world. Zhang *et al* found that operator estimation of DS was, in absolute terms, 16% higher than QCA in non-myocardial infarction patients and 10% in cases of myocardial infarction. Importantly, more than half (50,6%) of lesions assessed has > 70% DS visually had a QCA DS < 70%, clearly depicting the operators' tendency for overestimation. There were also significant differences across 57 physicians and 30 hospitals[86].

The above-mentioned findings, which are consistent over a large period of time (where improvements in image quality and digital systems have ensued), with increasing numbers of lesions, physicians and hospitals, therefore suggest that visual estimation of DS method is inherently prone to error.

Given the limitations of DS visual estimation and the high reproducibility and objectivity of QCA, one might expect it to have become the standard in CAG-based lesion severity estimation. However, that has not been the case. No data is available regarding the implementation of routine QCA in everyday clinical practice, i.e. how much of DS is based on visual estimation vs on QCA measurements. However, the very absence of such data

speaks to its likely very low usage. Furthermore, as a defendant of this thesis, over a 10-year period in catheterization laboratories across multiple countries and realities, my experience is that QCA generally is not routinely used, quite the opposite.

Moreover, major clinical trials for several decades have used DS based on visual estimation as the primary method of assessing stenosis severity from CAG images.

In the CASS trial, conducted during the late 1970's and early 1980's, the criteria for defining a lesion as significant was a DS ≥ 70% in any main vessel and 50% if left main[61].

During the 1990's, in the BARI trial (1990's), a significant lesion was defined as DS ≥ 50% in any main vessel[62,91]. In the 2000's, the COURAGE trial (2000's) defined a proximal DS ≥ 70% in a major epicardial artery plus evidence of myocardial ischemia (or ≥ 80% plus symptoms) as significant[56]. In the SYNTAX[92], FAME[93] and STICH trials[94], a DS ≥ 50% was used.

In the 2010's, the ISCHEMIA trial used a DS ≥ 50% cutoff as well[95], as did four[51,96–98] of five the trials regarding complete vs culprit-only revascularization in the context of STEMI (the CVLPRIT[99] trial used a ≥ 70% cut-off). And in our current decade, the FAVOR-III[100] and FIRE[53] trials used the ≥ 50% cutoff as well, while the REVIVED trial[101] used a DS ≥ 70% for the major epicardial arteries and ≥ 50% for the left main.

It is worth noting that all of the above trials used the visual estimation of DS method as either the criteria for proceeding with revascularization or considering it (with the additional presence of symptoms or demonstration of ischemia by either invasive or non-invasive methods), regardless of the revascularization modality (PCI or CABG) or clinical context (acute vs chronic coronary syndromes), clearly illustrating the ubiquity of this approach.

As a result, current guidelines[34,58,102–104] also adopt the visual estimation of DS as the primary method for either considering revascularization or engage in further testing to determine the severity of a lesion, despite acknowledging its limitations[103]. A DS ≥ 70% (≥ 50% if left main) threshold is suggested for proceeding with revascularization, especially in cases where further testing is limited (such as valvular heart disease[104]). Guidelines also suggest that intermediate lesions, starting with a DS ≥ 50%, may warrant additional testing (such as invasive physiology), namely in clinical settings such as chronic coronary syndromes, with American guidelines limiting the upper cut-off of what constitutes an intermediate lesion at 70%[102,103], while European guidelines suggesting that a more stringent

upper cut-off of 90% may also be used[34,58], given that even in the 70-90% range, 20% of coronary stenosis may not be significant, as assessed by FFR[105].

If QCA would show good correlation with other better-established modalities of CAD assessment, or an improvement in outcomes, then stronger support for its use in detriment of visual DS estimation might ensue. As it stands, QCA measurements are still advocated for in some international research consortiums[106] or scientific societies[107], but only seldomly.

QCA measurements have shown only modest correlation with IVUS or OCT severity of lesions in older studies[108–111], although more recent evidence suggest that current software might perform better with higher correlation[112]. With regards to ischemia testing, 2D-QCA has shown modest to good correlation with FFR[113–115], albeit with wide variability across the same DS values (Figure 6)[116]; modest to good correlation ischemia imaging tests[117]; and very good correlation with 3D-QCA [117], which partly enabled the development of angiography-derived FFR[118], as outlined below. Furthermore, the idea that DS assessed by 2D-QCA might better correlate to FFR than DS assessed by visual estimation has also been challenged[89].

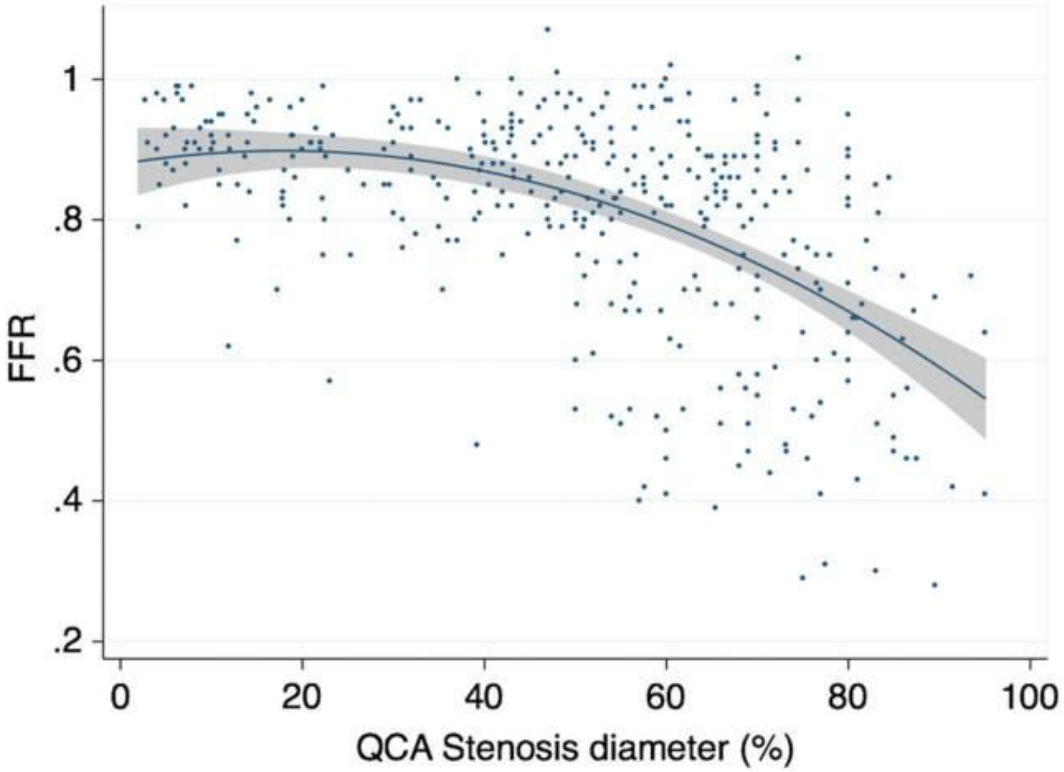


Figure 6: the relationship between FFR and 2D-QCA. Despite decreasing FFR values with increasing QCA Stenosis diameter, the wide variability of FFR measurements across the same DS limits FFR prediction from 2D-QCA. From Nijjer et al[116]. FFR: Fractional Flow reserve; QCA: Quantiativa Coronary Angiography.

Regarding clinical outcomes, while evidence is scarce, some studies suggest that the use of QCA might provide some prognostic information: a subanalysis of the PROMISE trial has shown that patients with obstructive disease as diagnosed by concordant QCA and visual estimation had higher event rates than those where QCA assessment did not find significant disease, even if visual estimation would classify it as such[119]. Furthermore, the recently presented (albeit not yet published) results of the GUIDE-DES trial suggest that optimizing PCI based on QCA may yield similar outcomes to IVUS-based optimization[120].

Thus, in light of all of the above, improvements in the assessment of the severity/significance of coronary lesions as viewed in CAG, rendering it more objective and reproducible, are evidently necessary. On one hand, the standard approach of performing DS visual estimations has obvious limitations which have persisted for decades, including the occulostenotic reflex, as coined by Eric Topol[121]. On the other hand, QCA is only semi (rather than fully) automatic, still hampered by somewhat limited evidence and only seldom used as a result.

2.2.2. Assessment of lesion morphology and complexity

The assessment of coronary lesions also includes characterizing their morphology. The most widely accepted classification is the AHA/ACC classification[122] (Table 3), which classifies lesions as type A, B or C. Each type represents an increasing level of procedural complexity and interventional failure rate. The following criteria are used for this classification: length, eccentricity, tortuosity, degree of calcification, ostial/bifurcation locations, complete total occlusion (CTO), presence of thrombus and native vs degenerated graft.

Table 3: coronary lesion morphology and classification according to AHA/ACC [122].

Type A lesions (minimally complex) – high success, > 85%; low risk	**Type B lesions (moderately complex) – moderate success, 60 to 85%; moderate risk**	**Type C lesions (severely complex) – low success, < 60%; high risk**
Discrete (length <10 mm)	Tubular (length 10 to 20 mm)	Diffuse (length >2 cm)
Concentric	Eccentric	Excessive tortuosity of proximal segment
Readily accessible	Moderate tortuosity of proximal segment	Extremely angulated segments >90°
Nonangulated segment (<45°)	Moderately angulated segment (>45°, <90°)	Total occlusions >3 mo old and/or bridging collaterals
Smooth contour	Irregular contour	Inability to protect major side branches
Little or no calcification	Moderate or heavy calcification	Degenerated vein grafts with friable lesions
Less than totally occlusive	Total occlusions <3 mo old	-
Not ostial in location	Ostial in location	-
No major side branch involvement	Bifurcation lesions requiring double guide wires	-
Absence of thrombus	Some thrombus present	-

2.2.3. Integrating severity and complexity as a whole

Combining the characteristics of all lesions is essential for fully appreciating CAD in a patient and ensuing with the necessary pharmacotherapy and revascularization strategies. Multiple scores may be used in the process.

The Leaman score considers the diameter stenosis of a lesion combined with the amount of myocardium supplied by the affected artery, taking into coronary dominance, by multiplying the relevance of a segment by a factor according to the severity of the obstruction[123] (Figure 7).

SEGMENT	RIGHT DOMINANT	LEFT DOMINANT	% LUMINAL DIAM. REDUCTION	WEIGHTING FACTOR
1	1.0	0.0	70-89	1
2			90-99	3
3			100	5
4				
5	5.0	6.0		
6	3.5	3.5		
7	2.5	2.5		
8	1.0	1.0		
9	1.0	1.0		
10	0.5	0.5		
11	1.5	2.5		
12	1.0	1.0		
13	0.5	1.5		
14	0.5	1.0		
15	0.0	1.0		

Figure 7: the Leaman Score. From Leaman *et al* [123].

A more complex score that also accounts for lesion complexity, diameter stenosis, location and disease extension is the SYNTAX (SYNergy between PCI with TAXUS™ and Cardiac Surgery) score, with higher values reflecting more complex and extensive disease. Its calculation is complex, and an online calculator is available (https://syntaxscore.org/calculator/syntaxscore/frameset.htm). This score was extensively validated in the SYNTAX trial, where three terciles of severity were observed: low (< 23), intermediate (23 – 32) and high (> 32). In general, in the landmark SYNTAX trial, higher scores favored CABG over PCI[92]. This is the most widely used score regarding decisions on revascularization modality and it has been employed in other landmark studies, such as the NOBLE[124] or EXCEL[125] trials.

Yet another score is worth mentioning: the British Cardiovascular Intervention Society - 1 Jeopardy Score (Figure 8). Like the Leaman Score, it considers the visually estimated diameter stenosis severity and lesion location to describe the severity of myocardium in

jeopardy and has been validated in the REVIVED trial[101]. Its' somewhat limited features contrast with the much-increased simplicity of use, as opposed to the SYNTAX score. However, it is far less widely used, as it has also been much less validated.

REVIVED **BCIS-1 Jeopardy Score**

	Coronary Artery/Graft	Instructions	Score
1	LMS	if ≥50% lesion, score 8 and go to row 11 if <50% lesion, score 0 and go to row 2	
2	Proximal LAD (Before DG)	if ≥70% lesion, score 6 and go to row 5 if <70% lesion, score 0 and go to row 3	
3	Mid LAD (After DG)	if ≥70% lesion, score 2 and go to row 4 if <70% lesion, score 0 and go to row 4	
4	Major DG	if ≥70% lesion, score 2 and go to row 5 if <70% lesion, score 0 and go to row 5	
5		If Cx dominant, go to row 8 If RCA dominant, go to row 6	
6	Proximal RCA (Before PDA)	if ≥70% lesion, score 4 and go to row 10 if <70% lesion, score 0, go to row 7	
7	PDA	if ≥70% lesion, score 2 and go to row 10 if <70% lesion, score 0, go to row 10	
8	Proximal Cx (Before OM)	if ≥70% lesion, score 6 and go to row 14 if <70% lesion, score 0, go to row 9	
9	Mid Cx (After OM)	if ≥70% lesion, score 2 and go to row 10 if <70% lesion, score 0, go to row 10	
10	Major OM	if ≥70% lesion, score 2 and go to row 14 if <70% lesion, score 0 and go to row 14	
11		If Cx dominant, score 4 and go to row 14 If RCA dominant, score 0 and go to row 12	
12	Proximal RCA (Before PDA)	if ≥70% lesion, score 4 and go to row 14 if <70% lesion, score 0 and go to row 13	
13	PDA	if ≥70% lesion, score 2 and go to row 14 if <70% lesion, score 0, go to row 14	
14		Previous CABG? If yes, go to row 15 If no, go to row 21	
15	LAD graft beyond DG	if <70% graft lesion, score -4, go to row 16 if >70%, poor run-off or n/a, score 0, go to row 16	
16	Major DG graft	if <70% graft lesion, score -2, go to row 17 if ≥70%, poor run-off or n/a, score 0, go to row 17	
17	Major OM graft	if <70% graft lesion, score -2, go to row 18 if ≥70%, poor run-off or n/a, score 0, go to row 18	
18	Cx graft beyond OM (Cx dominant system)	if <70% graft lesion, score -4, go to row 19 if ≥70%, poor run-off or n/a, score 0, go to row 19	
19	RCA graft (before PDA)	if <70% graft lesion, score -4, go to row 21 if ≥70% poor run-off or n/a, score 0, go to row 20	
20	PDA graft	if <70% graft lesion, score -2, go to row 21 if ≥70% poor run-off or n/a, score 0, go to row 21	
21	TOTAL SCORE	Add filled in scores and enter (range: 0 to 12)	

Figure 8: the BCIS-1 Jeopardy Score [101].

3. Invasive coronary physiology – a major companion to invasive coronary angiography

The limitations of assessing the significance of coronary artery lesions by CAG images alone led to the development of other methods capable of complementing the imaging findings. Ideally, such a method should provide an objective measurement index of some sort, while also being applicable in the setting of cardiac catheterization procedure.

The emergence of percutaneous coronary interventions led not only to therapeutic capabilities, but also the possibility of studying coronary flow and pressure from inside the coronary arteries *per se*. Grüntzig himself measured intra-coronary pressure across stenosis with his double lumen balloon, noticing the significant drop in pressure gradient after angioplasty[20]. In the 1980's, Wijns *et al* noticed the same, all the while correlating the pressure gradient with QCA-measured diameter stenosis, as well as ischemia assessed by thallium scintigraphy[126]. Assessing pressure with coronary dilatation balloons is, however, problematic, not only due to risk of iatrogeny, but also because the device itself may be partially occlusive, further dampening pressures and overestimating the pressure gradient[127]. Hence, the development of flow (via doppler or thermodilution) and/or pressure sensors integrated into coronary wires in the 1990's was a key step in enabling accurate physiological measurements[128–130].

3.1. Fractional Flow Reserve

Given the that measuring flow directly is cumbersome, a shift towards intra-coronary pressures as a surrogate of flow and measurement of the significance of epicardial disease (the subgroup of CAD that is effectively amenable to revascularization) soon became dominant. The most well-known and studied index, Fractional Flow Reserve (FFR), was introduced in 1993 in a seminal paper by Pijls *et al*, where the group conducted their experiments in five dogs[131]. The premise of FFR is based on the concept that, in the absence of resistance, flow is directly proportional to pressure, much like in Ohm's Law, voltage = current x resistance[132]. By abolishing micro-circulatory resistance with an agent (adenosine) that maximizes vasodilatation, this concept can be put into practice. FFR is defined as the ratio between pressure in the coronary arteries at the position of the transducer distal to a stenosis (Pd) relative to the pressure proximal to the stenosis (Pa), in practice measured in the aortic root, correlating strongly with flow as measure by a doppler wire[131]. The resulting value of FFR varies from 0

to 1, with a value of 0,80 and 0,75 indicating a 20% and 25% drop in pressure, respectively. Originally, FFR was described as myocardial FFR (FFRmyo), as the right atrial pressure (Pra) was taken into account ((Pd - Pra) / (Pa – Pra)). However, because the impact Pra is negligible and its measurement was an additional layer of complexity and invasiveness, a simplified formula is used (Pd/Pa)[133].

To define thresholds associated with clinically significant ischemia, FFR was first compared with non-invasive ischemia testing. Initially, treadmill stress tests were used. A pivotal study of 60 patients with single-vessel disease and preserved ejection fraction was undertaken. A cut-off ≥ 0,1 mV ST depression was used for the treadmill test. The authors found that an FFR cut-off < 0,75 had high accuracy for predicting a positive stress test[134]. On a later study, 24 patients underwent evaluation with FFR, treadmill, thallium scintigraphy and stress echocardiogram. 21 patients had a negative result on all three non-invasive tests. The three with reversible ischemia in at least one test underwent revascularization, after which all ischemia tests reverted to negative. The FFR cut-off of 0,75 was once again identified as the one with the best accuracy overall (93%). Over the next two decades, many other groups conducted relatively small studies (ranging from 20 to 232 patients) comparing FFR to non-invasive testing, noticing the existence of a so-called "gray-zone" of 0,75 – 0,80 as the ischemic threshold[132]. Considering than ischemia is a continuous rather than a binary concept, the finding is hardly surprising.

No less important than establishing an ischemic cut-off of clinical significance is determining whether the clinical use of FFR can have a meaningful impact on outcomes. Three landmark trials, from the group whom originally developed FFR, addressed this issue.

The DEFER trial tested whether PCI was beneficial in patients with an FFR > 0,75. There were no differences in either outcomes or symptoms in patients above the 0,75 cut-off, thereby suggesting that PCI above the FFR ischemic threshold is unnecessary[135]. Further large registries, including a large Portuguese registry, have shown similar findings[136].

The FAME trial compared angiography-guided (based on DS visual estimation) vs FFR-guided (based on a FFR value of ≤ 0,80) PCI. The latter group experienced improved outcomes, with a reduced rate of a composite outcome of death, myocardial infarction and repeat revascularization, driven primarily by the latter two endpoints, both of which narrowly missed statistical significance on their own[93].

The FAME-2 trial assessed whether FFR-based PCI is superior to medical therapy alone. The revascularization group experienced a sharp reduction (hazard ratio 0,32) in the primary endpoint of death, myocardial infarction or urgent revascularization[57]. The trial was strongly

criticized for its early termination given the large superiority of the FFR arm, which might have amplified its findings, as well as the fact that urgent revascularization was the main driver of the primary endpoint difference. Notwithstanding that, long-term follow-up at five years has shown that spontaneous myocardial infarction rates were reduced by 34% in the revascularization group[137], a finding strikingly similar to that of the ISCHEMIA trial[95], where ischemia was assessed by non-invasive methods, with very low usage rates of FFR. The aggregated results of these FFR trials, entirely focused on chronic coronary syndromes, with mostly one or two-vessel disease and PCI as the sole revascularization modality, strongly reinforced the role of FFR in guiding revascularization, while also highlighting the limitations of angiography-based assessment of coronary artery disease, as has been extensively discussed in previous sections.

FFR has also been studied in the setting of more extensive multivessel coronary artery disease in the FUTURE trial, which encompassed patients with two or three-vessel disease. The trial included patients with acute coronary syndrome (46% - 47%), where FFR was used in non-culprit lesions and was mandatory for all lesions with a DS ≥ 50%, as assessed by visual estimation. Any revascularization modality (PCI or CABG) could be used. Like the FAME-2 study, the FUTURE trial was also terminated early, but for the opposite reason - safety concerns in the FFR arm. No significant differences were found between groups, with a concerning but not significant trend towards increased mortality in the FFR group. The trial, however, had several major pitfalls. The result of FFR was not always respected, as more than one-third of FFR-negative lesions were revascularized. There were similarly low CABG rates across groups despite different syntax scores. Revascularization was incomplete in most patients, thereby downplaying its potential advantage. And, ultimately, the trial does not quite mimic standard practice, since any lesions with DS ≥ 50% had to undergo FFR, when in everyday practice only intermediate lesions are tested. This likely led to the maneuvering of pressure wires (harder to negotiate across lesions than standard wires) in relatively severe stenosis, which might explain the unusually high complication rates of FFR observed in this trial[138].

FFR has also been studied in the setting of acute coronary syndromes (ACS). Two of the major trials of complete vs culprit only revascularization in STEMI patients mandated the use of FFR in non-culprit lesions for deciding whether to proceed with revascularization or not[97,98]. The recent FIRE trial, which also included NSTEMI patients and focused on an elderly population also mandated physiology[53]. All of these have shown better outcomes in the revascularization group, reinforcing the role of an FFR-based strategy for revascularization decisions in non-culprit lesions in the setting of ACS. However, the only trial to have directly

compared an angiography vs FFR-based revascularization strategy failed to demonstrate an advantage of the latter strategy in STEMI patients. The very large confidence intervals of the results limit an exact interpretation of the results. However, given that microcirculatory vasodilatation may be impaired in this context, one might argue that FFR assessment could have led to an underappreciation of the severity of non-culprit lesions, due to incomplete elimination of microcirculatory resistance. However, the average 2,6 / 2,7 days between index procedure and FFR measurements should have attenuated this effect. In contrast, the FRAME-AMI trial, with a much smaller population which included both NSTEMI and STEMI patients, found better outcomes in the FFR group, driven by strikingly lower rates of all-cause death and myocardial infarction[52]. Thus, whatever the interpretation, the answer as to what the best strategy (FFR vs angiography) for non-culprit lesions in ACS patients is clearly not established, and recent ACS guidelines do not currently recommend the routine use of FFR in this setting[58].

3.2. Instantaneous wave-free ratio

As mentioned above, the calculation of FFR requires the use of adenosine, which is associated with patient discomfort, some contra-indications (namely asthma, known hypersensitivity or significant sinus or AV node disease) and cost, not to mention added time to the procedure[132]. Thus, efforts to obtain a resting index which would not rely on adenosine ensued and several are currently available today. The most widely studied and validated is instantaneous wave-free ratio (iFR). The premise of this index rests on the concept proposed by Davies et al that six waves drive coronary blood flow and there is a time period where no waves are present. The latter corresponds to a period of minimal and constant coronary microcirculatory resistance, with a consequent linear relationship between pressure and flow, with a nearly constant rate change[139]. As a result, theoretically, the fundamental premise that flow = pressure in the absence of resistance should be applicable in this phase of the cardiac cycle, and an index consisting of the Pd/Pa ratio during this particular period may be employed (Figure 9).

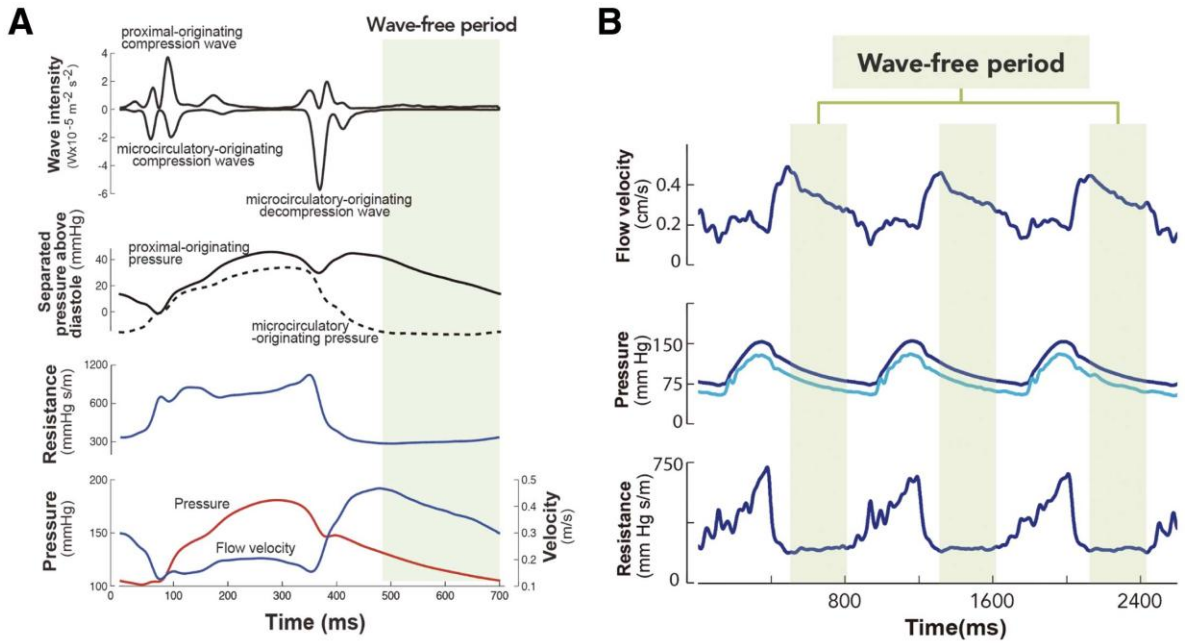


Figure 9: the six coronary waves during the cardiac cycle as described by Davies et al in 2006[139] (left). The demarcation of the wave-free period where iFR is measured vs the full cardiac cycle where FFR is measured during maximal hyperemia (right)[140]. iFR: instantaneous wave-free ratio; FFR: Fractional Flow Reserve.

In the seminal ADVISE study, the authors measured intra-coronary resistance in the wave-free period at rest and during maximal vasodilatation (using adenosine) and found similar measurements in 39 stenosis, thereby confirming their theoretical premise. Furthermore, the authors found a very strong correlation (R 0,9 ; $p < 0,001$) between iFR and FFR, testing in a pool of 157 lesions[141]. The fact that further research confirmed that with increasing stenosis severity, microcirculatory (distal) resistance is physiologically reduced, thereby generating a pressure gradient and maintaining flow, further reinforced the concept that Pd/Pa ratio during the wave-free period of diastole effectively measures the physiological impact of a stenosis[116].

The validation of iFR as compared to FFR ensued, as multiple studies (including from our group) continuously found good correlation with FFR, with an iFR cut-off of 0,89 closely matching the more widely accepted 0,80 FFR cut-off for significant ischemia[142–145], albeit with approximately 80% concordance. These findings were further confirmed against other invasive resting indexes combing pressure and velocity[146,147], as well Positon Emitron Tomography (PET), the gold-standard for non-invasive quantification of myocardial blood flow[148].

The wider introduction of iFR into clinical practice began after the results of the ADVISE-II study, an international multicentric validation study. The authors not only further validated the 0,89 iFR cut-off as the one more strongly correlated with an FFR of 0,80, but also confirmed that a grey zone exists between an iFR of 0,86 and 0,93. Outside this range, the

accuracy for an FFR ≤ 0,80 or > 0,80, respectively, was 94,2%[149]. In the absence of trials directly testing for outcomes, the authors proposed the hybrid strategy, whereby values of iFR ≤ 0,85 merited intervention, values ≥ 0,94 medical therapy alone, and values in the grey zone should undergo additional testing with FFR. This strategy permitted the avoidance of adenosine in 60 to 70% of patients[149].

iFR was then directly compared to FFR in two large non-inferiority trials, the DEFINE-FLAIR[150] and iFR SWEDEHEART[151], together encompassing 4529 patients. In these trials, an FFR value ≤ 0,80 and iFR value ≤ 0,89 would indicate revascularization (with either PCI or CABG as deemed appropriate), with values above those thresholds indicating medical therapy alone. Contrary to the FAME trials, both of these studies included patients with ACS (19%[150] - 38%[151]), with the interrogation of non-culprit lesions in STEMI patients occurring a minimum of 48 hours after the index procedure. No significant differences regarding clinical outcomes were noted between groups, either as isolated or composite endpoints at 1-year, for both trials. There were also no significant outcome differences between iFR or FFR according to clinical presentation (ACS or CCS)[152]. As expected, the avoidance of adenosine in the iFR arms led to a reduction of approximately 5 minutes in procedural time in the DEFINE-FLAIR trial (but not in the iFR-SWEDEHEART trial), with much less frequent chest discomfort in the iFR arms of both trials. Deferral of revascularization was more common with iFR vs FFR in both trials (45% vs 50%, p < 0,001 in a pooled meta-analysis), which led to concerns that it might translate in adverse events in the deferral arm, a finding not confirmed at 1-year[152]. Furthermore, the iFR-SWEDEHEART maintained its original outcome findings at 5-years[153], while a subgroup analysis of its deferral arm alone showed no increase in adverse events either[154], as did a very large SWEDEHEART registry data[155], suggesting that the increased deferral rates resulting from the use of iFR instead of FFR do not translate into harm. However, in the 5-year results of the DEFINE-FLAIR trial (not published but presented at EuroPCR 2023[156]), patients the iFR arm experienced an increase in all-cause mortality (IFR 9% vs FFR 6.2% p = 0.01), but not other cardiovascular events such as myocardial infarction, raising concerns with regards to the safety of an iFR based-strategy. Furthermore, two (not patient-level) meta-analysis of both trials found similar results[157,158], driven by the DEFINE-FLAIR data. Some authors[158] postulate that the fact that half of measurements were conducted in the LAD, where iFR and FFR are more commonly discordant (i.e. negative iFR but positive FFR)[159], may partially explain the finding. However, the wide confidence intervals of mortality data, the absence of patient-level data, and the absence of a non-irrelevant proportion of data regarding all-cause mortality at 5-years[157] render such conclusions difficult to interpret.

Lastly, the exact role of iFR for revascularization decisions of non-culprit lesions in the ACS setting is currently being investigated in the iMODERN trial, which is underway[160].

All of the above evidence pertains to diagnostic assessment of coronary lesions. However, both FFR and iFR have and are currently being studied for optimizing the results of PCI. Preliminary evidence suggests this approach might improve outcomes[161–163]. Because the focus of this thesis is the diagnostic interpretation of coronary artery disease, a detailed review of this approach falls outside its scope.

3.3. Common pitfalls of FFR and iFR

One of the debated issues of physiology is the assessment of tandem lesions. Because the presence of a distal stenosis leads to reduced hyperemic flow between that and a proximal one, the resulting individual FFR value is overestimated, leading to underestimation of the proximal stenosis severity. This so-called cross-talk phenomenon has long been described and multiple equations have been proposed to address this issue (Figure 10)[164,165].

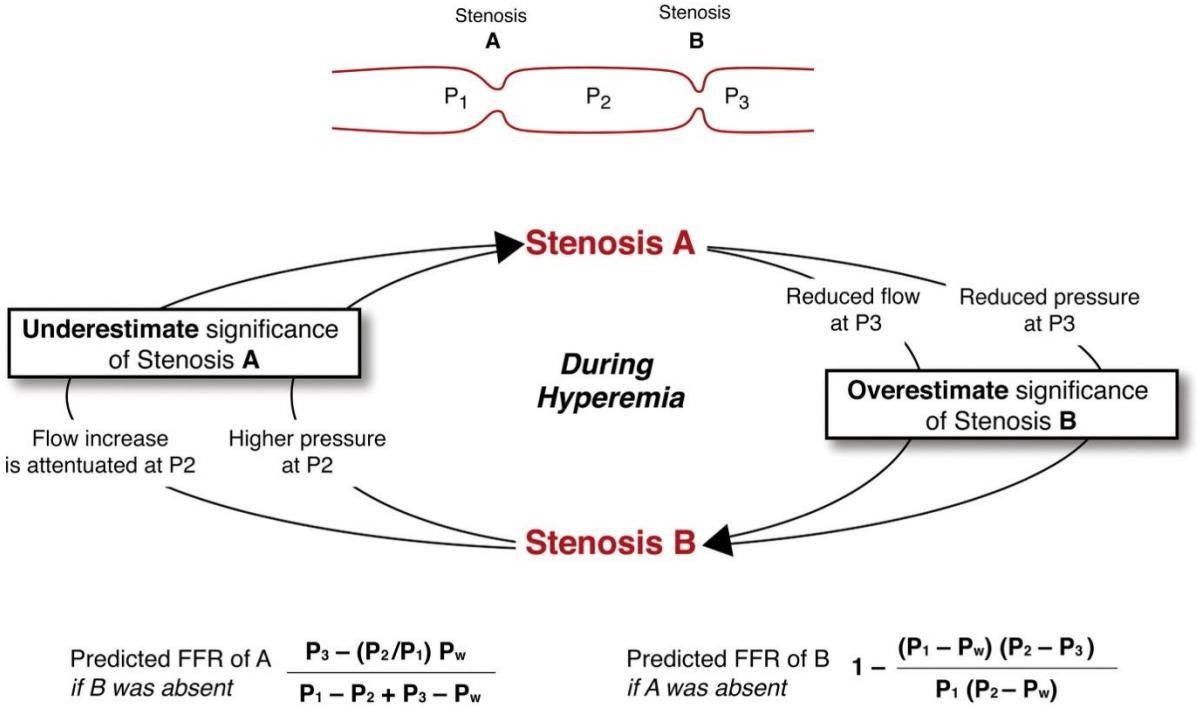


Figure 10: the cross-talk phenomenum. Tandem lesions can influence flow in hyperemic conditions. Stenosis B attenuates flow increase and increases pressure at P2. As a result, the resulting measured FFR in P2 is increased (underestimation of stenosis A). The proposed calculations enable recalculating FFR, but require balloon occulsion of the stenosis in question and are hence not usually used in clinical practice. From Nijjer *et al*[166].FFR: Fractional Flow Reserve.

In clinical practice a simpler method is applied: FFR is measured distally to all lesions, and pullback during hyperemia ensues. The largest pressure step-up permits identification of the most relevant segment, enabling the treatment of such regions first, while further measurements afterwards may ensue to assess the need for additional revascularization, given that treating a lesion may significantly influence the FFR of other lesions [167]. Using this approach, the concept of a pressure pullback gradient (PPG) was recently developed, enabling the identification of three patterns: focal, diffuse, or mixed[168–170]. This information can then be

used to better define the strategy of revascularization, although its exact impact on outcomes remains to be determined.

iFR, on the other hand, may theoretically be a better fit for such cases. Because, as mentioned, resting flow is relatively stable across stenosis, as demonstrated in the IDEAL study[116], the cross-talk phenomenon should be less relevant with resting indexes, as pressure changes should be more predictable across the vessel. Furthermore, for the same reason, PCI should not fundamentally change the resting flow of remaining stenosis and thus its resulting relative contribution to pressure drops (Figure 11)[166].

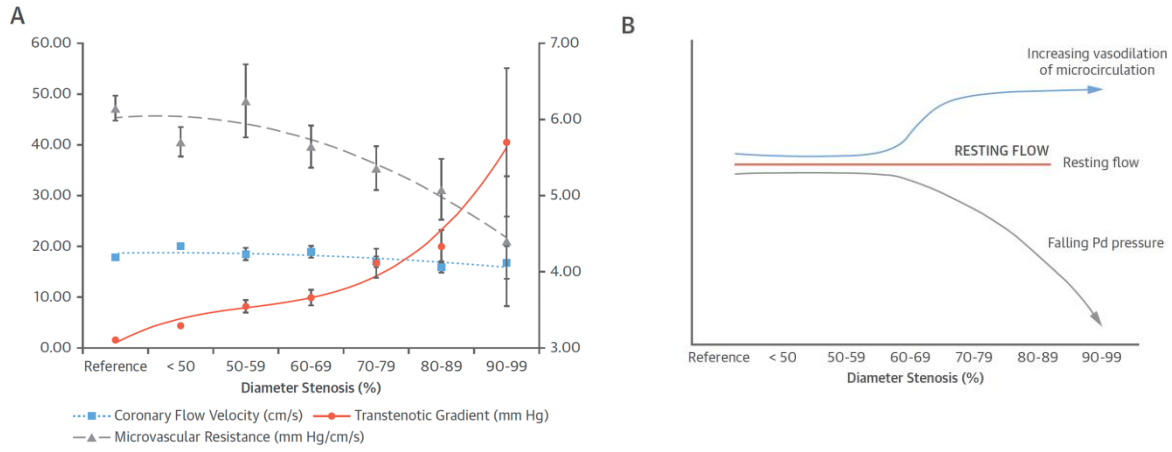


Figure 11: in resting (non-hyperemic conditions), flow across stenosis is maintained stable due to reduced microcirculatory resistance, with a resulting pressure drop [116]**.** Figure from Gotberg *et al*[132].

The fact that iFR is measured on a beat-by-beat (rather than averaged across 3-5 cycles) also enables a superior temporal resolution during pullback[132]. These theoretical observations have been supported by in-human studies, whereby an iFR pullback with fluoroscopy co-registration enables the assessment of the contribution of each stenosis to the overall distal iFR value. When operators assessed the predicted vs observed iFR after PCI, they found that the pullback predicted the physiological outcome with high accuracy, thus supporting the notion that iFR may better identify the individual contribution of each lesion in tandem disease (Figure 12)[171,172].

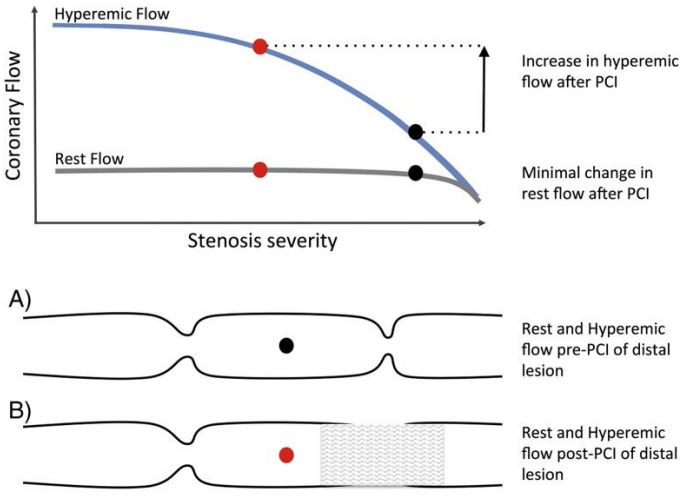


Figure 12: flow at rest remains relatively constant in tandem lesions. As a result, the exact prediction of iFR is not significantly affected by PCI. Thus, resting indexes may better assess individual contributions of stenosis severity in tandem lesions, leading to predictable iFR measurements pre and post-PCI [171,172]**.** From Nijjer *et al*[166]. iFR: instantaneous wave-Free Ratio; PCI: Percutaneous Coronary Intervention.

Despite this, some level of cross-talk does seem to exist even with resting indexes, since other studies have reported this phenomenon with iFR as well, albeit to a lesser degree[173]. Nonetheless, considering all of the above, together with the availability of easily usable fluoroscopy co-registration, iFR may be a better option in the setting of tandem disease.

Another pitfall of physiology is pressure drift, whereby small changes in intra-coronary pressure may lead to erroneous measurements. To account for this, operators should ideally perform a drift check by repositioning the coronary pressor sensor at the tip of the catheter, ensuring calibration is correct (i.e Pd/Pa = 1,0). Some reports have suggested that whole cycle indexes (i.e. FFR) are more prone to pressure drift, proposing that iFR may be advantageous given it is fully measured in only a part of diastole[174].

One highly debated issue of iFR and FFR is discordance. As mentioned above, when using a binary cut-off, the indexes disagree on approximately 20% of cases [142–145]. One common criticism of iFR that partially explains this is its limited range (i.e. less granularity), as well as distribution of measurements. Because iFR range is lower, with most measurements very close to the 0,89 cut-off with a small standard deviation and a tendency for mostly negative results (close to 70%)[150,151,174], the discriminatory ability of this index for classifying stenosis and assessing the degree of ischemia is lower than that of FFR. For the latter, the distribution of values is usually less close to 0,80, with a much wider range[174] and fewer lesions are classified as negative (close to 60%)[150,151]. Furthermore, lesions where a large amount of myocardium is

supplied[159,175], such as the proximal LAD or left main, might require hyperemia to truly unmask the significance of the coronary lesions, since the usual discordance pattern is an iFR > 0,89 with a FFR ≤ 0,80.

Additionally, both iFR and FFR have limited evidence in the setting specific clinical scenarios, two of which are of particular importance. In patients with severe aortic stenosis, systolic and hyperemic flow seem to significantly increase after treatment, contrary to flow during the wave-free period. Thus, the few small studies that have addressed the issue are more supportive for the use of iFR in this setting[176–179]. In contrast, for patients with heart failure (particularly with reduced ejection fraction), what little evidence is available is more supportive of the use of FFR, given that increased end-diastolic pressures may negatively affect iFR[133,180,181].

Lastly, the single greatest criticism of iFR is that, from a diagnostic perspective, it has always been validated as using FFR as reference. No trial has ever compared iFR based revascularization versus angiography in a "FAME-like" strategy. Hence, iFR can be considered a surrogate of FFR, rather than an index primarily tested in all clinical scenarios.

Despite all the pitfalls and gaps in evidence, it is clear that the bulk of currently available evidence supports the use of coronary physiology as a key adjunct to CAG images when making decisions regarding the significance of coronary lesions, particularly when considering revascularization. As a result, current European guidelines recommend the use of iFR or FFR in CCS (class I)[34]. In the setting of acute coronary syndromes, given the contrasting data currently available, physiology-guided revascularization of non-culprit lesions is only a class IIb indication[58]. Notwithstanding, physiology remains vastly underused, encompassing only 7-13% of procedures[182,183]. This is likely because it requires further invasive maneuvers, risk of iatrogeny, time consumption and cost. Thus, a non-invasive method of deriving physiology from CAG images alone using digital methods would be ideal.

3.4. Other indexes of physiology commonly used in clinical practice

Resting indexes other than iFR have also been developed. The Resting Full-Cycle Ration (RFR) measures the largest Pd/Pa difference encountered during the whole cardiac cycle. It has been validated in two large studies, with very similar measurement results when compared with iFR[184,185]. Other resting indexes, measuring Pd/Pa during selected phases of diastole have also been developed and yielded similar measurements to iFR[186,187]. In clinical practice, the choice between iFR and RFR are essentially a consequence of available technology in each

catheterization laboratory. However, it is worth noting that, unlike iFR, no large, randomized outcomes trial has been conducted for any of these indexes. Thus, one could argue that they constitute a surrogate of iFR, which in turn could be argued is itself a surrogate of FFR. Another index, contrast FFR (cFFR), has also been proposed as a surrogate of FFR with good accuracy, with the basic premise that contrast itself can induce significant hyperemia while avoiding adenosine[188]. However, given the larger evidence base for iFR, its use has received less attention.

Lastly, it is worth pointing out that when performing physiology of the coronary circulation, two major compartments can be assessed: macrocirculatory (epicardial arteries) and microcirculatory (pre-arteriolar vessels and arterioles). Because revascularization can only directly address epicardial disease, most of the focus of research has been on the above-mentioned epicardial indexes. Notwithstanding, a brief mention of non-epicardial indexes is warranted.

Coronary Flow Reserve assesses both epicardial and microvascular disease, consisting of the ratio between coronary flow during maximal hyperemia (with the administration of adenosine) versus at rest[189]. It can be measured by thermodilution or a doppler-wire. Although the latter may be more accurate, it is also more cumbersome[190]. Values greater than 2,5 (usually with doppler wire)[191] or 2 (usually with thermodilution)[192,193] have been deemed abnormal and associated with a worse prognosis.

Microcirculation is usually assessed by the Index of Microvascular Resistance, which is calculated as the distal coronary pressure divided by the inverse of the hyperemic mean transit time and thus represent a combination of pressure and flow (once again, by either thermodilution or doppler)[194]. A value ≥25 is considered pathological[195].

Further development of these is outside the scope of this thesis.

3.5. Non-invasive software-derived physiology from CAG images

In recent years, significant research effort has been undertaken to derive physiology from CAG images alone. Given that FFR is the most widely studied and evidence-based index, almost all groups primarily attempted to derive it non-invasively, rather than deriving iFR.

The Quantitative Flow Ratio (QFR) has been the most widely studied "virtual" index. Originally, for its calculation, a minimum set of prespecified orthogonal projections were acquired. Three-dimensional vessel reconstruction using 3D-QCA is then applied. Using computational flow dynamics (CFD), the contrast medium transport time was used for deriving

volumetric flow rate, which was then calculated during hyperaemia and at rest, enabling the derivation of CFR. Using the flow rate and the pressure at the catheter tip, together with the measured pressures with coronary pressure wires used for invasively measuring FFR, the authors were able to derive FFR based solely on the 3D-QCA reconstruction (Figure 13)[196].

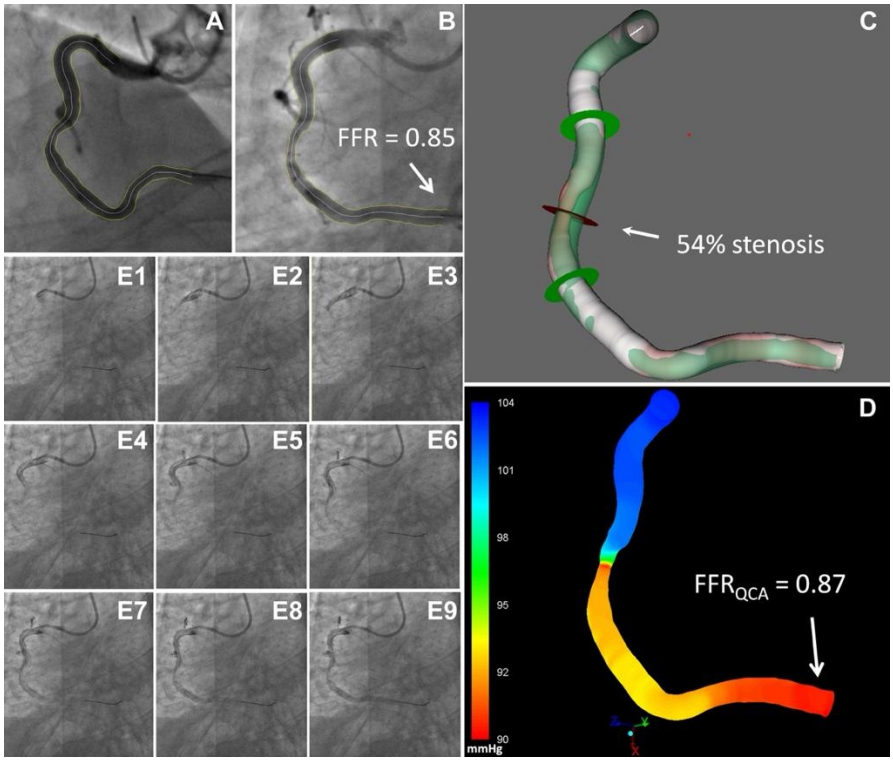


Figure 13: representation of the process of calculating QFR from 3D-QCA and CFD. From Tu et al [196]. QFR: Quantitative Flow Ratio; QCA: Quantitative Coronary Angiography; CFD: Computational Fluid Dynamics.

Three variants of QFR were then validated in a pilot study, with accuracies ranging from 80% to 87%[118], followed by two larger studies comprising more than 600 patients[197,198]. Once the good correlation of QFR with FFR was demonstrated, the next logical step would be a FAME-like trial where QFR-based revascularization would be compared with classical angiography-based (visual DS estimation) approach - the FAVOR-III trial. This was a Chinese multicentric trial with 3847 patients, where a QFR based revascularization was superior to an angiography-based strategy, resulting in a 41% reduction of myocardial infarction. The trial included mostly one or two-vessel disease, and the QFR threshold for revascularization was ≤ 0,80. Importantly, contrary to the FAME trial, no specific threshold for the angiography was mandated, as the protocol only specified "according to local standard"[100]. Nevertheless, the FAVOR III trial has been the first and only trial to date showing that non-invasive angio-derived physiology can improve outcomes without the limitations of invasive physiology.

Importantly, in the recent FIRE trial, where physiology-based revascularization of non-culprit lesions in elderly patients with ACS was compared with a conservative approach, QFR-based decisions were allowed, and comprised 35,2% of measurements[53]. It is plausible to consider that given the elderly condition of patients, the non-invasive nature of this method may have positively influenced outcomes, contrasting with the findings of the FLOWER-MI trial[199]. Currently, a large Euro-Japanese trial is underway directly comparing QFR to FFR as a method of revascularization decisions[200]. If positive for QFR, it could fundamentally change the paradigm of how to better assess coronary epicardial physiology.

Several other "virtual" FFR systems exist, of which two have been the focus of considerable research as well and merit mention. Both rely in methods somewhat similar to QFR, whereby a 3D-reconstruction based on 3D-QCA and CFD enables the non-invasive derivation of FFR.

FFRangio has been developed by the original FFR research group. In a pivotal trial and a pooled analysis of 5 additional cohort studies (> 700 patients), the authors reported an accuracy of approximately 92%[201,202].

Vessel FFR (vFFR) was validated in two studies comprising 434 patients and its accuracy was 90%[203,204]. Two large, randomized trial comparing vFFR to invasive FFR are underway[205,206].

In addition to the convenience and reduction of iatrogeny risks, another issue where non-invasive derivation of FFR may be of use is in determining the significance of non-culprit lesions in the acute-phase (i.e. ACS), given that the reliability and potential advantage of FFR in this phase has been challenged, as mentioned-above. A recent study compared QFR measurements of non-culprit lesions of STEMI at index event and 30 days later. They found that QFR did not significantly change over time (0.84 ± 0.09 vs 0.83 ± 0.10, p = 0,310), whereas FFR measurements significantly decreased (0,88 ± 0,08 vs 0,85 ± 0,10, p = 0,001)[207].

Even though all the above methods enable non-invasive physiology assessment, none are fully automatic. Indeed, several manual elements must be provided, such as defining normal vs diseased regions, marking the proximal and distal end of the target vessel, plus manually correcting vessel edges. Thus, in reality, these are semi-automatic systems. Furthermore, results may be less reliable outside an experienced core lab: in the FAST-II trial, the accuracy

of the vFFR measurements was 83% vs 90% at the core lab[204]. Fully automatic systems have only been described using artificial intelligence (AI) and are reviewed in the following chapter.

Given the characteristics of FFR, the stronger focus on its derivation is understandable, However, considering the potential advantages of iFR in certain contexts, particularly tandem lesions, its derivation would also be of clinical interest. To date, only a single study has focused on iFR rather than FFR derivation. Its rationale has been published[208] and final results are expected soon.

4. Artificial Intelligence and Medicine: a revolution in the making

4.1. Foundational Concepts

The creation of computer systems capable of human-like reasoning has been a long-dreamed ambition. In 1936, Alan Turing famously published what is regarded as a foundational concept paper – "On Computable Numbers, with an Application to the Entscheidungsproblem" (the German word broadly meaning "decision problem") - where he explores what ultimately would constitute a universal computer, while of course conceding that such a machine was well beyond the technological capabilities of the time[209]. In 1950, Turing published another impactful article, "Computer Machinery and Intelligence"[210], where he argued that the concept of a universal computer leads necessarily to the idea that it must be possible to program intelligence that is essentially indistinguishable from human intelligence, an idea that today would map into the concept of Artificial General Intelligence.

Today, the concept of Artificial Intelligence (AI) is quite broad, fundamentally consisting of computer systems capable of seemingly human-like reasoning – whether that be autonomous learning, decision-making or creativity. AI is thus an umbrella term, from which multiple other key concepts emerge[211–214].

Machine learning (ML) refers to the ability of a system to learn a skill it was not explicitly programed to perform. In a classical software system, a program (i.e. algorithm) is explicitly described, by instructions that specify how to receive an input and produce an output. In machine learning systems, both the input and output are provided, and the system is expected to produce the algorithm itself [211–214]. Although many approaches to machine learning exist, one of the most successful ones is based on the concept of a Neural Network (NN), trained by back-propagation of the error. The concept of back-propagation (in reality the computation of derivative using the chain rule) was made popular by a seminal paper dating back to 1986[215]. Since then, it has been greatly developed, and today NN form the basis of ML systems and are the fundamental technology used in the area known as Deep Learning (DL) and its networks referred to as Deep Neural Networks (DNN). In simple terms, an activation function enables each neuron to produce an output based on weighted input from other neurons. A cost function measures how right or wrong a model is. The performance of the model is then improved by a gradient descent function, optimizing the weights between neurons. [211–214]. This training process is computationally resource heavy, usually requiring very powerful graphics

processing units (GPU). Using already trained models into common computer systems is usually much less computationally demanding.

Two further concepts are of particular relevance to the application of AI in Medicine and this particular thesis. One is Convolutional Neural Networks (CNN), a particular type of DL NN, which have been shown to be of particular use in medicine, especially for image analysis [211–214]. The other is Transformers, a specific type of DL architecture that relies on attention mechanism and is regarded as a major breakthrough in recent years, exceling in image interpretation and, particularly, in natural language processing (NLP), made popular by systems such as ChatGPT[216].

Although machine learning can take multiple forms, two main classes are supervised learning and unsupervised learning. The former enables teaching known patterns, taking labelled/annotated data as input. The latter takes unlabeled input and is expected to find unknown patterns. It is also worth mentioning a third type, reinforced learning, a process where system is trained to execute sequential decisions by receiving only occasional environment feedback [211–214].

The explosion of AI in recent years, including in Medicine, can be construed as the emergence of a perfect storm: computational power is far greater than in the past; and the dawn of the internet has made big data, cloud sharing and open source easily possible[214].

Lastly, AI is not without its pitfalls and dangers. Given the many current (and likely future) capabilities of AI, ranging from image analysis to NLP (both written and spoken), AI is increasingly becoming able to "see", "hear" and "speak" in creative ways. As such, the current and potential applications to Medicine are immense, ranging from image interpretation, decision-making, electronic health records analysis or structuring and even patient engagement[213,217]. Such a prospect is both remarkable and daunting, but some would argue that these technologies may free physicians from mundane and repetitive tasks, paradoxically making Medicine more human. Such is the perspective of Eric Topol, which he defended in his acclaimed essay "Deep Medicine"[214]. One study suggests that while this hopeful perspective is dominant among Cardiologists, multiple concerns regarding cost, usability, limited knowledge, poor electronic health records and lack of trust have been raised[218]. Indeed, in Medicine, several limitations are of concern. Big data availability is necessarily constrained due to privacy and medical privilege, rendering training datasets much smaller than in other settings, potentially hampering the performance of existing and future models. Moreover, data annotation can be extremely cumbersome and those better qualified (i.e. physicians) can sparingly take part in such tasks, further limiting both the size and quality of training datasets.

Furthermore, the exact mechanism by which many AI systems function is not always entirely understood, raising ethical concerns as to its application in clinical practice.

4.2. Application of AI in Cardiovascular Medicine

Research in AI applied to Cardiology has increased exponentially, from 29 papers in 2012 to 921 in 2022 (Figure 14). More than half comes from the USA (58%), followed by Europe (28%) and Asia (11%). Nearly 90% of research focuses on diagnosis (50%) and prediction (38%), with the remainder related to screening or decision-making. DL methods were used in approximately half of cases (48%), which given the larger focus on diagnosis, ECG (37%) and imaging (22%), is expected[217].

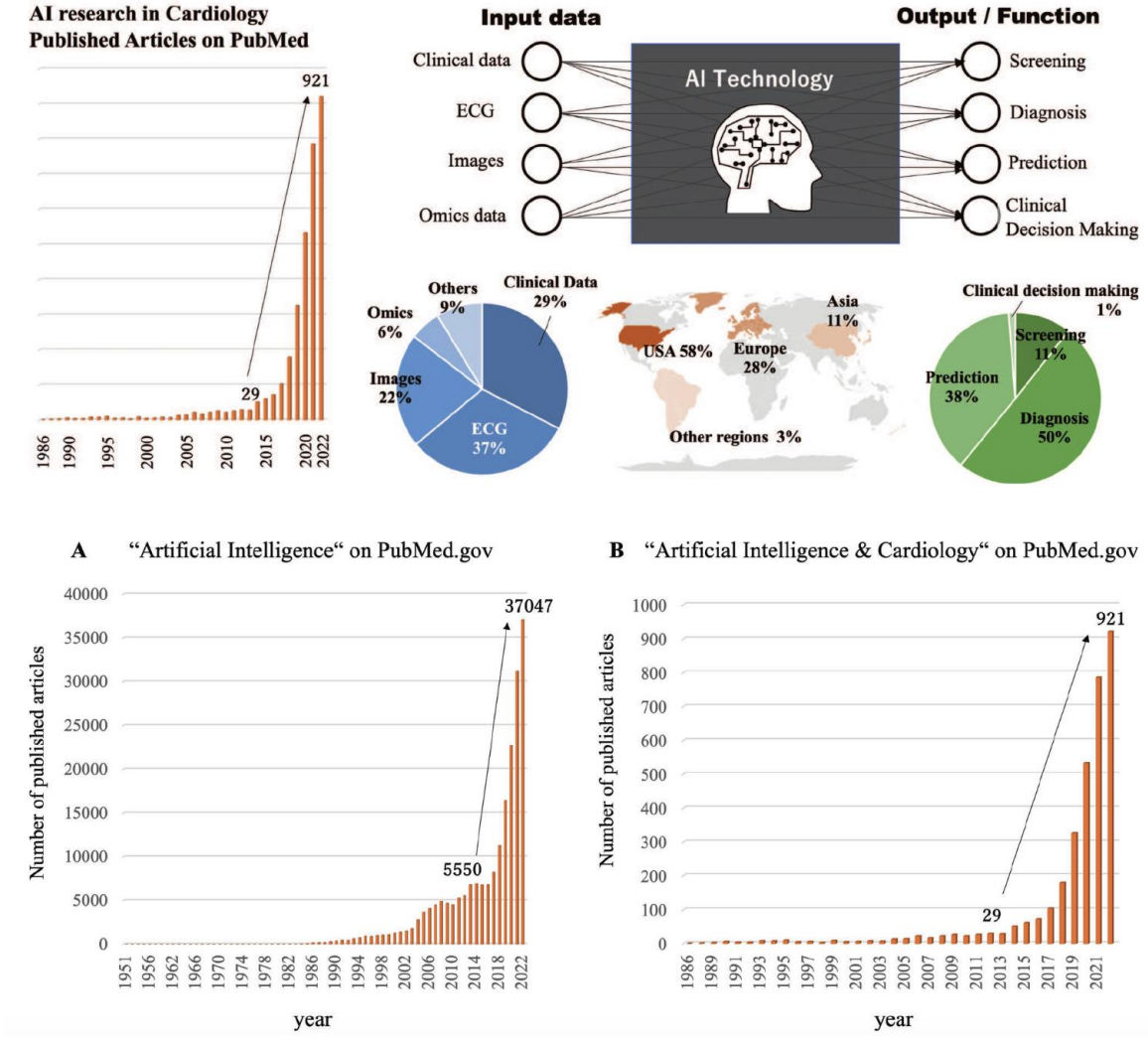


Figure 14: the exponential growth of AI research in Cardiolog in the last 10 years. From Makimoto et al [217]

Background

In the field of electrocardiography, multiple papers have been published regarding the automatic detection of arrythmias or hypertrophy (an example of identifying known patterns), as well as left ventricular disfunction (unknown patterns), among other conditions[219–222]. Our own research group has developed algorithms capable of highly specific acute pulmonary embolism diagnosis[223].

Echocardiography interpretation with AI has also been explored. Several groups have created automatic chamber quantification algorithms[224,225]. A Stanford group has developed a fully automated cardiac system especially focused on ejection fraction[226], and even conducted a blinded randomized trial comparing the performance of AI to experienced sonographers, with a clear demonstration of the non-inferiority of AI[227]. The authors have since announced the deployment of these systems in their echocardiography laboratory. Other groups have developed models capable of detecting valvular heart disease, such as mitral regurgitation[228] or aortic stenosis[229].

A great deal of effort has been placed in Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), given that the high reproducibility of images, together with excellent resolution, make them ideal candidates for image analysis. In Coronary CT (CCTA), models for automatic calcium score calculation[230] and stenosis quantification[231] have already been developed. Furthermore, FFR derivation from CT images (FFRCT) has been achieved for a few years now[232], but improvements using ML algorithms have since been developed[233,234]. In cardiac MRI (CMR), fully automated segmentation and quantification of cardiac chambers has been described and is available on commercial software already, greatly speeding the process of CMR analysis[235]. Research efforts have also been conducted regarding the analysis of patterns specific of myocarditis or hypertrophic cardiomyopathy, among others, by automatically detecting and classifying patterns of late gadolinium enhancement, as well as T1/T2 mapping[236,237].

The application of AI to clinical activity has also been explored. In heart failure, the use of wearable technology has enabled the assessment of compensated vs imminent and pre-clinical decompensation[238,239]. Furthermore, efforts are underway to provide voice technology into clinical practice, for both retrieving and providing information[213,240].

All the above developments, however, pertain largely to non-invasive cardiology. Much fewer research efforts have been placed in Interventional Cardiology, with the exception of OCT images interpretation, where currently available commercial software already provides a great deal of automatic analysis regarding sizing, stent apposition and lesion characterization, with meaningful impact on clinical outcomes in two recent major trials[39,241]. Thus, despite the

promising applications of AI in decision making, image interpretation and even augmented reality, there is still much room for research and improvement[212,242]. Notwithstanding, the basic diagnostic assessment every coronary procedure starts with is CAG. The following section briefly reviews the application of AI to CAG interpretation.

4.3. AI applied to CAG

The application of AI to CAG interpretation can be construed in a number of ways. Arguably, the first step would be to correctly separate the coronary tree from noise (i.e. background). This might be achieved by means of exact demarcation of the arteries (i.e. semantic segmentation) or a broader assessment, with a perhaps greater focus in lesion detection (i.e. object detection). The second step would be to interpret the findings, such as location, thrombus, calcium or stenosis. Given that, as explained in previous chapters, one of the greatest conundrums of CAG interpretation is the heterogeneous interpretation of stenosis severity assessment, as well as the fact that physiology-based severity estimation seems to have superior impact in clinical outcomes, an AI tool should also ideally improve the ability of operators to classify the severity of lesions, while additionally providing insight into its physiological significance.

Most published papers are essentially technical and thus dwell in engineering repositories, focusing primarily on segmentation. Most groups have used CAG images annotated by either physicians, medical students or technicians, with datasets ranging from hundreds to a few thousand images. The annotations may be either manual or based on existing clinical software. Using known image analysis architectures, such as ImageNet[243], U-Net[244], or other variations [245], segmentation models are developed for this specific task. To assess the performance of these, the authors usually compare the superimposition of annotated vs AI-segmented images, often using the Dice Score as gold standard. This score is calculated as (2 x true positives) / ((2 x true positives) + false positives + false negatives)), thus emphasizing the identification of true positives while penalizing false negatives. Some authors have reported Dice scores of approximately of 90% in CAG segmentation [244,246]. A more detailed review of technical papers, however, falls outside the scope of this thesis, given its medical nature.

Few publications in medical or biology journals regarding AI-based CAG interpretation and segmentation are available. Tianming Du *et al*[247] developed a segmentation model with a very large dataset of 13373 images annotated by 10 qualified analysts (the exact role or experience in clinical practice was not specified). While the Dice Score was not used, the authors report a segmentation accuracy of 98,4%, when considering pixel overlap. The authors report the models were also able to correctly identify thrombus, calcium and lesion location, but lesion severity is not specified by this algorithm (Figure 15).

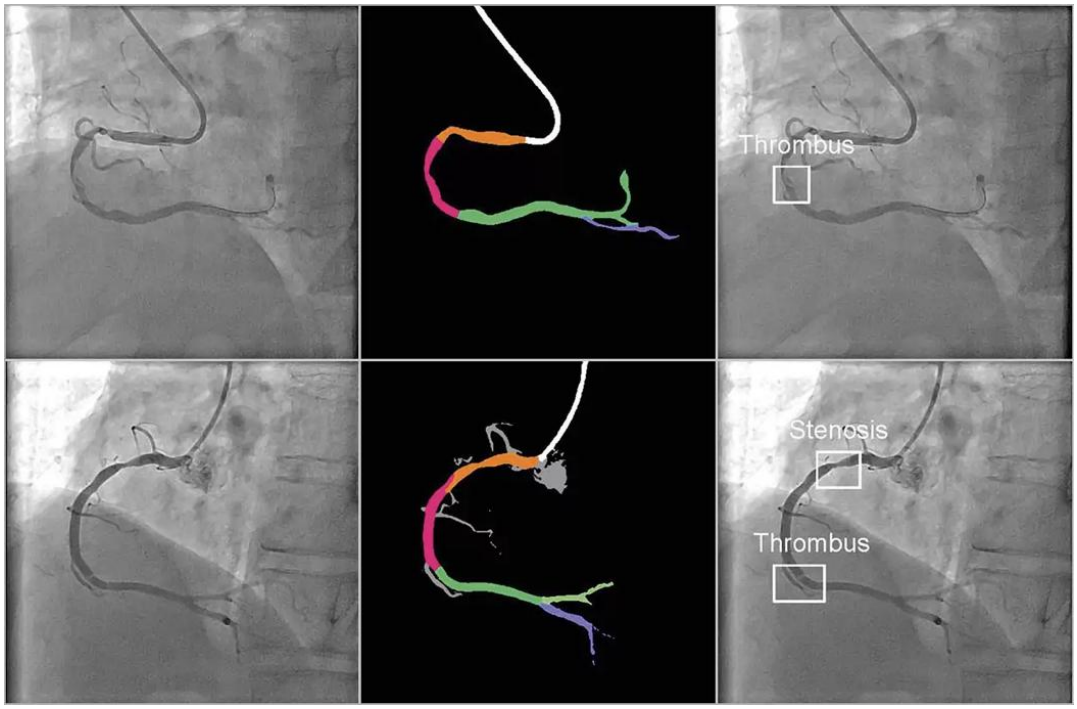


Figure 15: an AI segmentation algorithm by Du *et al*[247].

Another group has also developed segmentation models, either focused on stenotic regions alone, or the whole coronary tree, with similar performance, but using smaller (3302 to 4904 images) datasets[248–250]. Lastly, some models have combined non-AI based computer vision methods with DNN models to perform segmentation on smaller datasets, with good accuracy as well (Dice Score of 87,4%)[251].

Despite the seemingly impressive performance of these models, no evaluation regarding the quality of the segmentation, from a clinical perspective, was employed. Thus, if very small imperfections are present, but affect a critical part of the coronary tree (such as incorrectly segmenting a proximal LAD), very high accuracy and Dice scores can still be achieved, but a major flaw, from a clinical standpoint, would be present. Furthermore, no exact method for conducting such an assessment as either been developed or validated.

With regards to stenosis severity assessment with DS, one group has developed a fully automated auto-QCA system with reasonable correlation to existing systems (Pearson's R = 0,765)[252].Two other groups developed models capable of identifying lesions based on bounding boxes, either of any severity or with a DS ≥ 70% as assessed by QCA[253,254].

Lastly, very few data regarding the derivation of physiology from CAG images using solely AI methods has been published.

To our knowledge, only one fully automatic software - AutocathFFR - has been made public but is not yet commercially available. It relies solely on artificial intelligence and is therefore quite different, from a methodological standpoint, from the above-mentioned systems. Manual input is not only unnecessary but also not possible – the system automatically detects lesions. The authors reported an accuracy of 90% in a small pilot study with only 31 patients[255] and a 94% accuracy in a very recent three-center retrospective validation study of 304 vessels from 297 patients[256]. Regarding the training process, the authors consecutively used more thatn 13000 invasive coronary angiogram procedures and 1500 FFR measurments, from centers across Israel, USA, Japan and India. No lesion types were excluded. QCA software was used for labelling the lesions and all FFR measurements, except if off-label, were used. Additionally, the authors state that "*we used mild and extreme lesion data for cases in which invasive FFR is not usually in use*", which presumably means classifying these lesions as FFR negative or positive. However, the exact details regarding the exact AI training architecture were not provided, nor was any detailed data regarding the 13000 angiograms and 1500 FFR measurements used for training made available (Figure 16) [256].

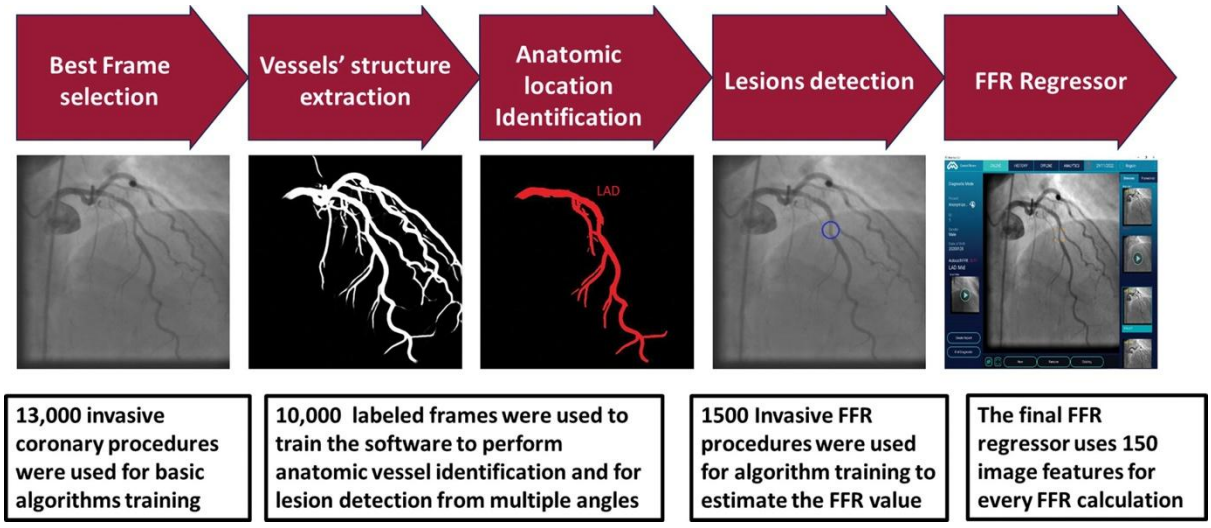


Figure 16: workflow of the AutocathFFR system and training process outline. From Ben-Assa et al [256].

Another group has also worked on fully automatic FFR derivation from angiogram. Using AI methods, the authors created a binary classifier (using the 0,80 FFR cut-off) with

82% accuracy[257]. Other researchers have launched an initiative for fully automated AI-based PCI guidance based on FFR[258], but no results have been published so far.

Lastly, no results data regarding the derivation of iFR from CAG images using AI has been published, to our knowledge. Thus, despite existing and potential applications of AI in Cardiovascular medicine, it's clear that a great deal of further research is necessary for developing and deploying AI tools for aiding CAG interpretation in the catheterization laboratory.

5. Major Gaps and Pitfalls in the interpretation of Coronary Angiography and the application of AI

The limitations in the interpretation of CAG, as those of currently available AI technologies in this context, have been explored in detail in previous subchapters. This section very briefly provides their systematization:

1. Issues with Coronary Artery Segmentation
    a. Segmentation is a fundamental step in the development of AI systems for CAG interpretation
    b. There is a paucity of segmentation methods capable of simple, fast and fully automatic CAG segmentation
    c. No methods are available to assess the quality of the segmentation from a clinical perspective
2. Issues with the interpretation of lesion severity
    a. Visual estimation of percentage diameter stenosis remains the cornerstone of lesion severity assessment and, thus, consideration for revascularization
    b. Visual assessment of lesion severity is prone to operator variability and overestimation of stenosis
    c. QCA enables the objective and reproducible assessment of lesion severity, but is understudied and seldom used
3. Issues with the use of coronary physiology in clinical practice
    a. Coronary physiology implementation in everyday practice remains beneath desired levels
    b. Despite the benefits of invasive coronary physiology, its impact in improving outcomes has not been uniformly demonstrated due to neutral clinical trials in more complex scenarios, namely acute coronary syndrome and extensive multivessel disease. The increased risk of complications in such clinical contexts, together with the limitations of invasive physiology imposed by the acute setting, likely contributed to these findings
    c. Non-invasive derivation of coronary physiology has been developed and may help overcome the above-mentioned limitations, but evidence is still insufficient.

Furthermore, currently available software is only semi-automatic and its reliability may be hampered by operator heterogeneity and inexperience

d. Non-invasive derivation of coronary physiology has focused almost entirely on FFR. Given that no single index is ideal and iFR may pose some advantages, the derivation of iFR from CAG images is desirable
e. Non-invasive derivation of coronary physiology primarily with AI technology has seldom been explored and may eliminate or significantly reduce the limitations of currently available non-AI systems

CHAPTER 2: OBJECTIVES

This dissertation aims to fulfil part of the pitfalls and gaps in evidence pertaining to the interpretation of coronary angiography, exploring the potential role of AI as the primary tool for the task.

In accordance, four research projects were developed, with the following objectives:

Part 1: to create fully automatic AI models capable of accurate CAG segmentation and develop a scoring system capable of assessing their quality from a clinical perspective, in addition to those of conventional metrics.

Part 2: to validate both the AI models and said score in a multicentric external validation dataset.

Part 3: to assess the impact of AI-based segmentation models in the assessment of coronary lesions severity, namely its impact in inter-operator heterogeneity and overestimation tendency.

Part 4: to develop AI-based validation models capable of deriving iFR data from CAG images alone in fully non-invasive fashion.

To achieve the proposed objectives, four observational studies, two of which multicentric, were undertaken, as outlined in the following chapter.

CHAPTER 3: RESEARCH PROJECTS

Part I. Development of deep learning segmentation models for coronary X-ray angiography: quality assessment by a new global segmentation score and comparison with human performance

Miguel Nobre Menezes[1,2*], João Lourenço Silva[3], Beatriz Silva[1,2],Tiago Rodrigues[1,2], Ana Rita Francisco[1,2], Pedro Carrilho Ferreira[1,2], Arlindo L. Oliveira[3], Fausto J. Pinto[1,2]

(1) Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon, Faculdade de Medicina, Universidade de Lisboa, Av Prof. Egas Moniz, 1649-028 Lisboa, Portugal

(2) Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Av Prof. Egas Moniz, 1649-028 Lisboa.

(3) INESC-ID / Instituto Superior Técnico, University of Lisbon

Rev Port Cardiol. 2022 Dec;41(12):1011-1021. doi: 10.1016/j.repc.2022.04.001.

1. Introduction

Artificial intelligence (AI) has shown great potential in Medicine, in applications such as predictive data analysis[259], decision making support[260] or even medical education/awareness improvement[261], but especially in image analysis. Several publications have demonstrated impressive results with regards to ECG[262], echocardiography[224,225] or MRI[235,263].

Notwithstanding, the use of AI in Interventional Cardiology (IC) is still vastly underexplored. In particular, its application to coronary angiography (CAG) has been explored in very few medical/biology publications[247–250]. The possibilities are nonetheless many[264], ranging from automatic anatomical identification, stenosis analysis, lesion subset characterization and perhaps even physiological indexes derivation. Regardless of whichever task one aims to fulfil, arguably the first step in applying AI to CAG is separating and identifying relevant information –the coronary tree – from non-relevant information (bones, other structures). This task is called segmentation[265].

In this paper, we explore the development of AI models capable of automatic coronary artery segmentation from CAG, and assess the results from a clinical perspective, using a new set of criteria and score clinically defined by a panel of Interventional Cardiologists.

2. Methods

2.1. Dataset selection

2.1.1. Inclusion criteria

We retrospectively and randomly included patients who had undergone CAG and invasive physiology assessment (Fractional Flow Reserve and/or other indexes) during the procedure in a single center (tertiary university hospital).

These patients have at least intermediate lesions in one or more vessels. Around one third usually undergo revascularization due to the severity of their disease[150,151]. Therefore, a dataset focusing on these patients comprises a wide spectrum of obstructive coronary artery disease in a relatively balanced way.

2.1.2. Exclusion criteria

We excluded cases where any of the following applied:

1) Major occluded vessels (acute or chronic)
2) Poor image quality
3) Less than two orthogonal views in the left coronary artery (LCA) - one caudal and one cranial - or absence of at least one left oblique (LAO) view - either cranial or simple - in the right coronary artery (RCA)
4) Patients with previous cardiac surgery, cardiac devices or other sources of potential artifact

2.1.3. Image selection

For each diagnostic angulation incidence from each patient, a single best frame was selected.

2.1.4. Dataset size

The dataset size was the result of a trade-off between two opposing criteria: dimension large enough for successful training of a deep convolutional neural network, estimated from published data[244,248,249,266] vs expected time required to complete the annotation. We estimated the latter based on a short period of annotation testing prior to formal dataset creation. The trade-off pointed out to a training set size of roughly 400.

We then randomly and consecutively selected patients until a total of at least 400 annotated images were obtained.

2.2. Baseline annotation process

Baseline human dataset images were annotated by two senior Cardiology Fellows (TR/BS) previously trained in CAG interpretation, under the supervision of an Interventional Cardiologist (MNM), who also annotated. Images were periodically reviewed and perfected by all three, such that any initial heterogeneity between annotators was corrected by consensus. The small size of the team was aimed at reducing heterogeneity, as we noticed, during a preparatory phase, that some operators tended to annotate too much (supplementary figure 1) while others did the opposite (supplementary figure 2).

Both the catheter (labelled red) and the coronaries (labelled white) were to be segmented.

The coronary tree was to be fully segmented up to branches of approximately 2 mm in calliper at their origin (as the vessel became smaller, it was to be segmented until discernible), using the catheter as reference (without formal measurements - eyeball appreciation was used). There were several reasons for this: (1) when performing percutaneous coronary intervention, vessels smaller than 2 mm are usually approached conservatively, as the risk of target lesion failure increases significantly[267,268]; (2) human annotation is cumbersome - segmenting every single vessel would increase the risk of errors significantly; (3) including very small vessels might increase the chances of artifacts from bone or other structures when training and applying AI models.

2.3. Baseline AI model training

We performed segmentation using an encoder-decoder fully convolutional neural network based on the U-Net[269], commonly used in medical image segmentation. As their name suggests, these neural networks are composed of an encoder, responsible for extracting image features, and a decoder, which processes those features to produce segmentation masks. To derive the best approach for this task, we conducted a comparative study of encoder and decoder architectures, which resulted in the proposal of the EfficientUNet++, a computationally efficient and high-performing decoder architecture[270] that, in this work, we combine with an EfficientNet-B5 encoder[271] (Figure 17).

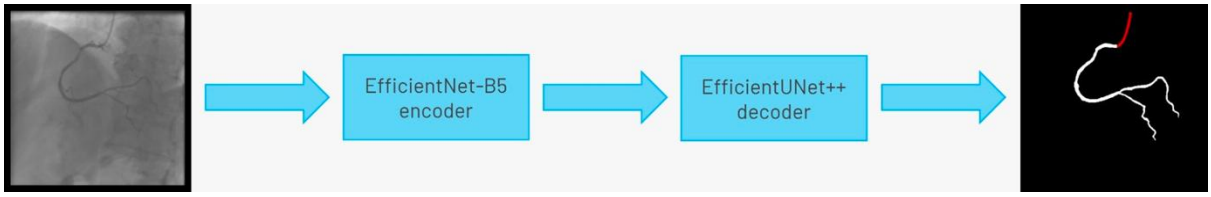


Figure 17: Segmentation model composed of an EfficientNet-B5 encoder and an EfficientUNet++ decoder.

To ensure fair evaluation, it was necessary to guarantee that each model was tested on data it had not seen during training. Therefore, the dataset was split, at the patient level, into 13 subsets of approximately 32 angiograms each. Each subset's segmentation was performed using a neural network trained exclusively on the remaining data. This enabled the assessment of the segmentation results for the entire cohort, as the usual splitting into a training and testing dataset would have yielded a much smaller group of images for result assessment.

The training hyperparameters, namely the number of training epochs and the learning rate decay schedule, were set on the first train-test split, using 1 of the 12 training data subsets for validation. The selected values were then used on every other train-test split, and to train the model on the whole training set of the first split.

2.4. Enhanced human model

The results of the baseline AI training were reviewed by the annotating team, without formal grading as later on (see below). For each image, both human and AI segmentation were compared with the original. Every annotation was then perfected using a mixture of the best of baseline human segmentation and baseline AI, with additional *de novo* manual segmentation as needed.

2.5. Enhanced AI model

The neural network architecture and training procedure were identical for both the baseline and enhanced AI model (Figure 17). The sole difference was the dataset. The baseline AI model was trained using the baseline human annotations, whereas the enhanced AI model was training using the enhanced human annotations. Figure 18 outlines development stages.

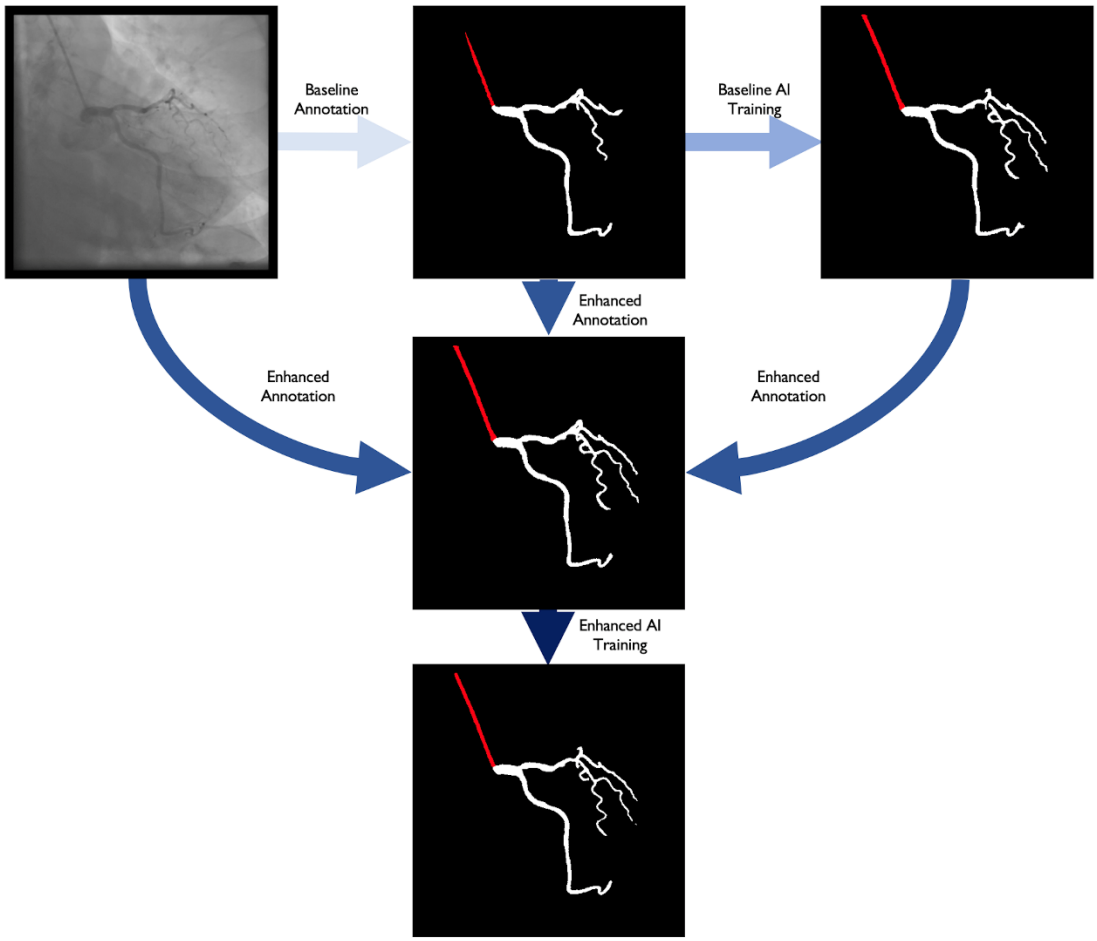


Figure 18: annotation and training process.

2.6. Performance assessment

2.6.1. Non-medical metrics

AI models were evaluated using the Dice Similarity Coefficient (DSC) and Generalized Dice Score (GDS), measures of the overlap between segmentations. Given two segmentations, the DSC has a value between 0 - no overlap - and 1 - total overlap - corresponding to the ratio between the area of their intersection and the sum of their areas. The GDS[272] is a weighted sum of each class's DSC that attributes the same importance to all classes, regardless of their frequency. While the DSC and GDS alone do not reflect clinical usefulness, they are helpful and entirely objective metrics that allow simple comparison between models.

2.6.2. Clinical performance criteria

The DSC objectively assesses model performance. However, it does not provide a medically meaningful impression of whether segmentation is appropriate. Also, because the DCS can only be calculated based on previously annotated images, it cannot be applied to new, unannotated datasets in the future. To overcome these limitations, we created a set of criteria to assess performance as interpreted by expert physicians.

The following 11 criteria are as objectively defined as possible and were analysed for each image. Each was independently met or not. A “perfect” example is shown in Figure 19. Supplementary figures 3 to 13 show error examples for each.

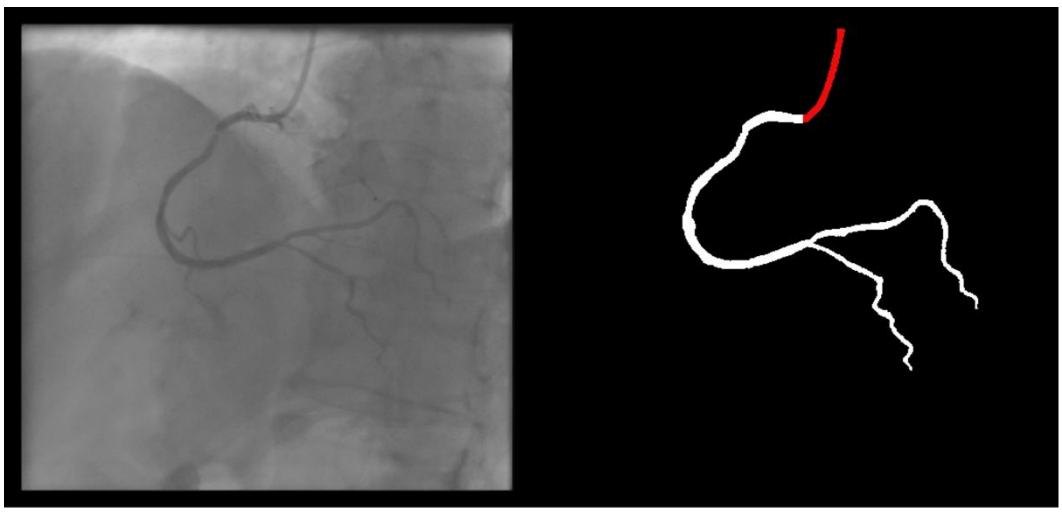


Figure 19: a segmentation case fulfilling all 11 criteria.

Research Projects

1) **Catheter segmentation**:
    a. **Main segmentation**: The distal part of the catheter (i.e. the closest discernible portion to the coronary artery in the ascending aorta) is correctly segmented and labelled (supplementary figure 3). If minor gaps are present, this criterion should be scored as met.
    b. **Gaps** (minor) are absent (supplementary figure 4)
    c. **Catheter thickness** is accurate, by visual appreciation (supplementary figure 5)
    d. **Location**: if parts of the catheter far from the coronary ostia (ascending and/or descending aorta) are segmented, there are no major gaps or artifacts (supplementary 6)
2) **Vessel segmentation**:
    a. **Main vessels** are correctly segmented and labelled. For the RCA, this includes the segments from the ostium to the crux (supplementary figure 7). For the LCA, this includes the segments from the left main ostium to the visually discernible distal segments of the left anterior descending or the circumflex (or most important obtuse marginal branch), depending on incidence. Branches are excluded from this criterion. If minor gaps are present, this criterion should be scored as met.
    b. **Branch segmentation**: branches with a luminal diameter of at least approximately 2 mm (using the catheter size as reference) are correctly segmented and labelled (supplementary figure 8). Size is estimated by visual appreciation. If minor gaps are present, this criterion should be scored as met.
    c. **Main vessel gaps** (minor) are absent (supplementary figure 9).
    d. **Branch gaps** (minor) are absent (supplementary figure 10).
    e. Catheter to artery **transition**: correct labelling of the catheter tip vs coronary artery origin (supplementary figure 11).
3) **Artifacts**
    a. **Coronary**: no non-coronary structures are labelled as part of the coronary tree (supplementary figure 12).
    b. **Catheter**: no non-catheter structures are incorrectly labelled as part of the catheter (supplementary figure 13).
        - These two artifacts' criteria are not applicable to the small catheter-artery transition area.

In order to provide an objective assessment, these criteria were scored by a panel of three Interventional Cardiologists (MNM, ARF, PCF), of whom two (ARF, PCF) took no part in any stage of the annotation/training process. Discrepancies were solved by agreement. All images were graded across all groups: baseline human segmentation, enhanced human segmentation, baseline AI and enhanced AI. During the grading process, the image's group was blinded.

Lastly, because the above-mentioned criteria are not equally important, a Global Segmentation Score (GSS – 1,5 to 100 points) was devised, taking into account the relevance of each criterion as defined by the three experts (Table 4). The panel was also asked to select which of the two AI models was preferred for each image, regardless of final score.

Table 4: scoring metrics for application of the Global Segmentation Score.

Criteria	Cathether Vs Coronary Relative Weight	Individual Criteria Relative Weight	Points
Main Vessel Segmentation	70%	40%	28,0
Main Vessel Gaps		10%	7,0
Catheter to Artery Transition		15%	10,5
Branch Segmentation		20%	14,0
BranchGaps		5%	3,5
Coronary Artifacts		10%	7,0
Catheter Segmentation	30%	40%	12,0
Catheter Gaps		10%	3,0
Catheter Artifacts		15%	4,5
Catheter Location		5%	1,5
Catheter Thickness		30%	9,0
Total			100

2.7. Statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. To assess the association between qualitative (categorical) variables the Chi-Square test was used. To assess differences in quantitative variables we used the Mann-Whitney test (two independent groups) or the Kruskal Wallis test (multiple independent groups). A p-value < 0,05 was used for statistical significance, except for multiple groups comparisons, where we used a p value of < 0,01. SPSS 27 was used for analysis.

2.8. Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local Ethics' Committee.

3. Results

3.1. Baseline dataset

We included 416 images from 69 patients (Table 5). With two human and two AI datasets, 1664 processed images were generated.

Table 5: baseline clinical characteristics of patients from whom images were analysed.

Factor	N +/- SD or N(%)
Age	67 +/- 11
Sex (male)	54 (78%)
Hypertension	56 (81,2%)
Diabetes mellitus	27 (39,1%)
Dyslipidemia	39 (56,5%)
Smoker (past or present)	26 (37,7%)
Chronic coronary syndromes	50 (72,5%)
Acute coronary syndrome	19 (27,5%)
Revascularization during/after CAG	21 (30,4%)

3.2. Performance assessment

3.2.1. Non-medical metrics

Results are outlined in Table 6. These scores indicate that the enhanced AI was generally superior to the baseline AI. Segmentation performance was good and consistent across arteries, as indicated by the high mean and low standard deviation of the DSC. For the catheter, performance was lower and much less consistent.

Table 6: Generalized Dice Score (GDS) and class-wise Dice Similarity Coefficient (DSC) obtained by the baseline (BAI) and enhanced AI (EAI) models. Results presented as mean ± standard deviation.

	BAI	EAI
GDS	0.9234 ± 0.0361	0.9348 ± 0.0284
Artery DSC	0.8904 ± 0.0464	0.9134 ± 0.0410
Catheter DSC	0.7526 ± 0.1998	0.7975 ± 0.1836

3.2.2. Clinical performance

Overall performance – individual criteria assessment (Supplementary table 1)

Coronary segmentation: the main vessels were correctly segmented in almost all cases across groups. Minor gaps occurred rarely in the baseline human segmentation and both AI models, although there was a small but non-significant improvement with the enhanced AI vs baseline AI.

Branches' segmentation was also correct almost always in all groups, albeit less so than main vessel segmentation. There was a small yet significant improvement with the enhanced AI vs baseline AI.

Minor branch gaps were quite common, with very significant differences between AI and human models. While the Enhanced AI performed numerically better than the baseline AI, it still produced small gaps in nearly 2/3 of cases.

Coronary artifacts were very uncommon in human annotations and were usually minor imperfections in catheter/coronary crossovers. They were common and usually minor in both AI models, although there was a very significant improvement with the enhanced AI vs baseline AI (14,4% vs 25,7%).

Catheter/artery transition: the baseline human segmentation failed in 12% of cases and the enhanced human segmentation missed 3,8%. The baseline AI produced a higher error rate (19,7%), but the enhanced AI was numerically more often correct than the baseline human segmentation, sometimes correctly identifying the transition where humans failed (Figure 20).

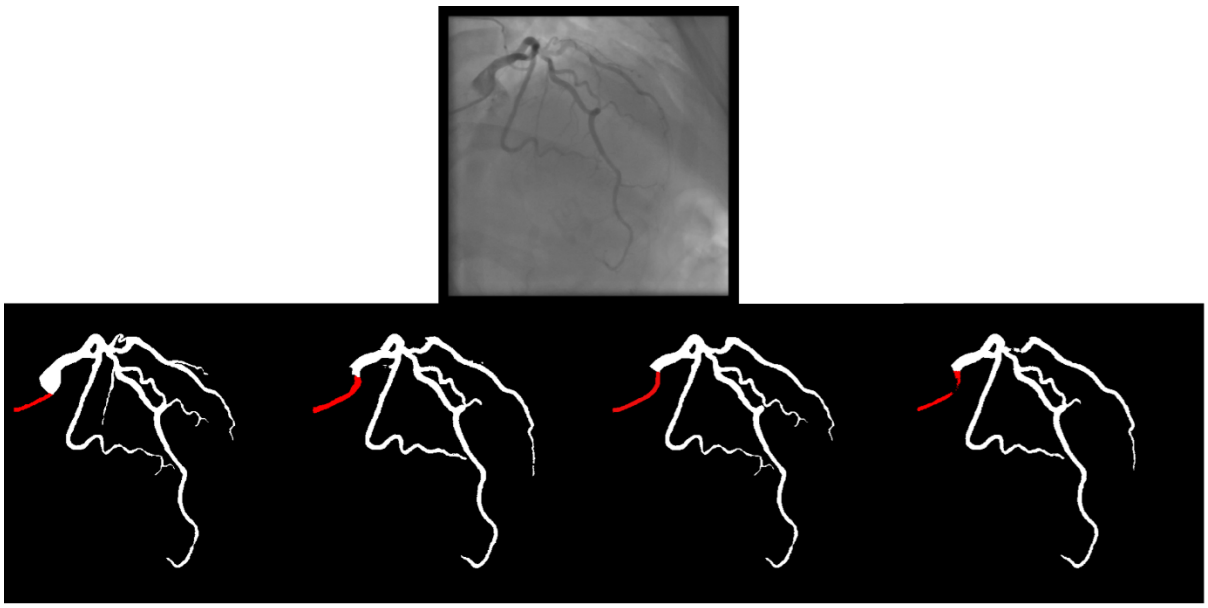


Figure 20:(left to right): the first human segmentation incorrectly labels contrast backflow as coronary. The baseline AI model improves on the human segmentation, but is still not perfect. The enhanced human model perfectly segments the transition. The enhanced AI model is hampered in catheter segmentation, but identifies the transition correctly.

Catheter segmentation: the baseline human segmentation produced thickness imperfections (usually mildly engorged catheter) in 13,9% of cases, but otherwise, the segmentation was almost always correct regarding other criteria.

The baseline AI produced low error rates in main body segmentation. However, artifacts, usually quite minor and in the vicinity of coronary segments, occurred very frequently (41,1%). Another common error was catheter thickness (36,3%), often overestimating catheter size.

The enhanced human segmentation significantly improved on thickness issues, although imperfections persisted in 6,2% of cases.

The enhanced AI produced better results than the baseline AI model regarding catheter thickness (correct in 96,4%), also surpassing both human models (although the difference was not statistically significant when compared to the enhanced human segmentation). However, the performance of the enhanced AI otherwise decreased in all other criteria, especially regarding minor gaps, which became much more common (3,1% in the baseline AI model to

23,3%). Even main body segmentation was significantly affected, although successful in the vast majority of cases (86,5%). Despite this, most failure cases would still permit catheter identification, as major gaps often occurred distally in areas of contrast backflow. Artifacts and location issues also numerically worsened slightly in the enhanced AI vs baseline AI.

Overall performance – GSS assessment and expert preference (Table 7)

Human models outperformed AI models. Enhanced models surpassed baseline models. The difference was statistically significant for all comparisons. GSS was very high for both AI models- the enhanced AI reached an average of 90 points.

With regards to expert preference, the enhanced AI model was preferred in 300 (72%) cases, the baseline AI model in 100 (24%) and in 16 (4%) cases no AI model was preferred.

Table 7: performance by group according to GSS (significance at p <0,05 for paired differences and p < 0,01 for multiple comparisons).

GSS	Group				p-value						
	BH	EH	BAI	EAI	Between all*	BH vs EH**	BAI vs EAI**	BH vs BAI**	EH vs EAI**	BH vs EAI**	EH vs BAI**
Mean +/- SD	96,9 +/-5,7	98,9 +/- 3,1	86,1+/ -10,1	90 +/- 7,6	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**<0,001**
Median (IQR)	100 (9)	100 (0)	87,5 (9)	92 (9,5)							

Performance by coronary artery – individual criteria assessment (Supplementary table 2)

There was a trend towards better performance in the RCA, both regarding human and AI groups. The most notable and statistically significant differences occurred in catheter transition (regarding both AI models and the baseline human segmentation) and catheter segmentation (both AI models performed better in the RCA). Branch gaps were quite less frequent in the RCA with the enhanced AI model. Other differences, even if statistically significant, were very small.

Performance by coronary artery – GSS assessment (Supplementary table 3)

All models scored very high for both arteries. There were very minor statistically significant differences for the baseline AI model only, favouring RCA segmentation.

Considering expert preference:

- RCA: the enhanced AI was preferred in 109 (68,6%) cases, the baseline AI was preferred in 43 (27%) and in 7 (4,4%) cases no AI model was preferred.
- LCA: the enhanced AI was preferred in 191 (74,3%) cases, the baseline AI was preferred in 57 (22,2%) and in 9 (3,5%) cases no AI was preferred.

Performance according to angulation incidence – individual criteria assessment (Supplementary table 4, Supplementary table 5)

Given the large amount of data, the fact that there were no significant differences in the vast majority of cases and for the sake of readability, only statistically significant differences are shown in the tables. Overall, the impact of incidences in model performance was limited, and affected almost exclusively the AI models.

Performance according to angulation incidence – GSS assessment (Supplementary table 6, Supplementary table 7)

Differences were minor and only statistically significant for human performance in less common incidences (PA views for the LCA and PA cranial for the RCA).

4. Discussion

4.1. Overall considerations

Baseline human segmentation was generally correct. Catheter/coronary transition and catheter thickness errors were the most common. Poor individualization due to contrast backflow, catheter curves and human fatigue all likely contributed.

Enhanced human segmentation was nearly perfect. Mild transition issues persisted, emphasizing the task's difficulty. As this model was actually a combination of the best of baseline human segmentation and baseline AI, it also demonstrates how AI can help improve human performance.

These human imperfections, even if small, highlight the need for rigorous quality control during and after the final results, rather than assuming a "perfect" ground truth (human annotation). This an inherent limitation to the annotation of medical images, as the sheer amount of cumbersome work is error prone.

The baseline AI performed CAG segmentation successfully, yet was affected by the same two issues of the baseline human segmentation – transition and catheter thickness. The effort to correct these when developing the enhanced AI was fruitful regarding the former, but produced mixed results for the latter. Impact on transition performance was impressive, as the enhanced AI at times even achieved correct assessments where humans failed (figure 20). However, it seems the gain in catheter thickness accuracy was offset by losses in other catheter segmentation tasks. Lastly, every aspect of coronary segmentation improved in the enhanced AI, which performed better than baseline AI. The differences between the two AI models also highlight how relatively small differences in the ground truth can impact relevantly on AI training.

It may seem surprising that catheter segmentation was less successful than coronary segmentation. However, while intuitively one may think that catheter segmentation is an easier task and therefore the results would have been better for this task, from a machine learning perspective that is not the case. In particular, segmentation performance is highly dependent on the frequency of each class. Rarer classes, or ones that occupy smaller areas, are interpreted by the model as being less likely to appear. Furthermore, during training, the lower the number of pixels belonging to a particular class, the lower the penalty for segmenting that class incorrectly. Even though we used a loss function designed to mitigate this phenomenon, the

poorer segmentation of less common classes (the catheter, in this case) is still evident in the results.

RCA segmentation was easier than LCA, but differences were quite small and less than expected, considering its greater anatomical simplicity. Angulations also had a relatively small impact both in human and AI performance and small observed differences may be attributed to specific issues that are more common in certain incidences: contrast backflow (less problematic in PA or RAO caudal); coronary/catheter crossovers (such as *spider* or extreme RAO cranial – Figure 21); proximity of bone (such as RCA LAO views); smaller samples of some incidences, such as PA cranial; uncommon catheter "pathways", such as the femoral approach, which sometimes produces a central vertical outline.

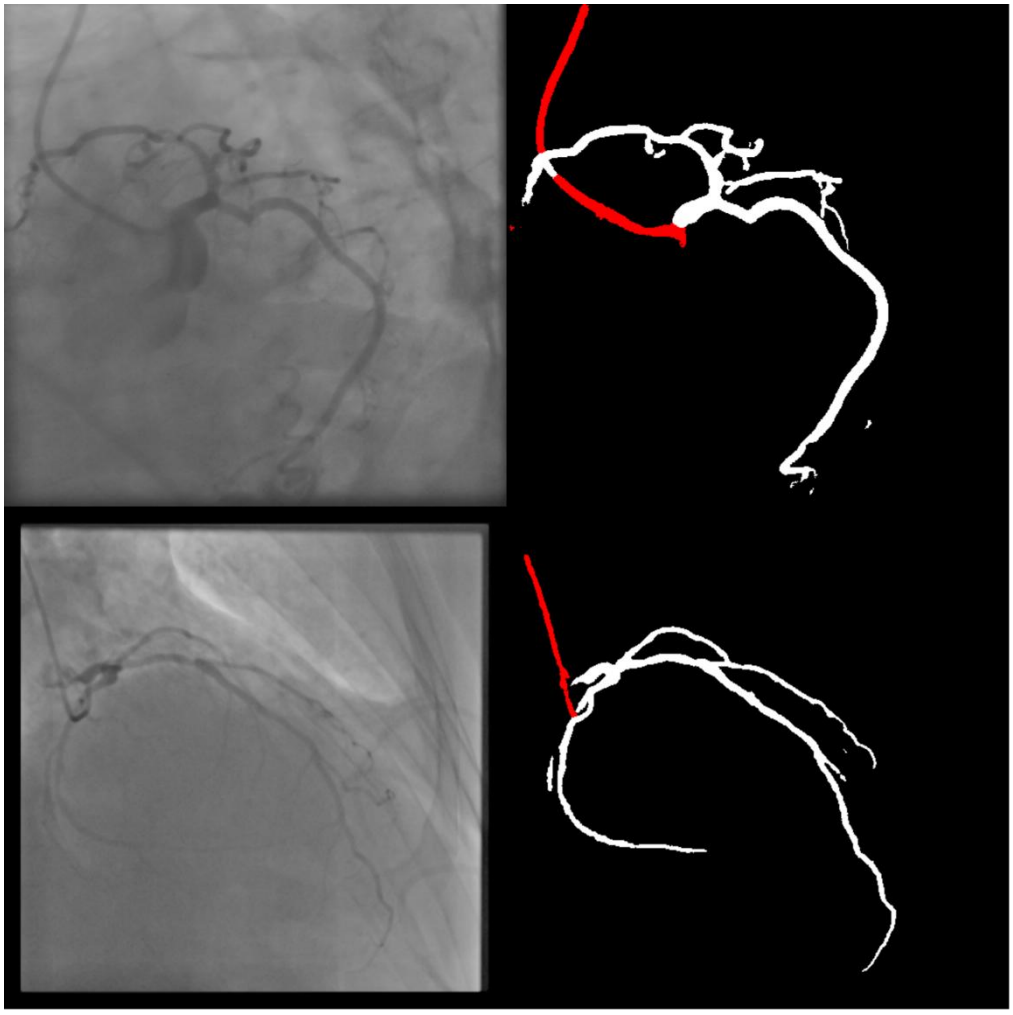


Figure 21: crossovers in spider (above) and extreme RAO cranial (below) views generating artifacts.

Globally, both AI models achieved a very high DSC, with higher performance in artery segmentation than in catheter segmentation, supporting the results of qualitative clinical assessment. When factors are weighed in terms of their perceived relevance – as assessed by the GSS - both performed very well. The enhanced AI scored an average of 90 points, meaning it provided 90% of what experts deemed most relevant when looking at a CAG. By all measures, the enhanced AI was the better model. However, the fact that differences between the two AI models were not large and that the enhanced AI was preferred in most, but not all cases, highlights the difficulty in improving an already good performance.

4.2. Other studies with AI applied to CAG segmentation/interpretation

Few studies regarding coronary artery segmentation based on AI technologies have been published in medical/biology journals so far.

Su Yang *et al*[248] successfully developed AI models capable of segmenting CAG. Their dataset was larger (3302 images / 2042 patients) and was also annotated by (two) expert physicians. Different incidences were also used. However, they focused solely on segmenting specific segments of major vessels with at least mild (> 30%) stenotic lesions. Nor the branches nor the catheter were segmented, leading to a much simpler problem than the one addressed in this work.

Two other works[249,250], from the same baseline dataset, also developed AI-based CAG segmentation. Their dataset was also larger (4904 images from 170 videos). However, the annotations were performed by medical students and no details are provided regarding patient subset, target vessel or incidences.

Very recently, Tianming Du *et al*[247] published the results of a broad work. They focused on two tasks: CAG segmentation and special lesion morphology identification (calcium, thrombus, among others). For the former task, which overlaps with ours, they used a very large dataset of 13373 images distributed across 10 incidences (six LCA and four RCA), annotated by 10 qualified analysts. This was an all-comers study, rather than focusing on patient subsets. They too annotated catheter/arteries, and additionally marked different coronary segments. Their model is impressive as judged by shown images, even distinguishing between contrast backflow, catheter and coronary. However, they did not specify the exact criteria for

segmenting the coronary tree and their exact metrics make it difficult to assess exactly how their models performed in detail regarding segmentation.

While all the above-mentioned groups have worked with datasets larger than ours, our work has several unique features: (1) there was a medical rationale for vessel size segmentation; (2) results were assessed from a set of criteria defined by experts, capturing the quality of the segmentation from Interventional Cardiologist's eyes; (3) human annotations were also graded, rather than assuming a perfect human ground truth; (4) specific segmentation tasks were assessed individually, enabling insights into strengths and weaknesses of AI and human models alike; (5) results were also assessed globally with the GSS, by factoring the relevance of each criteria, permitting a broad, simple appreciation of results. Furthermore, the ability to perform high-quality segmentation in a system trained using less data provides relevant evidence that more advanced AI systems can be effectively applied even in situations where the available data is limited.

4.3. Limitations

This is a single center retrospective dataset, with a single image per projection and a smaller sample size than some previously published manuscripts. The images comes from the same angiography devices (Siemens Artis) and thus we have not yet tested on our models on images obtained with other equipment or image settings.

We have not yet conducted formal assessment on how well the models perform in segmenting specific stenosis severities. Our models are also yet to be tested for specific vessel disease types (calcium, thrombus), clinical settings (chronic total occlusion, ST-elevation myocardial infarction).

We have not yet assessed the performance of AI models on an external "validation" cohort. There are several reasons for this. We aimed to compare AI and human results in detail first and assess the exact performance of AI models for each segmentation task. A validation dataset would comprise a new set of images which would not undergo human segmentation, thus impeding comparison with human performance. Also, validation implies that a metric be available for comparing results. Because the Dice methods require a ground truth human annotation for comparison, and the GSS was developed and applied for the first time for this paper, we felt a suitable metric was not yet available for performing validation prior to the current analysis. Lastly, AI models are continuously and dynamically improving. As we are currently working on further testing and enhancing current AI models (view Future Directions

and Implications section bellow), we felt performing external validation at this stage was premature.

The exclusion of cardiac devices/cardiac surgery and other foreign objects renders our models not yet applicable to such cases. We did not, however, excluded cases with previously implanted stents.

Lastly, focusing specifically on patients undergoing invasive physiology assessment may have created a bias, limiting a broader application of the models to other patient subsets.

We are currently working to address all these issues in future work.

4.4. Future directions and implications

CAG segmentation in itself is not a final objective but rather an arguably essential milestone for developing AI systems capable of CAG analysis and interpretation. These results should therefore be regarded as a first step, rather than a final deployment tool. While not yet mature for immediate clinical application, the results of both AI models are already relevant, providing a framework for the future which can be built upon.

Further steps include testing the models for stenosed segments, which will be critical for clinical. We then aim to test our models with a "validation" cohort using new angiograms. Sub-segmentation, automatic anatomical identification and physiology are also areas for future research.

We will also strengthen the capabilities of our models further by broadening our training base to other patient and lesion subsets and focusing on particular issues where there is still room for improvement, as identified by our uniquely detailed analysis.

Our results also provide insight on which human tasks are most challenging, which can be of use to others.

The GSS is the first score of its kind for assessing the quality of segmentations in CAG. By providing a reasonably objective and quantitative clinical measurement, it may be used as a benchmark for comparing and validating results across research groups.

Lastly, while conventional segmentation software does exist, it is not without its limitations, and only by developing AI systems can we compare and improve both in the future. The potential implications of AI to Interventional Cardiology are immense, and we envisage a cath lab of the future where all of these insights render the human eye more objective, improving patient care.

5. Conclusions

We successfully developed two AI models capable of good quality automatic CAG segmentation as assessed by both the GDS, DSC and the GSS. The latter and its individual criteria provided a feasible, reasonably objective and quantifiable way of assessing the results from an expert's perspective.

The enhanced AI model outperformed the baseline AI model in coronary segmentation tasks as well as globally. With regards to catheter segmentation tasks, the enhanced AI model improved on the task of catheter thickness but performed less well in other catheter segmentation tasks. Both human segmentations were superior to both AI models, but only the enhanced human segmentation, built by combining the best of baseline human segmentation and baseline AI, achieved a near perfect GSS.

These results provide a relevant framework for building upon, potentially enabling future clinical application.

6. Acknowledgements

MNM was responsible for conceptualization, data gathering, processing and analysis, results interpretation and paper drafting. JLS was responsible for technical and AI tasks, data and image processing, model implementation and training. TR and BS were responsible for data gathering, processing and analysis. ARF and PCF were responsible for data analysis and results interpretation. ALO was responsible for supervising the work of JLS, having directly taken part in the same tasks. FJP was responsible for supervising the work of MNM, having directly taken part in the same tasks.

All authors revised the paper critically for important intellectual content, gave final approval for its publication and agree to be accountable for all respects of its accuracy and integrity.

Part II. Coronary X-ray Angiography Segmentation using Artificial Intelligence: a Multicentric Validation Study of a Deep Learning Model

Miguel Nobre Menezes[1,2*], João Lourenço Silva[3], Beatriz Silva[1,2],Tiago Rodrigues[1,2], Cláudio Guerreiro[4], João Pedro Guedes[5], Manuel Oliveira Santos[6,7], Arlindo L. Oliveira[3], Fausto J. Pinto[1,2]

(1) Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon, Faculdade de Medicina, Universidade de Lisboa, Av Prof. Egas Moniz, 1649-028 Lisboa, Portugal

(2) Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Av Prof. Egas Moniz, 1649-028 Lisboa.

(3) INESC-ID / Instituto Superior Técnico, University of Lisbon

(4) Centro Hospitalar de Vila Nova de Gaia

(5) Unidade de Hemodinâmica e Cardiologia de Intervenção, Serviço de Cardiologia, Centro Hospitalar Universitário do Algarve, Hospital de Faro

(6) Unidade de Intervenção Cardiovascular, Serviço de Cardiologia do Centro Hospitalar e Universitário de Coimbra, Praceta Professor Mota Pinto, 3004-561 Coimbra

(7) Faculdade de Medicina da Universidade de Coimbra, R. Larga 2, 3000-370 Coimbra

Int J Cardiovasc Imaging. 2023 Jul;39(7):1385-1396. doi: 10.1007/s10554-023-02839-5.

1. Introduction

The application of artificial intelligence (AI) to coronary angiography (CAG) has only been ascertained in very few medical/biology publications[247–250]. While the possibilities of such an approach are vast, the first step is arguably to produce accurate segmentation of CAGs, i.e., clearly identifying the coronary tree while excluding other structures.

We have previously published the first results of deep learning models capable of good quality CAG segmentation[273]. In this paper, we aim to validate the results, by applying the model to a new, previously unseen, dataset of coronary angiographies from multiple centers. A well-known validated software was used as reference for segments with non-occlusive lesions, where detailed measurements were undertaken, while also applying the previously described Global Segmentation Score for broad assessment of segmentation quality[273].

2. Methods

2.1. Participating centers and equipment

Four centers from across Portugal participated in this study. Images were acquired in Siemens Axiom Artis and Philips Azureon equipment.

2.2. Inclusion criteria

Retrospective selection of consecutive patients who had undergone CAG and percutaneous coronary intervention (PCI) and/or underwent invasive physiology assessment (Fractional Flow Reserve and/or other indexes), within a 1-month period of 2022, regardless of clinical context (i.e. both acute and chronic coronary syndrome). This ensures the model was tested in a real-world context where revascularization was either being considered or performed, thereby excluding a population with normal or near-normal coronary arteries.

2.3. Exclusion criteria

We excluded cases where any of the following applied:

5) Patients with previous cardiac surgery, cardiac devices or other sources of potential artifact
6) Absence of coronary lesions 50-99% stenosis by visual estimation (i.e. single-vessel ST-elevation myocardial infarction – STEMI - or chronic total occlusions – CTO alone)
7) Poor image quality
8) Unclear individualization of lesion outline with no overlapping vessels
9) Unsuccessful automatic measurements with validated software (details below)
10) Unsuccessful software extraction and superimposition of lesion markers on segmented image (details below)

2.4. Image selection

For each selected lesion, a single end-diastolic frame with clear outline definition of the vessel and target lesion was selected. More than one segment per patient and/or image could be used. With an original training dataset of 416 images as previously published[273] we aimed to have a validation dataset of at least 100 images.

2.5. Brief description of previous work and AI model

In our previous work[273] we trained AI models for CAG segmentation using 416 images from patients undergoing physiology or PCI in a single center. The images were manually annotated by a small group (two Cardiology Fellows and an Interventional Cardiologist, who both annotated and supervised the process) and continuously reviewed and corrected, in order to minimize heterogeneity and errors.

We then performed segmentation using an encoder-decoder fully convolutional neural networks based on the U-Net[269], commonly used in medical image segmentation. These are composed of an encoder for extracting image features and a decoder to process those features and produce segmentation masks. To derive the best approach for this task, we conducted a comparative study of encoder and decoder architectures, which resulted in the proposal of the EfficientUNet++, a computationally efficient and high-performing decoder architecture[270], which obtained the best results when combined with an EfficientNet-B5 encoder[271].

To ensure fair evaluation and minimize any bias induced by the input data, each model was tested on data it had not seen during training. The dataset was thus split at the patient level, into 13 subsets of approximately 32 angiograms each. Each subset's segmentation was performed using a neural network trained exclusively on the remaining data. This enabled the assessment of the segmentation results for the entire cohort, as the usual splitting into a training and testing dataset would have yielded a much smaller group of images for result assessment. The training hyperparameters, namely the number of training epochs and the learning rate decay schedule, were set on the first train-test split, using 1 of the 12 training data subsets for validation. The selected values were then used on every other train-test split, and to train the model on the whole training set of the first split. We also considered cross-validation, but it would be very compute-heavy.

This process resulted in an early AI model, which was then further improved by a second round of manual annotation, where the annotators corrected the resulting imperfections, thereby producing a final training dataset. An “enhanced” model was then trained once again using the same process with the new improved annotated dataset, yielding superior results to the early model, with a final Generalized Dice Score of 93,48/ +/- 2,84%. While we continue to work on improving our model, because the aim of this study is to validate the aforementioned “enhanced” model as previously published[273], no additional training was performed.

2.6. Original images analysis and segmentation

A well-established and validated software (CAAS Workstation 8.5.1) capable of semi-automatic segmentation and Quantitative Coronary Angiography (QCA) was used to generate a reference dataset for comparison. Because it is especially important for a model to correctly segment diseased segments, QCA analysis was performed in selected segments with a stenosis severity of 50-99% by visual estimation. For QCA measurements, calibration was performed either automatically (based on the DICOM information) or by measuring the catheter (5 or 6 Fr), provided it was clearly visible and measurable. The region of interest was then selected and automatic QCA measurements were undertaken.

For each region of interest where successful automatic QCA measurements were undertaken, the lesion diameter, reference diameter, diameter at proximal obstruction border and diameter at distal obstruction border were recorded. The diameter stenosis percentage was calculated as follows: ((reference diameter – lesion diameter) / reference diameter) x 100[107]. No manual adjustments were accepted, in order to exclude human bias or human-induced imperfection. If the automated outline and measurements were not clearly accurate by visual inspection, the case was excluded (Figure 22).

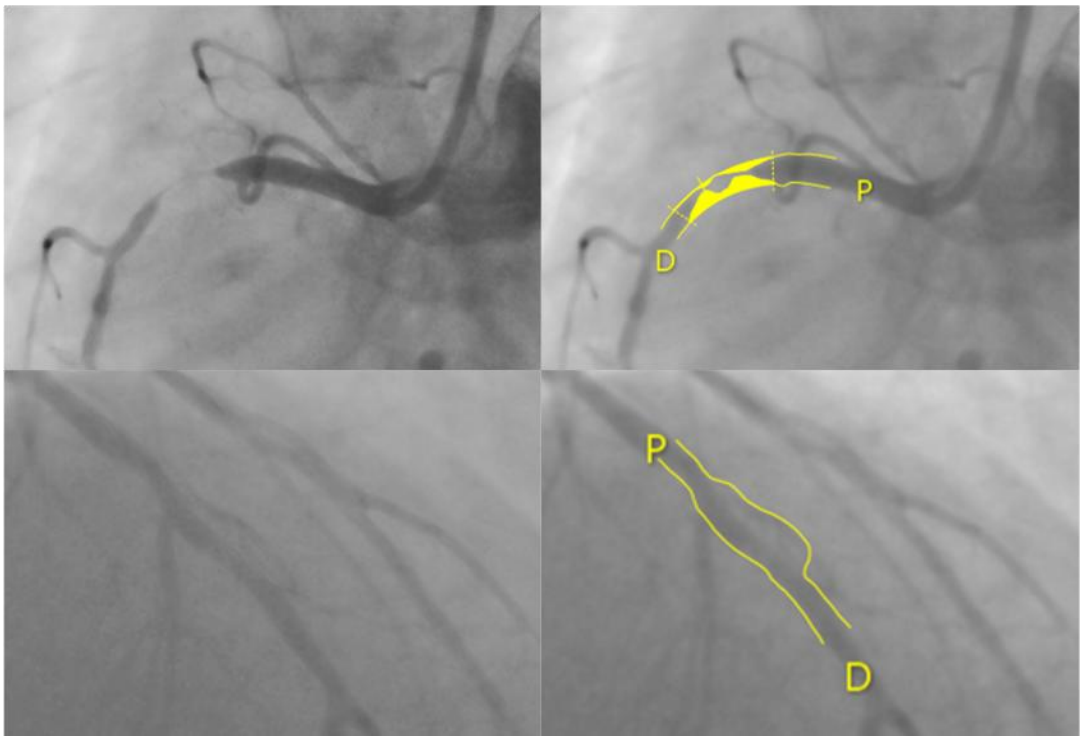


Figure 22: Two examples of failed auto-QCA analysis. In the right coronary artery, a subocclusive lesion is visible (upper left image). The software fails to track the lesion accurately (upper right image). In the left anterior descending artery, the software tracks a collateral rather than the main vessel on the left border (original image - bottom left, failed tracking - the bottom right).

The original images (i.e., without the measurement annotations generated by the CAAS software) were then segmented using our best AI model to date[273], which segments the coronary tree in white and the catheter in red. This process is fully automatic and the only required human input is the image itself. These images were used for testing only, not training.

2.7. Performance assessment

2.7.1. Diameters and percentage diameter stenosis

A dedicated python script was written to extract the CAAS markers and superimpose them on the segmentation obtained by the model. The lesion diameter, diameter at proximal obstruction border and diameter at distal obstruction border were then measured using a dedicated python script as well, by verifying the superimposition of the markers with the coronary tree. Because the reference diameter does not exist in the segmented image (which only contains the coronary artery tree and catheter), the CAAS-generated value was used. Percentage stenosis was then calculated using the same equation. Finally, we also compared the measured catheter diameter on the original image versus the segmented image with another adaption of the same script, by measuring the distance between the two parallel lines generated in the original image from the CAAS software. The resulting measurements obtained in the original and the segmented images were then compared.

2.7.2. Overlap between original and segmented images

A dedicated python script was also used for assessing the overlap between the original and the segmented images in the region of interest, using the CAAS output as reference. Pixels were then classified as follows:

- True positive (TP): a pixel marked as coronary in both the segmented and original image
- False positive (FP): a pixel marked as coronary only in the segmented image
- True negative (TN): a pixel marked as non-coronary in both the segmented and original image
- False negative (FN): a pixel marked as non-coronary only in the original image

Using this classification, the following parameters were calculated:

- Accuracy: ([TP+TN]/[TP+TN+FP+FN])
- Sensitivity: TP / (TP + FN)
- Specificity: TN/(TN+FP),
- Positive Predictive Value: TP / (TP + FP)
- Negative predictive value TN/(TN+FN)
- Intersection over Union (IoU): TP / (TP + FN + FP)
- Dice Score: 2TP / (2TP + FN + FP)

2.7.3. Global Segmentation Score

While the above-mentioned criteria offer a detailed account of the model's accuracy, they do not provide a broad overview of the quality of segmentation as assessed by experts in CAG interpretation (i.e. Cardiologists). As a result, we have previously developed the Global Segmentation Score (GSS), which we have previously applied on the original CAG dataset used to train the AI model (details on its application on supplementary data file)[273].The GSS was scored by consensus by four Interventional Cardiologists (one from each contributing center). Figure 23 summarizes the above-mentioned steps for assessing coronary segmentation.

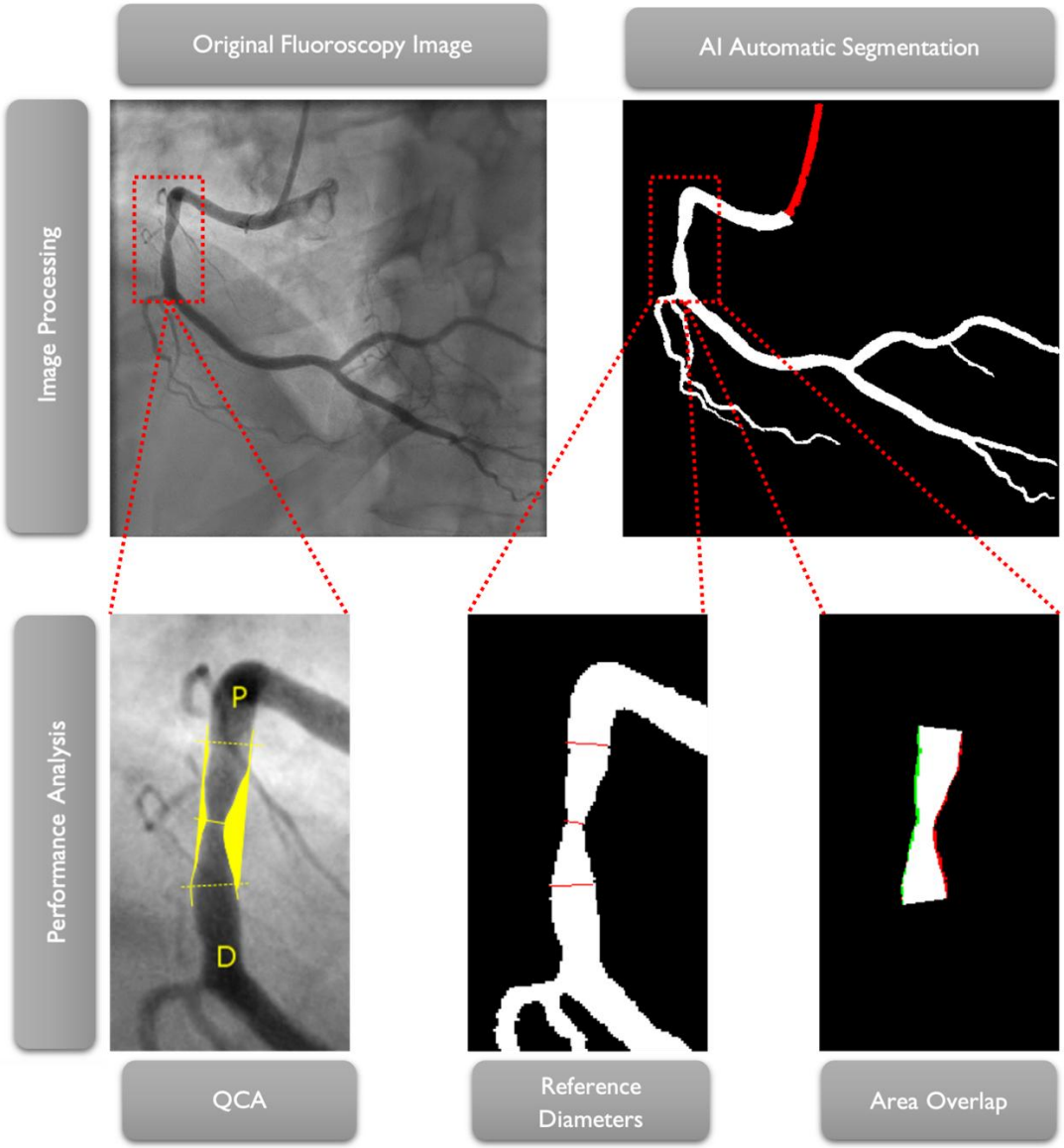


Figure 23: Overview of the segmentation and analysis process. Top left: Baseline CAG of a right coronary artery. Top right: AI automated segmented image. Bottom left: automatic QCA analysis image output in detail. Bottom middle: transposition of the lesion markers on the segmented image in detail. Bottom right: area overlap between the region of interest in the auto-QCA and the segmented image; white pixels are true positives; green pixels are false negatives; red pixels are false positives.

2.8. Statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. Quantitative variables are shown in average ± standard deviation (if normally distributed) or median (interquartile range) if non-normally distributed. If distribution was normal, we used the paired samples T-test to assess for differences in related samples quantitative variables. If distribution was not normal, we used the Mann-Whitney test (two independent groups) or the Kruskal Wallis test (multiple independent groups) to assess for differences in quantitative variables. A p-value < 0,05 was used for statistical significance. SPSS 27 was used for analysis.

2.9. Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local Ethics' Institutional Review Board.

3. Results

3.1. Baseline characteristics

We included 123 measurements from 117 images, from a total of 90 patients (flowchart in Figure 24; clinical data on Table 1).

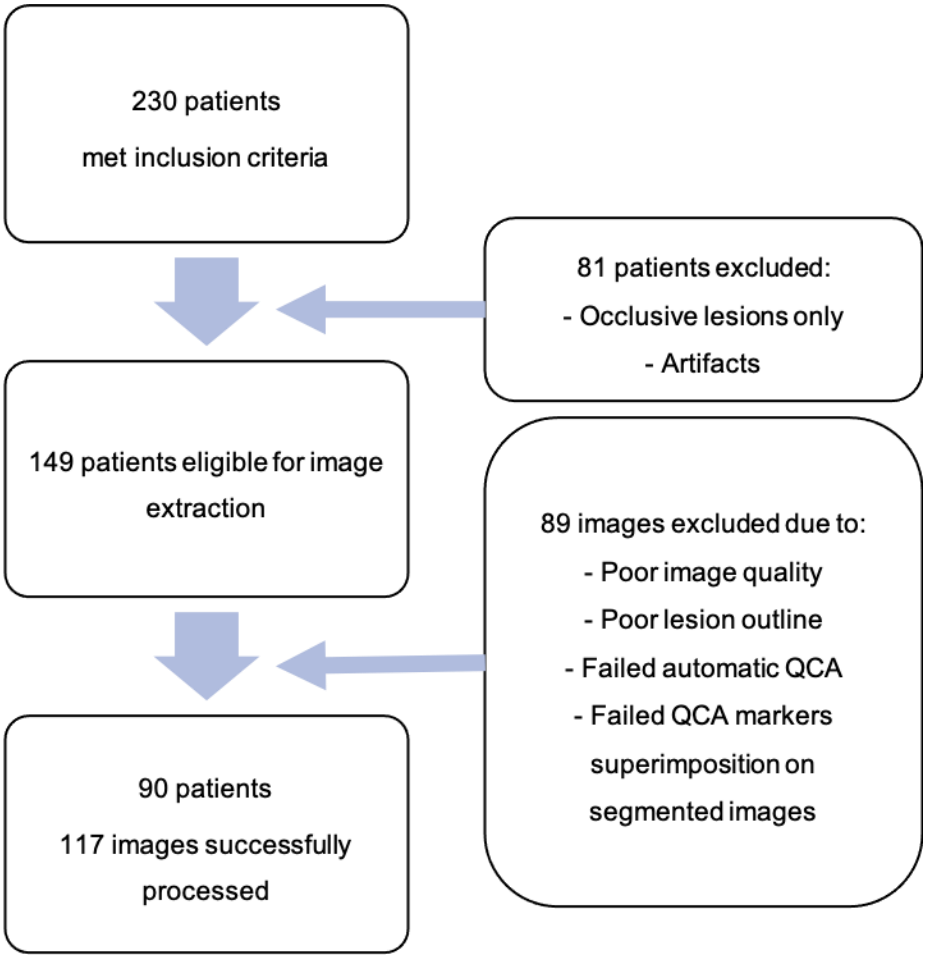


Figure 24: flowchart of patient and image selection.

Table 8: Clinical characteristics of included patients.

Factor	N +/- SD or N(%)
Age	65 +/- 12
Sex (male)	73 (81%)
Hypertension	62 (68,9%)
Diabetes mellitus	28 (31,1%)
Dyslipidemia	55 (61,1%)
Smoker (past or present)	50 (55,6%)
Chronic coronary syndrome	37 (41,1%)
Acute coronary syndrome	53 (58,9%)
Revascularization during/after CAG	76 (84,4%)
Invasive Physiology during procedure	19 (21,1%)

The left anterior descending artery (LAD) was the most common target vessel (three measurements were taken on diagonals, two emerging proximally and one emerging in the middle segment of the LAD; all were taken on the proximal segment of the collateral), with measurements taking place more frequently in the middle and proximal segments. As measured by QCA, most lesions had a 50-69% diameter stenosis, with a minority of ≥ 70% lesions (Table 9).

Table 9: distribution of target vessel and lesion severity. LAD: Left Anterior Descending Artery; RCA: Right Coronary Artery; CX: Left Circunflex Artery

Parameter			N (%)
Target Vessel	LAD	Proximal	19
		Middle	23
		Distal	8
		Total	50 (41)
	RCA	Proximal	9
		Middle	24
		Distal	8
		Total	41 (33)
	CX	Proximal	9
		Middle	21
		Distal	2
		Total	32 (26)
Lesion severity		≥70%	22 (18)
		50-69%	58 (47)
		< 50%	43 (35)

3.2. Performance

3.2.1. Diameters and percentage diameter stenosis

Detailed metrics of images (Figure 25) are depicted in Table 10 and Table 11. There were no significant differences for all parameters except for diameter at proximal obstruction border, where the median difference between groups was 0,19 mm. All difference parameters (Table 10) had a non-normal distribution, with the interquartile range demonstrating that there is a clear predominant difference towards the lower-end values, as the 25th quartile is either 0 or very close to 0.

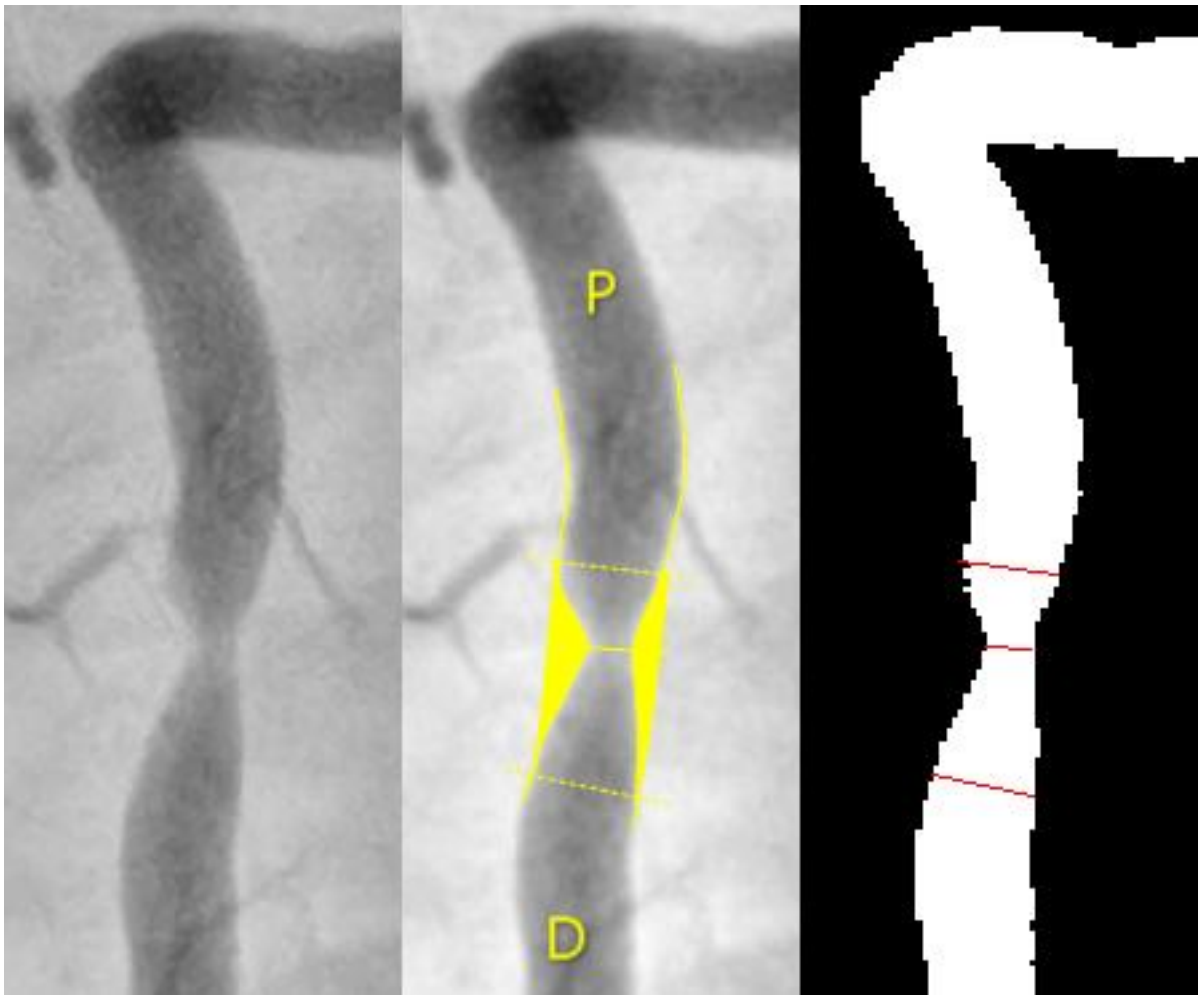


Figure 25: comparative view of a right coronary artery (56% stenosis by QCA). Left-to-right: original image, auto-QCA, transposition of lines (proximal border diameter, lesion diameter and distal border diameter) to segmented image. QCA: Quantitative Coronary Analysis.

Table 10: detailed measurements between the original and the segmented images. Values shown as mean ± standard deviation. AI – artificial intelligence. *Paired samples T-test

Parameter	Original Image	AI Generated Segmented Image	P-value*
Diameter Stenosis (%)	56 ± 13	55 ± 13	0,071
Diameter at lesion (mm)	1,06 ± 0,39	1,08 ± 0,37	0,146
Diameter at proximal obstruction border (mm)	2,27 ± 0,54	2,09 ± 0,53	**< 0,01**
Diameter at distal obstruction border (mm)	2,19 ± 0,56	2,15 ± 0,58	0,133

Table 11: Median differences between the original and segmented images. Values shown as median (IQ 25th – 75th).

Parameter	Difference
Diameter Stenosis (%)	4,5 (0 – 7,7)
Diameter at lesion (mm)	0,10 (0 – 0,17)
Diameter at proximal obstruction border (mm)	0,19 (0,09 – 0,28)
Diameter at distal obstruction border (mm)	0,10 (0 – 0,19)

There were no significant differences across stenosis severity (Supplementary table 8) or target vessel (Supplementary table 9). There were also no significant differences considering across centers (Supplementary table 10 and Supplementary table 11).

With regards to the catheter diameters (Figure 26), results are shown on Supplementary table 12. A significant number of cases (26/117 - 22%) had to be excluded, either because of collimation (rendering the catheter not visible – 8 cases) or segmentation gaps leading to inaccurate border definition (18 cases). The latter occur because the model focuses especially on segmenting the distal part of the catheter for correctly identifying the transition between catheter and coronary, whereas in the original images calibration occurred predominantly in less distal portions. Because the presence of two groups (5 and 6 Fr) of catheters renders the overall distribution of the sample non-normal, the two groups were analysed separately. There were no significant differences between the original and segmented images. Again, the difference parameter had a non-normal distribution, with the interquartile range demonstrating that there is a clear predominant difference towards the lower-end values.

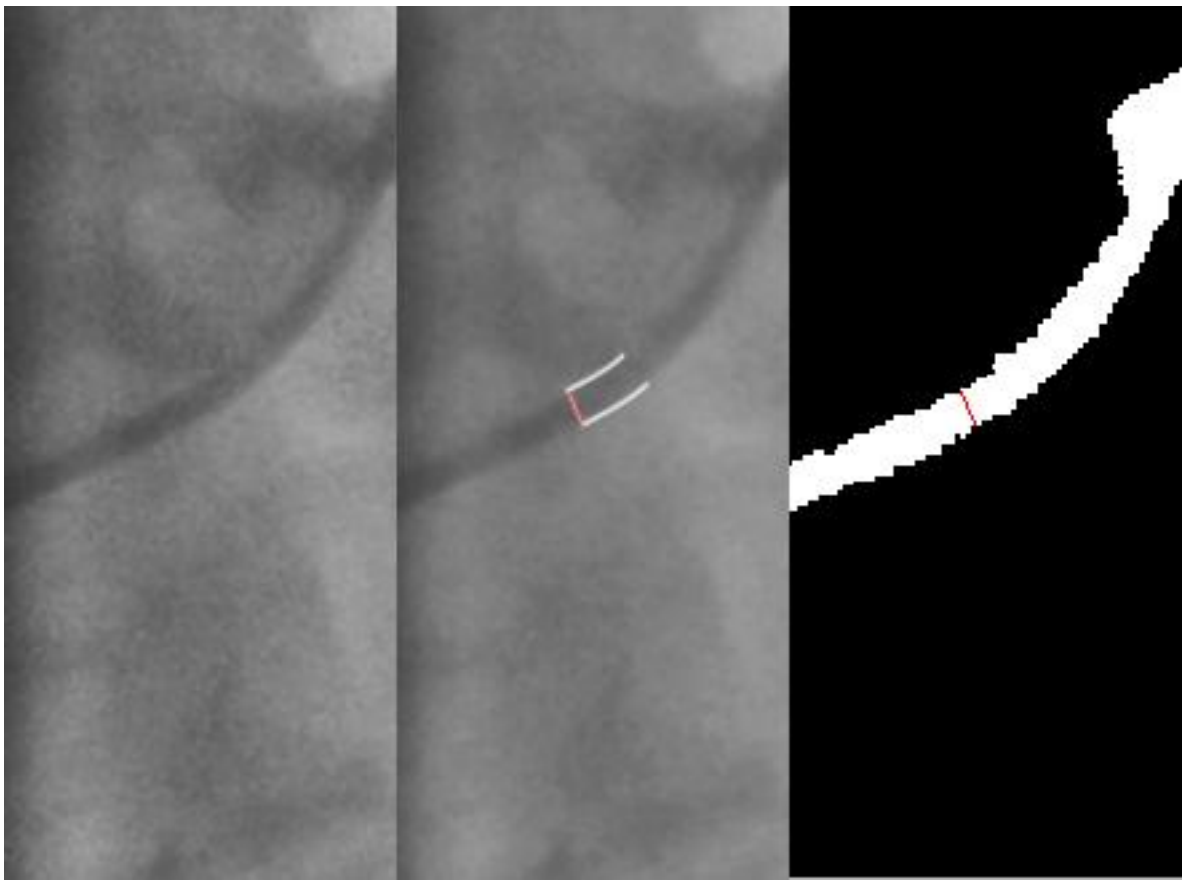


Figure 26: catheter segmentation assessment. Left-to-right: original image, auto-border detection by reference software, transposition of lines in proximal border to segmented image.

3.2.2. Overlap between original and segmented images

Results are detailed on table 12. The model scored ≥ 90% in all metrics (Figure 27). There were some significant differences between target vessel (Supplementary table 13) and stenosis severity (Supplementary table 14) which, in absolute terms, were between 1 and 3%. There were no differences between centers (Supplementary table 15).

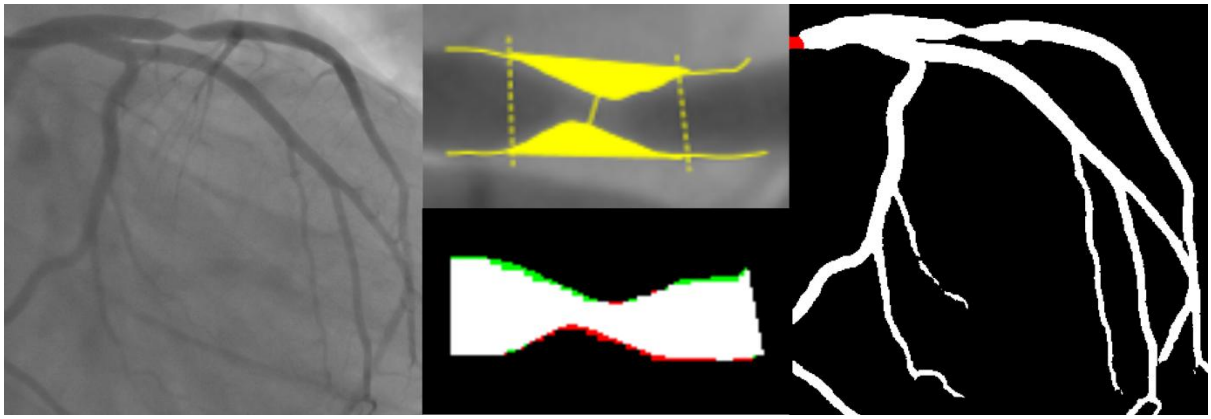


Figure 27: Area overlap in a and left anterior descending 64% stenosis (as measured by Quantitative Coronary Analysis).

Table 12: Overlap metrics. Values shown as median (IQ 25th – 75th).

Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Intersection over Union (%)	Dice Score (%)
99,9 (99,9 – 99,9)	95,1 (92,8 – 96,4)	99,9 (99,9 – 99,9)	94,9 (93,1 – 96,5)	99,9 (99,9 – 99,9)	90,1 (87,6 – 91,7)	94,8 (93,4 – 95,7)

3.2.3. Global Segmentation Score

Results are shown on Supplementary table 16. The model scored well above or close to 90% in most criteria. Catheter gaps were common, usually due to contrast backflow impeding proper visualization of such portions. Catheter artifacts were common and mild gaps in distal parts of small collaterals were quite common as well.

N is lower than overall measurements due to assessment of more than one lesion per image and 8 cases of collimation where the catheter could not be scored, thereby excluding those cases from assessment.

4. Discussion

4.1. Main findings

A deep learning AI segmentation model was capable of fully automatic accurate CAG segmentation, as checked by a reference segmentation obtained with validated software and also when assessed by a broad assessment score we previously developed[273].

Diameters at both healthy segments (proximal and distal lesion borders) and diseased segments (diameter at maximum obstruction zone) were similar between the two groups, with statistically significant differences only at the proximal obstruction border. However, in absolute terms, the difference was very small (0,19 mm, a < 10% difference considering the proximal diameter in either group) and we therefore believe it is unlikely to be of clinical significance. The stenosis severity as assessed by percentage stenosis only differed by < 5% in absolute terms, a difference not meaningful either statistically or clinically. The latter is perhaps the single most important finding, as percentage diameter stenosis is the fundamental criteria assessed in clinical practice for proceeding with either revascularization or functional testing, as recommended in current guidelines[34]. Importantly, there were no significant differences in performance regarding target vessel, stenosis severity or centers.

When considering the overlap between the segmented image and the original image, accuracy, specificity and negative predictive value scored close to 100%. This was expected, because most of the image is composed by background rather than artery. As a result, we believe metrics that do not take into account true negatives provide a more faithful indication of the actual model performance. In that regard, sensitivity and positive predictive value still scored quite high, at approximately 95%. The metric that more directly assesses the true overlap between the original and segmented images in the region of interest (correctly identifying all of the vessel while avoiding non-artery pixels) is the intersection over union criteria, which fell just short of 90%. Lastly, the Dice Score puts greater emphasis on the fundamental task of segmentation – correctly identifying the target structure i.e. true positives – in this case, the coronary tree. With an average score of approximately 95%, while also considering all the remaining metrics, we believe our model can be described as accurate. Importantly, the Dice Score in our previous study was 93%, thus very similar to what we now found[273]. There were statistically significant differences in the IoU and Dice Scores between

target vessel stenosis severity. Notwithstanding, the absolute differences were very minor (around 1-2%) and therefore of little or no clinical relevance.

With regards to the GSS, our model achieved a high score with a median of 92/100 points, exceedingly similar to what we had previously described in the dataset used to train and develop the model. The model scored very high in almost all tasks, while maintaining minor imperfections with regards to mild gaps in collateral branches, which were very frequent. Catheter segmentation was not as good as coronary segmentation, as usually small catheter artifacts or gaps in the vicinity of the coronary tree origin were common. This was due not only to contrast backflow, but also because of how AI models are trained and function. Indeed, performance is very dependent on class frequency. Because the catheter is a less frequent class (i.e. corresponds to much fewer pixels), the models receive less penalty for errors regarding its segmentation when compared to the coronary tree. This is partly mitigated by the use of an appropriate loss function, but the imbalance nevertheless persists to some extent. Once again, this was very similar to what we saw in the training dataset[273]. With regards to precise catheter measurements, the differences between original and segmented images (for both 5 and 6 Fr catheters) were not statistically significant, suggesting the catheter's segmentation, from a calliper precision point of view, is accurate. However, due to the above-mentioned limitations and to a small number of images where only a small portion (or none at all) of the catheter was discernible, our sample was somewhat reduced, thereby limiting this assessment.

4.2. Other studies in the field

There are very few studies published in medical/biology journals to date where a comparison with our results can be made. With regards to the GSS in particular, no similar application has ever been undertaken, to our knowledge.

The largest published study[247] included a dataset of 1050 images distributed across all incidences and vessels for performance evaluation. An average 98% accuracy was obtained. While specificity and negative predictive values scored very highly, sensitivity and positive predictive value came closer to 80%. The performance was slightly inferior in more distal vessels. Intersection over union or Dice Score were not reported. Importantly, however, that study's evaluation used the baseline human annotation as reference, rather than an external validated software, thereby not enabling the identification of bias or imperfections which might have become embedded in their AI model. In our previous study, we demonstrated that even with a small group of annotators and continuous review of the annotations, there is always

some degree of imperfection in human annotation[273], hence the relevance of comparing against an automated and validated external software. Additionally, the reported accuracy focuses on the overlap across the entire coronary tree rather than the percentage stenosis of diseased segments. This is advantageous in the sense that a globally accurate performance can be tested. Notwithstanding, we believe testing only for diseased segments actually renders the comparison more demanding. This is because the segmentation of stenotic segments is harder from a technical point of view and also due to the fact that the number of true positive pixels is necessarily smaller in such segments – leading to a lower likelihood of true positives. Whichever interpretation is made, it is clear that an exact comparison with Du *et al*[247] is not possible. However, broadly speaking, the accuracy of both models seems quite high and our model seems at least as accurate, if not more.

Su Yang *et al*[248] also produced AI models for CAG segmentation. Their validation dataset was somewhat larger (181 images), but their performance seems slightly lower, with all overlap metrics generally scoring just short of 90% and a Dice Score of 89%. Importantly, they also only segmented diseased segments, with a minimum lesion of 30% and used the same reference software as we did. Thus, their results are more directly comparable to ours and our model seems to have superior performance. Two other works[249,250], from the same baseline dataset, also went on to develop AI-based CAG segmentation, this time with a validation dataset of 550 images. While the model performed well, with an accuracy of 98% and a sensitivity of 87%, they also based their validation dataset on human annotation of the coronary tree without using external software. Thus the above-mentioned considerations for Du *et al*[247] also apply.

Recently, Gao *et al*[251] published the results of a CAG segmentation model trained on only 130 images. Their methodology, however, is somewhat different, since they combined features from deep learning segmentation models' features and non-AI image filters to perform pixel-wise classification using gradient-boosting decision trees[274] and deep forests[275]. Their results also show good performance, with a Dice Score of 87,4%, sensitivity of 90,2% and specificity of 99,2%. This highlights that merging deep learning with traditional computer vision methods can yield good results, when working with relatively small datasets. However, no external validation software was used and the whole coronary tree was evaluated. As a result, once more, the previous considerations for Du *et al*[247] apply.

Other works in the application of AI to coronary segmentation are primarily technical and featured in engineering publications. A detailed review of these falls outside the scope of

this paper and can be consulted in our previous technical publication[270]. However, some considerations regarding these provide further contextualization of our findings.

Xian *et al*[246] used a very large dataset of 3200 manually annotated images and experimented with the U-Net architecture as well, with a sensitivity of 90,1%, positive predictive value of 89,8% and Dice Score 90%. However, the annotations were undertaken with a specific software for the purpose of coarsely signaling the vessel route, and focused only on the main vessels. Since we achieved higher performance metrics, it seems a smaller but higher quality dataset, with very precise and cumbersome manual annotations, may be a better approach.

Yang *et al*[243] have obtained a sensitivity, positive predictive value and Dice Score of 91,3%, 92,5% and 91,9%, respectively, by using popular image classification backbones pre-trained on ImageNet instead of the U-Net's encoder, while also using a modified generalized dice loss function. Their findings were influential in our training method, as we used a combination of their proposed loss function and the focal loss[276]. Other authors have explored the use of dense connections, improving on the performance of the standard U-Net[244]. This approach is also present in the U-Net ++[245], which we used in our approach.

In all of the above studies, metrics regarding vessel diameters were not performed. Thus, a direct comparison with this study regarding those is not possible. M'hiri *et al* addressed the issue of CAG diameter measurements, when dealing with the issue of diameter variation during the cardiac cycle due to vessel distensibility. They focused mainly in measuring specific segments of the coronary tree, as we did. However, they used a graph-based segmentation method, then tracked the changes across the cardiac cycle using a spatio-temporal segmentation method. They obtained a Dice Score of 98%, with a very small diameter mean error (0,18 mm) [277]. However, they did not focus on diseased regions. While this study is not focused on AI methods, it highlights that other methods may be of use for accurate CAG segmentation, potentially in combination with AI tools[251].

In light of all these studies, the performance of our model seems at least as good, if not better, than previously proposed AI models. We believe this is related to its neural network architecture, which was carefully chosen over a series of experiments[270], taking into consideration the invaluable contributions of previously mentioned studies. In addition to that, we also believe that our manual annotations methodology was essential, as it allowed us to

obtain a highly reliable training dataset: a small number of annotators (to reduce heterogeneity) well trained in the interpretation of coronary angiograms; very careful review of annotations with recurrent iterations of quality checks and improvements; and further manual improvement of the already accurate segmentation images produced by an earlier AI model, thus combining the best of AI and human annotations into a final training dataset, as mentioned in the methods section and previous publication[273].

4.3. Limitations

Our study is not without limitations. Despite the multicentric approach, our dataset is relatively small when compared to previously published studies. We also only tested the model performance against validated software in diseased locations, rather than on the whole coronary tree. Therefore, we cannot affirm that the performance would be identical in the remaining areas. However, as previously explained, segmenting zones with lesions is actually more challenging for the model than segmenting broad, mostly healthy segments. In addition to that, we did not find differences regarding target vessel or lesion severity. Plus, considering the results of the GSS, the overall performance regarding CAG segmentation was quite appropriate. Thus, we believe that it is unlikely that performance would be significantly different had we tested for the whole coronary tree. Importantly, if we had chosen to segment whole vessels, it would be very likely that some manual corrections had to be undertaken, which might induce bias or imperfections in the reference images. Hence, the decision to proceed as described was deliberate. The assessment of catheter segmentation was also more limited than that of the coronary tree, as described above.

The exclusion of potential sources of artifacts from devices or previous cardiac surgery means our model is not yet applicable to such patients. Notwithstanding, we didn't exclude cases with previous implantation of stents, but we did not perform detailed measurements on such segments.

The total number of patients/images who fully met exclusion criteria was somewhat high, thereby limiting the final amount of available images for analysis, which may raise questions as to whether this sample is representative of everyday CAGs and an therefore constitutes an adequate validation dataset. This was the result of somewhat stringent criteria, which we felt were nonetheless necessary due to basic feasibility (such as excluding single-vessel complete occlusion cases where QCA is not applicable, or excluding imaging artifacts for which the models are not yet trained), reduction of bias (such as not allowing for manual QCA correction), or excluding patients with normal/near-normal arteries (where testing would be

much less challenging or useful in future clinical application). Notwithstanding, we included patients consecutively rather than selectively and the clinical characteristics of included patients are in agreement with everyday clinical practice. We therefore believe our sample to be reasonably representative of real-world practice. Furthermore, we exceeded the minimum validation target of 100 images, yielding relative rates of training vs validation cases in agreement with other AI studies[247,248,250].

The imbalance in sample size limits the comparison between centers.

It has long been established that operators significantly differ in their interpretation of lesion severity and have a tendency to overestimate the importance of a stenosis[85,88,89,119,252], as we also saw in this study. Indeed, while visually all lesions were interpreted as > 50% stenosis, a significant amount of the sample actually had a < 50% stenosis, which further reflects the real-world nature of the dataset.

Lastly, the distance between the 2D centerline and the distance to the closest edge would have also been a good metric for assessing model performance in this setting. We did not perform such testing.

In light of all of the above, concerns may be raised regarding generalization from this dataset. However, we believe that the absence of statistically significant differences across all subgroups at least partially attenuates this concern.

4.4. Future directions

We are currently working in automatic anatomical interpretation, lesion severity based on auto-QCA and integration with physiology. We believe without effective segmentation models, none of these will be possible. Much like for human interpretation of CAG, separating the coronary arteries from everything else in the image is an essential first step. Our ultimate goal is to produce an intelligence augmentation tool that helps physicians perform a more objective and streamlined interpretation of CAG, hopefully contributing for better patient outcomes. As we continuously improve its performance, while also adding new capabilities, clinical application will potentially be possible in the near future, opening a new perspective and potentially more accurate method to assess coronary artery disease.

We are also continuously working to expand and improve the model, as segmentation alone is not a final goal in itself, but rather a fundamental step. We hope to release a public version in the near future, which other researchers may use for whichever application they may deem useful. Importantly, comparing or even merging with future models from other groups may also be very relevant. Since it uses an inherently data-hungry deep learning model, our

coronary artery segmentation system would surely benefit from training on a larger volume of data. Manual annotation of coronary angiography images, however, is very cumbersome and time-consuming, and therefore it is difficult to obtain much larger labeled datasets. Hence, significant improvements to the model could probably be achieved, for example, by using self-supervised learning on existing very large volumes of unlabeled data. These possibilities are described in detail in our previous technical publication[270].

4.5. Data Availability

Detailed full-scale study data cannot currently be made publicly available due to limitations imposed by national data protection regulations, as this is a retrospective study and no informed consent was obtainable regarding this particular analysis. Both our research team and others in the national scientific community are working to develop a framework where such would be possible. However, independent replication of our analysis is possible, given that the detailed description of our experimentations and relevant code is publicly available[270].

5. Conclusions

Our AI model was capable of accurate CAG segmentation when applied to a multicentric validation dataset, with no differences between target vessels or stenosis severity. This paves the way for future research and implementation for its clinical uses.

6. Acknowledgements

MNM was responsible for conceptualization, data/image gathering, processing and analysis and paper drafting. JLS was responsible for technical and AI tasks, data and image processing, model implementation and training. CR, JPG and MSO were responsible for data gathering and image analysis. TR and BS were responsible for data gathering, processing and analysis in the development of the original dataset. ALO was responsible for supervising the work of JLS, having directly taken part in the same tasks. FJP was responsible for supervising the work of MNM, having directly taken part in the same tasks.

All authors revised the paper critically for important intellectual content, gave final approval for its publication and agree to be accountable for all respects of its accuracy and integrity.

Part III. Segmentation of X-Ray Coronary Angiography with an Artificial Intelligence Deep Learning Model: Impact in Operator Visual Assessment of Coronary Stenosis Severity

Miguel Nobre Menezes[1,2*], Beatriz Silva[1,2], João Lourenço Silva[3], Tiago Rodrigues[1,2], João Silva Marques[1,2], Cláudio Guerreiro[4], João Pedro Guedes[5], Manuel Oliveira-Santos[6,7], Arlindo L. Oliveira[3], Fausto J. Pinto[1,2]

(1) Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon, Faculdade de Medicina, Universidade de Lisboa, Av Prof. Egas Moniz, 1649-028 Lisboa, Portugal

(2) Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Av Prof. Egas Moniz, 1649-028 Lisboa.

(3) INESC-ID / Instituto Superior Técnico, University of Lisbon

(4) Centro Hospitalar de Vila Nova de Gaia

(5) Unidade de Hemodinâmica e Cardiologia de Intervenção, Serviço de Cardiologia, Centro Hospitalar Universitário do Algarve, Hospital de Faro

(6) Unidade de Intervenção Cardiovascular, Serviço de Cardiologia do Centro Hospitalar e Universitário de Coimbra, Praceta Professor Mota Pinto, 3004-561 Coimbra

(7) Faculty of Medicine, University of Coimbra, Pólo das Ciências da Saúde, Unidade Central, Azinhaga de Santa Comba, Celas, 3000-354 Coimbra, Portugal

Catheter Cardiovasc Interv. 2023 Oct;102(4):631-640. doi: 10.1002/ccd.30805.

Epub 2023 Aug 14.

1. Introduction

The assessment of the severity of coronary stenosis is essential for revascularization decisions. In clinical practice, operators often begin by assessing the percentage diameter stenosis (%DS) of lesions, which can either be estimated visually ($\%DS_{VE}$) or by means of direct semi-automatic measurement with quantitative coronary angiography (QCA - $\%DS_{QCA}$) [107]. However, multiple studies have shown that visual inspection tends to result in average higher percent diameter stenosis than QCA, with heterogeneity across operators and/or hospitals[85–89]. Disagreements of lesion severity by visual inspection vs QCA may be clinically relevant, as they have been associated to the likelihood of clinical events[119].

Furthermore, the assessment of the functional significance of stenosis by means of fractional flow reserve (FFR) has been proven to be superior to that of angiography alone regarding clinical outcomes, despite the fact that physiology-guided revascularization results in lower rates of percutaneous coronary intervention, once again highlighting that operators tend to overestimate the severity of lesion severity by visual estimation[93]. As a result, current guidelines strongly emphasize the role of physiology or ischemia testing in the assessment of coronary lesion severity, rather than angiography alone[34,103]. Despite this, the adoption of physiology remains low[182,183].

As a result, non-invasive and automatic tools that reduce the heterogeneity of coronary angiography (CAG) interpretation are desirable[278]. Artificial Intelligence (AI) may be of use for such a task, but few studies are available in medical/biology publications regarding its application for CAG[247–250]. We have recently developed AI models capable of accurate CAG segmentation[273,279]. In this study, we sought to evaluate how CAD lesion severity is perceived by operators when CAGs are viewed in AI-segmented vs fluoroscopy images, using QCA as reference.

2. Methods

2.1. Previous work and study population

We have previously trained AI models for CAG segmentation based on manual CAG annotation of patients undergoing invasive physiology assessment (Fractional Flow Reserve and/or other indexes) or PCI, with an original sample of 416 images[270,273]. Recently, we published the results of our validation study with an additional dataset of 117 images[279]. Briefly, consecutive patients who had undergone (PCI) and/or invasive physiology assessment in four centers from across Portugal were selected. The images were then automatically segmented with our AI model. Lesions were measured by QCA with a validated software (CAAS Workstation 8.5.1) in the original images, which were then compared to the segmented images. We have shown that the AI-generated segmentation was highly accurate, with no significant differences between percentage diameter stenosis in fluoroscopy vs segmented images, across all degrees of lesion severity, target lesion or fluoroscopy equipment[279].

In this study, we chose to use the validation cohort for assessing the impact of segmentation in the perception of lesion severity, given that significant differences between segmented and original images are thus *a priori* excluded, as they have all been previously measured. As a result, potential differences can be attributed to visual perception rather than actual dataset discrepancies.

2.2. Stenosis severity assessment

One operator from each participating center was shown all images consecutively in random order. First, the fluoroscopy images were shown, followed by the AI segmented images. Two sessions, with at least a one-week interval, were scheduled for each of the two datasets and the order of randomization was different for each of the datasets, to avoid a carry-over effect. Operators had not seen any of the images prior the sessions and were blinded regarding clinical data, equipment or originating center. $\%DS_{VE}$ was visually estimated for each target lesion where QCA had been previously measured and operators were asked to provide a specific %DS value of their choosing, rather than a range interval.

Differences were then assessed both in terms of absolute %DS values overall, as well as according to three $\%DS_{QCA}$ Strata, which were defined as < 50%, 50 – 70% and > 70%.

2.3. Statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. Quantitative variables are shown in average ± standard deviation or median (interquartile range). To assess for differences in related samples quantitative variables we used the Wilcoxon test (paired samples) or the Friedman Test (multiple related samples). To assess differences in qualitative variables we used the Chi-Square test.

For further illustrating discrepancies between %DS across the angiography and segmentation groups, we calculated the $\%DS_{VE}$ / $\%DS_{QCA}$ plotted in a scatter graphic in the Y axis, against the $\%DS_{QCA}$ in the X Axis.

A p-value < 0,05 was used for statistical significance. SPSS 27 was used for analysis.

2.4. Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local Ethics' Committee.

3. Results

3.1. Baseline characteristics (Table 8, Table 9)

A total of 123 measurements (117 images), from a total of 90 patients, were included. Each operator performed %DS by visual estimation in both the angiography and segmented images, generating a total of 984 $\%DS_{VE}$ estimates. Most lesions had a $\%DS_{QCA}$ of 50-70%.

3.2. Lesion Severity Assessment

3.2.1. Overall Results (Table 13, Table 15, Figure 28)

Lesion severity by $\%DS_{VE}$ was estimated to be higher with angiography than with segmentation. While there were significant differences between $\%DS_{VE}$ and $\%DS_{QCA}$ in both groups, the overall absolute difference in %DS was lower in the segmentation group. Additionally, no statistical difference was found between the segmentation $\%DS_{VE}$ and the $\%DS_{QCA}$ for two of the four operators in the segmentation group, whereas the difference was significant for all operators in the angiography group.

When lesions were grouped in three strata of severity by QCA (< 50%, 50 - 70% and > 70%), agreement with QCA was generally low, albeit significantly higher (approximately double) with segmentation than angiography, both considering the overall sample and individual operators.

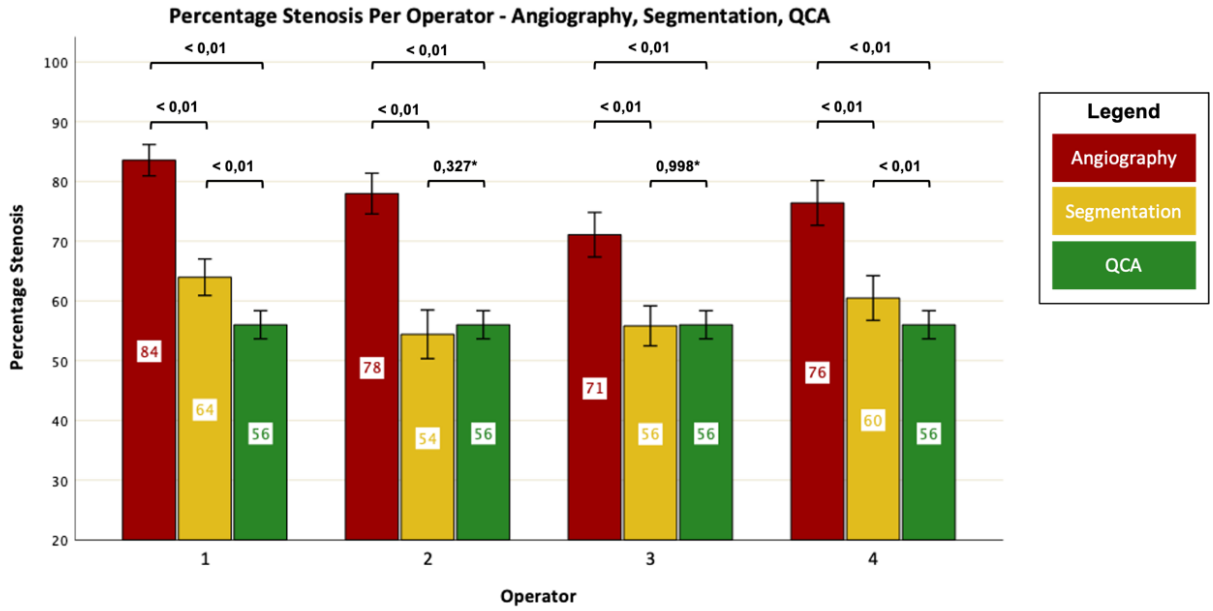


Figure 28: $\%DS_{VE}$ according to operator for both angiography and segmentation, compared with $\%DS_{QCA}$. Significant differences are always found between angiography and both other groups, but not between segmentation $\%DS_{VE}$ and $\%DS_{QCA}$ for operators 2 and 3. $\%DS_{VE}$: percentage diameter stenosis by visual estimation; $\%DS_{QCA}$: percentage diameter stenosis by QCA (Quantitative Coronary Analysis).

3.2.2. Results per QCA Severity Strata (Table 13, Table 14, Figure 29, Figure 30, Supplementary figure 14 to Supplementary figure 17)

For lesions with $\%DS_{QCA}$ > 70%, there was a statistically significant higher $\%DS_{VE}$ estimation in both the angiography and segmentation groups. In the segmentation group, $\%DS_{VE}$ values were lower and closer to $\%DS_{QCA}$ with no statistically significant differences for one operator.

There was a clear strata agreement between visual estimation and QCA with either angiography (100% agreement) or segmentation (88,6%).

For lesions with $\%DS_{QCA}$ of 50-70%, there was a very large and significant difference between $\%DS_{VE}$ with angiography, with median estimates of 80% or 90% across all operators and very low rates of strata agreement (0 – 19%).

When $\%DS_{VE}$ was undertaken with segmentation, differences with $\%DS_{QCA}$ were small and significantly different only for two of the four operators. The rates of strata agreement between visual estimation with segmentation and QCA were significantly higher when compared with angiography in the overall sample, with individual rates across operators between 39,7% and 51,7%.

For lesions with $\%DS_{QCA}$ < 50%, there were also significantly higher estimates $\%DS_{VE}$ with angiography for all operators and low rates of strata agreement (11,6 – 32,6%), but to a lesser degree than in the $\%DS_{QCA}$ 50-70% strata.

For $\%DS_{VE}$ with segmentation, there was no statistically significant difference with $\%DS_{QCA}$ overall and for two of the four operators, with higher rates of agreement (34,9 – 69,8%) than in the angiography group.

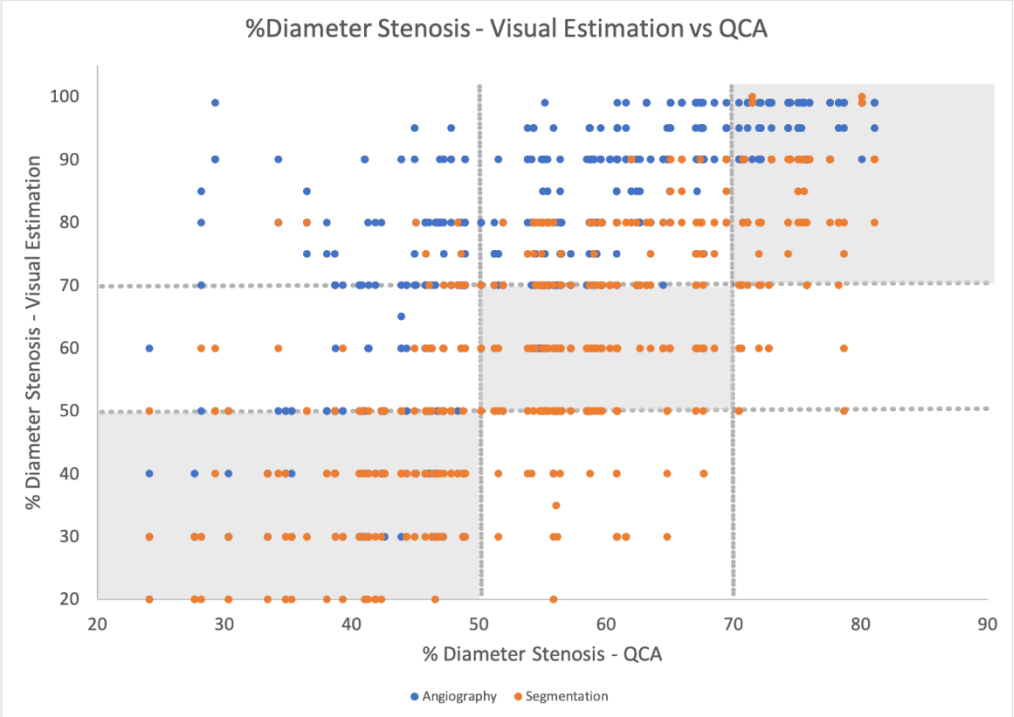


Figure 29: combined scatterplot of angiography and segmentation $\%DS_{VE}$ vs $\%DS_{QCA}$. The difference between the two $\%DS_{VE}$ is visually clear. $\%DS_{VE}$: percentage diameter stenosis by visual estimation; $\%DS_{QCA}$: percentage diameter stenosis by QCA (Quantitative Coronary Analysis).

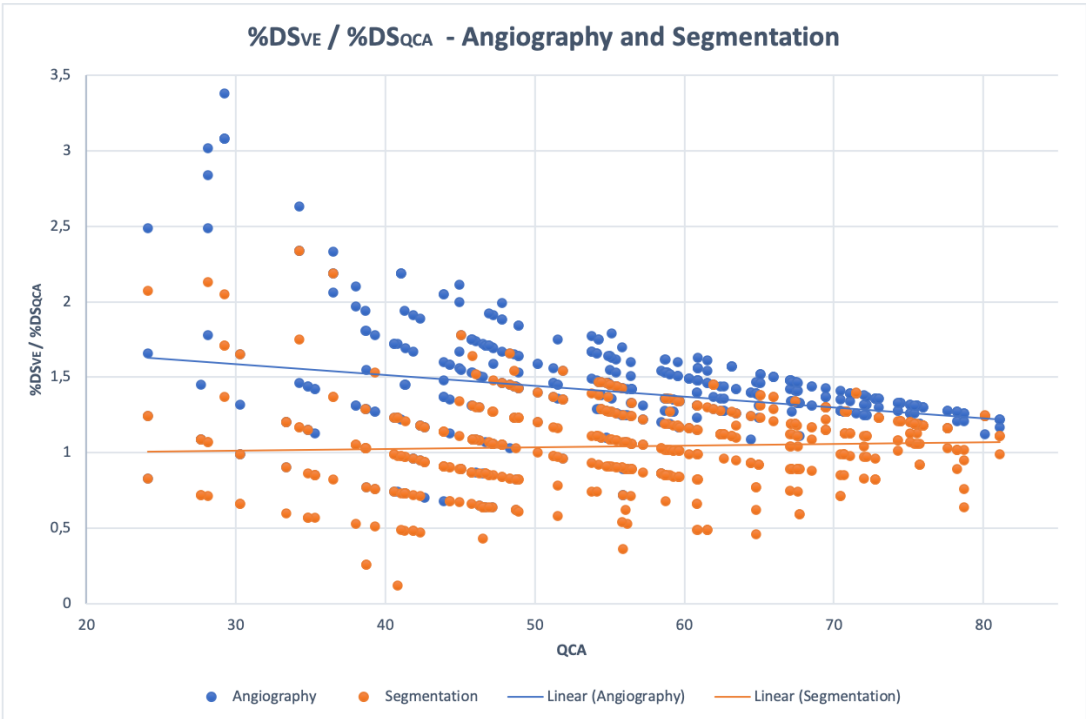


Figure 30: combined scatterplots of angiography and segmentation $\%DS_{VE}$ / $\%DS_{QCA}$ plotted in the Y axis vs $\%DS_{QCA}$ plotted in the X axis. The difference between the two $\%DS_{VE}$ is visually clear. $\%DS_{VE}$: percentage diameter stenosis by visual estimation; $\%DS_{QCA}$: percentage diameter stenosis by QCA (Quantitative Coronary Analysis).

3.2.3. Operator Heterogeneity (Table 13 to Table 15)

There were no significant differences across operators in the DS_{QCA} > 70% strata for either the angiography or segmentation groups. Considering the overall sample and the remaining strata, there were significant differences between operators in absolute $\%DS_{VE}$ values both in the angiography and segmentation group. However, in absolute terms, the differences were smaller in the segmentation group.

When considering agreement by QCA strata rather than absolute %DS values, there were significant differences in the angiography group overall and in the $\%DS_{QCA}$ 50 - 70% strata. For the segmentation group, significant differences were only found in the $\%DS_{QCA}$ < 50% strata.

Table 13: diameter stenosis assessed by visual estimation and QCA across operators, lesion severity strata and overall. QCA: Quantitative Coronary Angiography; $\%DS_{VE}$ – percentage diameter stenosis by visual estimation; $\%DS_{QCA}$ – percentage diameter stenosis by QCA; *Wilcoxon Test Paired Samples; **Friedman Test; significant p-values in bold.

		$\%DS_{VE}$ - Angiography					$\%DS_{VE}$ - Segmentation					$\%DS_{QCA}$
$\%DS_{QCA}$ Stratum	Group / Parameter	Operator 1	Operator 2	Operator 3	Operator 4	All	Operator 1	Operator 2	Operator 3	Operator 4	All	
All cases	%DS	84 +/- 15 90 (80 - 95)	78 +/- 19 80 (50 – 90)	71 +/- 21 70 (50 - 90)	76 +/- 21 80 (70 - 90)	77 +/- 20 80 (70 – 90)	64 +/- 17 60 (50 – 80)	54 +/- 23 60 (40 – 70)	55 +/- 19 50 (40 – 70)	60 +/- 21 60 (40 – 75)	59 +/- 20 60 (40 – 75)	56 +/- 13 56 (46 – 67)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	0,327	0,998	**< 0,001**	**< 0,001**	-
	Inter-Operator Difference**	**< 0,001**				-	**< 0,001**				-	-
> 70%	%DS	95 +/- 4 95 (90 - 99)	96 +/- 4 99 (90 – 99)	94 +/- 5 90 (90 - 99)	96 +/- 4 99 (90 - 99)	95 +/- 4 97 (90 – 99)	82 +/- 12 80 (78 – 90)	82 +/- 12 83 (70 – 90)	79 +/- 14 80 (70 – 90)	82 +/- 11 83 (74 – 90)	81 +/- 12 80 (71 – 90)	74 +/- 3 74 (71 – 76)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**0,008**	**0,016**	0,088	**0,004**	**< 0,001**	-
	Inter-Operator Difference**	0,115*					0,334*					-
50-70%	%DS	88 +/- 8 90 (80 - 95)	83 +/- 12 80 (75 – 90)	76 +/- 15 80 (70 - 90)	84 +/- 12 90 (75 - 90)	83 +/- 13 85 (75 – 90)	66 +/- 14 60 (58 – 80)	60 +/- 17 60 (50 – 70)	59 +/- 13 60 (50 – 70)	67 +/- 13 70 (60 – 75)	63 +/- 15 60 (50 – 75)	60 +/- 5 59 (55 – 64)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**0,002**	0,707	0,837	**< 0,001**	**< 0,001**	-
	Inter-Operator Difference**	**< 0,001**					**< 0,001**					-
< 50%	%DS	71 +/- 17 80 (60 - 80)	62 +/- 20 65 (40 – 80)	53 +/- 17 50 (40 - 70)	57 +/- 20 50 (40 - 70)	61 +/- 20 60 (43 – 80)	52 +/- 13 50 (40 – 60)	33 +/- 12 30 (20 – 40)	40 +/- 10 40 (30 – 50)	40 +/- 17 40 (30 – 50)	41 +/- 15 40 (30-50)	41 +/- 7 43 (38 – 47)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	0,116	0,538	0,380	-
	Inter-Operator Difference**	**< 0,001**					**< 0,001**					-

Table 14: agreement between $\%DS_{QCA}$ strata and $\%DS_{VE}$ strata between angiography and segmentation, across operators and overall. QCA: Quantitative Coronary Angiography; $\%DS_{VE}$ – percentage diameter stenosis by visual estimation; $\%DS_{QCA}$ – percentage diameter stenosis by QCA; NA – Not Applicable; *Chi-Square Test; significant p-values in bold.

$\%DS_{QCA}$ Strata / Group		Angiography All	Segmentation All	Angiography Operator 1	Segmentation Operator 1	Angiography Operator 2	Segmentation Operator 2	Angiography Operator 3	Segmentation Operator 3	Angiography Operator 4	Segmentation Operator 4
All cases	$\%DS_{VE}$ / $\%DS_{QCA}$ Strata Agreement (n / %)	148 (30,1)	297 (60,4)	27 (22,0)	65 (52,8)	35 (28,5%)	79 (64,2%)	47 (38,2)	80 (65,0)	39 (31,7)	73 (59,3)
	Angiography vs Segmentation* P Value	**< 0,001**		**< 0,001**		**0,007**		**< 0,001**		**< 0,001**	
> 70%	$\%DS_{VE}$ / $\%DS_{QCA}$ Strata Agreement (n / %)	88 (100)	78 (88,6%)	22 (100)	29 (90,9)	22 (100)	19 (86,4)	22 (100)	18 (81,8)	22 (100)	21 (95,5)
	Angiography vs Segmentation* P Value	NA		NA		NA		NA		NA	
50-70%	$\%DS_{VE}$ / $\%DS_{QCA}$ Strata Agreement (n / %)	17 (7,3)	107 (46,1)	0 (0)	30 (51,7)	2 (3,4%)	23 (39,7%)	11 (19)	32 (55,2)	4 (6,9)	22 (37,9)
	Angiography vs Segmentation* P Value	**0,009**		NA		0,243		**0,008**		0,113	
< 50%	$\%DS_{VE}$ / $\%DS_{QCA}$ Strata Agreement (n / %)	43 (25,0)	112 (65,1)	5 (11,6)	15 (34,9)	11 (25,6)	37 (86,0)	14 (32,6)	30 (69,8)	13 (30,2)	30 (69,8)
	Angiography vs Segmentation* P Value	**< 0,001**		**0,001**		0,590		**0,003**		**0,004**	

Table 15: significance of Inter-Operator differences overall and across $\%DS_{QCA}$ strata, in the angiography and segmentation groups. QCA: Quantitative Coronary Angiography; $\%DS_{VE}$ – percentage diameter stenosis by visual estimation; $\%DS_{QCA}$ – percentage diameter stenosis by QCA; NA – Not Applicable; *Chi-Square Test; significant p-values in bold.

$\%DS_{QCA}$ Strata	$\%DS_{VE}$ - Angiography P Value*	$\%DS_{VE}$ - Segmentation P Value*
All	**0,045**	0,183
> 70%	NA	0,521
50 – 70%	**< 0,001**	0,159
< 50%	0,109	**< 0,001**

4. Discussion

4.1. Main findings

When considering QCA as reference, operators generally tended to overestimate lesion severity in angiography images, but much less so in segmented images. Indeed, the overall rate of agreement in severity strata was approximately double in the segmentation group.

For lesions with a $\%DS_{QCA} > 70\%$, visual estimation was usually in agreement with QCA in both the fluoroscopy and segmented datasets. These were lesions operators deemed as very severe, as is evident by mean $\%DS_{VE} > 80 – 90\%$ in both groups. However, even in this stratum, absolute $\%DS_{VE}$ values were less discrepant with $\%DS_{QCA}$ in the segmentation group.

When considering lesions with $\%DS_{QCA}$ of 51-69%, overestimation was very frequent and pronounced in the angiography group, but not in the segmentation group. Indeed, the mean $\%DS_{VE}$ difference in the angiography group exceeded that of QCA by approximately 15 to 30 percentage points vs 1 to 7 percentage points in the segmentation group.

For the least severe lesions, the findings are somewhat similar to those in the intermediate group, albeit to a lesser degree in all respects: the discrepancy between %DS is not as large and the rates of agreement not as low, in the angiography group. In the segmentation group, differences in %DS are small and overall not significant, with much higher rates of agreement with the $\%DS_{QCA}$ lesion severity strata.

The scatterplots (Figure 29, Supplementary figure 14, Supplementary figure 15) clearly illustrate these findings, as the discrepancy outside $\%DS_{QCA} > 70\%$ stratum is much more evident in the angiography group than the segmentation group. Additionally, when values are plotted based on %DS ratios, it is clear that the trendline for angiography only closely matches that of QCA in the $\%DS_{QCA} > 70\%$ stratum, whereas in the segmentation group it is constantly much closer to 1 (Figure 30, Supplementary figure 16, Supplementary figure 17).

Significant differences across operators were found for both the angiography and segmentation groups. However, the absolute differences in %DS were much lower in the latter, with overall rates of agreement across $\%DS_{QCA}$ severity strata not significantly different in the segmentation group.

The results of our study therefore suggest that visualization of segmented images seems to render visual estimation of stenosis severity more objective, significantly reducing the tendency to overestimate, while possibly reducing operator heterogeneity as well (Figure 31).

The reason for these findings is not entirely clear. The most likely explanation is that segmented images display a stenosed segment in homogeneous fashion, with the transition between artery (white) and background (black) very clearly visible. In contrast, angiography images display the artery in shades of grey, with stenosed regions with poorer contrast filling and less clear demarcation of artery and background. As a result, the human eye seems more prone to underestimating the actual size of the artery lumen in these segments, thereby estimating the stenosis as more severe.

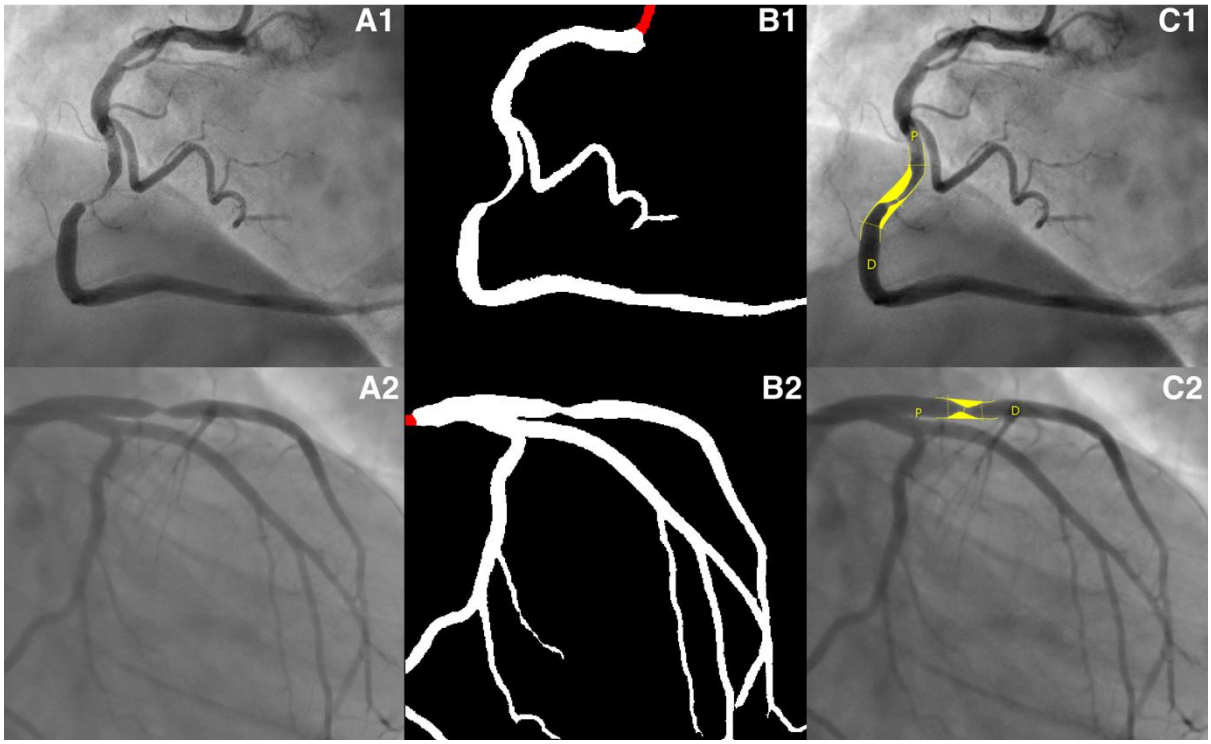


Figure 31: Image examples. The upper three images represent a lesion with $\%DS_{QCA}$ of 78% (C1), with average operator $\%DS_{VE}$ of 90% based on angiography (A1) vs 80% based on segmentation (B1). The lower three images represent a lesion with $\%DS_{QCA}$ of 63% (C2), with average operator $\%DS_{VE}$ of 88% based on angiography (A2) vs 69% based on segmentation (B2). $\%DS_{QCA}$: percentage diameter stenosis by QCA (Quantitative Coronary Analysis).

4.2. Other studies in the literature

The above-mentioned findings for stenosis severity assessment in fluoroscopy images have long been described in the literature for more than 30 years[87] and continue to be found today across the world, even in very large cohorts[88,89,252]. The idea that visual estimation is inaccurate was one of the triggers that spawned the development of QCA analysis and software. Indeed, core-lab QCA in PCI trials continues to be advocated for in international research consortiums[106] or scientific societies[107]. However, in clinical practice, the estimation of stenosis severity continues to be undertaken mostly by visual estimation[252]. The fact that correlation between QCA and FFR has not been demonstrated to be superior to that of visual estimation[89] or that clinical trials often don't require QCA measurements for including patients[57,101], further reinforces this approach.

However, discrepancies between QCA and visual estimation are likely to be of clinical relevance. In a subanalysis of the PROMISE trial of patients who underwent CAG, patients without obstructive disease by QCA, but classified as such by visual estimation, had lower event rates than those with obstructive disease by both criteria[119]. As operators perceive these patients as high risk, revascularization – sometimes potentially unwarranted – may become more likely. And, indeed, the decision to proceed with physiology, which has been shown to be superior to angiography-guided revascularization[93] and is recommended in current guidelines[34,103], implies that the lesion be classified as intermediate rather than severe. Thus, the persistence of the visual estimation approach and the resulting overestimation tendency may arguably be contributing to the seemingly low usage rates (6-13%) of physiology in cath labs[34,182,280]. Lastly, the growing relevance of imaging for decision-making in the cath lab further adds to this issue, as recently highlighted in a major clinical trial[281], further points to the importance of avoiding the so-called "oculostenotic reflex".

4.3. Practical Clinical Implications

In keeping with all of the above, the findings of our study may have an important clinical application: if visual estimation of segmented images makes operators less prone to overestimating lesions, unwarranted revascularization may become less likely and the use of physiology/ischemia/imaging testing may increase. The advantage of AI-based segmentation is that it is fast, fully automatic and requires no human input, other than the image itself. The operator can be simply and immediately be exposed to the segmented image almost effortlessly. Conversely, QCA software is semi-automatic, as it requires the manual annotation

of the region of interest and may require manual adjusting of vessel contours. The simple fact that it is seldom employed in clinical practice further emphasizes the point.

If these exploratory results are confirmed in subsequent studies, AI-based segmentation of the coronary tree may become a relevant tool capable of improving CAG interpretation, hopefully contributing to improved outcomes, without complicating or lengthening the procedure.

4.4. Other studies in AI applied to CAG

Few medical papers regarding the application of AI to CAG have been published in medical literature so far. Two studies focused on segmentation alone, with accurate results[248,249]. We too have previously published our results regarding the development of AI models capable of highly accurate segmentation, encompassing more than 500 images as a whole[273,279]. The largest published study to date focused on developing models capable of segmentation with high accuracy as well, with the added feature of lesion type recognition (i.e. calcium, thrombus, among others), thus potentially enabling future clinical application[247]. Lastly, a small exploratory study also tested the hypothesis of estimating FFR from CAG images using AI, but the study population was small and the group has published no further data in medical journals[255].

Thus, our study is one of the few and first in the field of AI applied to CAG with potential clinical implications.

4.5. Limitations

Our study has important limitations. The dataset size is small and composed predominantly of 51-70% lesions, with a small amount of lesions > 70%, as assessed by QCA. The fact that lesion assessment was based on the evaluation by only four operators is also a limitation. We did not perform sub analysis regarding clinical context, risk factors or lesion characteristics (such as presence of calcium), due to the small sample size. The concept of the study is thus exploratory in nature, requiring external validation in further works.

4.6. Future Directions

We are currently working not only in improving the segmentation capabilities of our model, but also exploring other applications of AI to CAG. Automatic anatomical interpretation, lesion severity based on auto-QCA and integration with physiology are areas of

active research. We aim to develop an integrated suite of AI models capable of enhancing CAG interpretation, improving decision-making and hopefully patient outcomes.

5. Conclusions

When considering QCA as reference, the visual estimation of stenosis severity was much less discrepant with automatic AI-segmented images than angiography images, with reduced inter-operator discrepancies as well. Operators were essentially prone to overestimation of lesion severity, especially of intermediate lesions (%DS_{QCA} of 50-70%). The visual assessment of coronary lesions with segmented images may therefore lead to a lower likelihood of unwarranted revascularization, while potentially increasing the use of functional assessment, as recommended by current guidelines.

6. Acknowledgements

MNM was responsible for conceptualization, data/image gathering, processing and analysis and paper drafting. JLS was responsible for technical and AI tasks, data and image processing, model implementation and training. CR, JG and MSO were responsible for data gathering and image analysis. JSM was responsible for image analysis. TR and BS were responsible for data gathering, processing and analysis. ALO was responsible for supervising the work of JLS, having directly taken part in the same tasks. FJP was responsible for supervising the work of MNM, having directly taken part in the same tasks.

All authors revised the paper critically for important intellectual content, gave final approval for its publication and agree to be accountable for all respects of its accuracy and integrity.

Part IV. Coronary Physiology Instantaneous Wave-Free Ratio (iFR) derived from X-ray angiography using Artificial Intelligence Deep Learning Models: a Pilot Study

Miguel Nobre Menezes[1,2*], João Lourenço Silva[3,4], Beatriz Silva[1,2], Rita Marante de Oliveira[5], Tiago Rodrigues[1,2]Arlindo L. Oliveira[3,4], Fausto J. Pinto[1,2]

(1) Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon (CCUL@RISE), Faculdade de Medicina, Universidade de Lisboa, Av Prof. Egas Moniz, 1649-028 Lisboa, Portugal

(2) Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Av Prof. Egas Moniz, 1649-028 Lisboa.

(3) INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1000-049 Lisboa

(4) Neuralshift Inc. Av. Duque d'Ávila 23, 1000-138 Lisboa

(5) Faculdade de Medicina, Universidade de Lisboa, Av Prof. Egas Moniz, 1649-028 Lisboa, Portugal

J Invasive Cardiol 2024;36(3). doi:10.25270/jic/23.00285.

1. Introduction

The use of invasive coronary physiology has been extensively studied and is today clearly recommended in clinical guidelines[34,103]. The most widely studied index is Fractional Flow Reserve (FFR). Three major trials established its use in selecting lesions where revascularization had an additional benefit to medical therapy[57,93,135], alongside large observational data. More recently, another index gained importance – instantaneous free-wave ratio (iFR). It was initially studied using FFR as gold standard, with high accuracy[147,149]. Our own experience showed similar results[143]. iFR achieved non-inferiority as compared to FFR in 2 major outcome trials[151], including over a 5-year follow-up period[153].

Coronary physiology is, however, vastly underused, ranging from 7-13%[182,183] procedures. The risk of iatrogeny, time consumption and cost are likely reasons for this. Thus, calculating a physiological index (either FFR and/or iFR) digitally from coronary angiography (CAG) images is desirable, as it would bypass these limitations and potentially broaden physiology adoption. While this has already been achieved with several different software approaches relying on 3D vessel reconstruction and complex fluid dynamics computational algorithms[197,202,204], some limitations remain.

In medicine, artificial intelligence (AI) has shown great potential, especially so in imaging, as several publications have demonstrated excellent results with regards to ECG[223], echocardiography[224,225], or MRI[235,263].

However, the use of AI regarding physiology estimation derived from CAG has seldom been explored[255,257]. The potential advantages of using AI for this task could be either fully automating the process with minimal user input and/or improve the reliability of current systems, either functioning as standalone approach or an added layer to existing software.

In this pilot study, we aimed to develop fully automated AI models capable of binary iFR lesion classification from CAG images alone, using measured invasive iFR as reference.

2. Methods

2.1. Inclusion criteria

Single center retrospective selection of consecutive patients in a three-year period (2017 to 2019), who had undergone both CAG and invasive physiology assessment with iFR (Philips Volcano System), regardless of clinical context (i.e. both acute and chronic coronary syndrome).

2.2. Exclusion criteria

We excluded cases where any of the following applied:

Imaging criteria:

- Patients with cardiac devices or other sources of potential imaging artifacts overlapping with the coronary tree image
- Poor image quality
- Unsuccessful segmentation with AI models
- Unclear individualization of lesion outline with overlapping vessels

Clinical criteria

- History of coronary bypass surgery (CABG) or valvular intervention (surgical or percutaneous)
- Significant left heart valvular disease (severe aortic stenosis or regurgitation, severe mitral regurgitation, moderate or severe mitral stenosis, moderate valvular disfunction in both aortic and mitral valves)
- Target culprit vessel culprit of acute coronary syndrome (ACS)
- Target vessel non-culprit of ACS in patient with ST segment elevation myocardial infarction (STEMI) within 48 horas of presentation
- Previous transmural myocardial infarction in target-vessel
- Chronic total occlusion not previous treated with PCI in any vessel
- Left ventricular systolic disfunction, defined as an ejection fraction < 50%
- Cardiogenic shock
- Haemodynamic instability
- Left main lesions

All the exclusion criteria were selected because of their potential impact on physiology indexes measurements, which could generate confoundment in the training process of an AI algorithm.

2.3. AI models development

We have previously trained AI models capable of fully automatic AI segmentation of CAG images[270,273,279,282]. We used these models to segment all the CAGs of included patients. Images were then annotated with the target vessel and the location of the pressor sensor wires used to measure iFR, using a single telediastolic frame where the target vessel was outlined. All available projections were annotated for each case.

Preliminary testing for AI models showed that performance was better when a single image, which best outlined the target vessel, was used, rather than using a combination of different projections. As a result, we proceeded to train our models as such. The model was given only the CAG image, both original (greyscale) and manually annotated after automatic segmentation (Figure 32). No further details, clinical or otherwise (such as the target vessel), were provided. The models were trained to binarily classify whether a given target had an iFR $\leq 0,89$ or $> 0,89$, henceforth defined as a positive or negative iFR result, respectively.

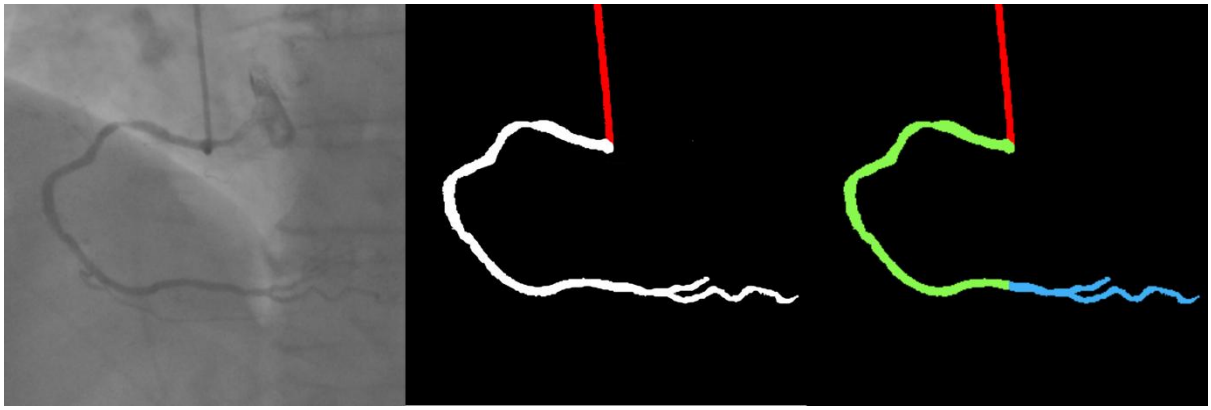


Figure 32: original diagnostic CAG image (left), automatic AI-based segmentation (center) and annotated image with coronary vessel proximal to wire sensor position in green and distal in blue (right).

A total of 3 models were trained. The first model uses as input the sequence of diameters along the main branch of the analyzed vessel, automatically computed by a preprocessing algorithm. This sequence is then processed by a transformer encoder[216], with a classification head on top, which inherently takes into account the sequential nature of the data and allows predicting the iFR value at any given point within the artery. During training, the loss is only

computed at the point for which the ground-truth iFR is available, using a Cross-Entropy Loss, weighted by the inverse of each class' frequency. The Transformer encoder used has 6 layers, 8 heads, a hidden dimensions of 768, GeLU activation, 0,3 dropout and a maximum of 1024 linear positional embeddings, and is followed by a linear classification layer with 0,3 dropout.

Models 2 and 3 are CNNs which take as input the concatenation of the single-channel angiography image and its segmentation, with one channel per class. Model 2 uses a simple Cross-Entropy Loss, and model 3 uses a Cross-Entropy Loss weighted by the inverse of each iFR class' frequency, aimed to mitigate the negative effects of class imbalance. The CNNs use was an EfficientNet-B5[271], with which we had already had success in previous work[270,273,279,282], followed by a linear classification layer.

Theory suggests model 1 should be much more suitable for this task. Unlike it, the CNNs in models 2 and 3 cannot inherently consider the one-dimensional characteristic of the artery and must learn to do so during training, possibly requiring more data to achieve the same level of performance. Additionally, the transformer in model 1 operates on a much lower-dimension input space than CNNs, making the former's task theoretically much easier. Finally, since the location of the iFR predicted by models 2 and 3 is directly tied to the input segmentation, they require an additional inference per iFR prediction.

2.4. Performance assessment and statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. Quantitative variables are shown in average ± standard deviation (if normally distributed) or median (interquartile range) if non-normally distributed. The chi-square test was used for statistically comparing the binary classification of measured iFR vs that of the models. A p-value < 0,05 was used for statistical significance.

The results of the models' classification of target lesions as either iFR positive or negative were compared to those of the true (i.e. real) invasive iFR measurements, as follows:

- True positive (TP): both estimated and real iFR were positive
- False positive (FP): positive estimated iFR and negative real iFR
- True negative (TN): both estimated and real iFR were negative
- False negative (FN): negative estimated iFR and positive real iFR

Using this classification, the following parameters were calculated:

- Accuracy: ([TP+TN]/[TP+TN+FP+FN])
- Sensitivity: TP / (TP + FN)
- Specificity: TN/(TN+FP),
- Positive Predictive Value (PPV): TP / (TP + FP)
- Negative predictive value (NPV): TN/(TN+FN)

To ensure proper evaluation of the results, the models could not be tested on data already seen during training. Hence, we used a cross-validation split, at the patient level, into 10 subsets, retaining the relative distribution of target vessel and iFR classification per split. The models' iFR classification in each subset was then undertaken using neural networks trained exclusively on the remaining data. This enabled the assessment of the models' performance for the whole cohort, whereas the usual splitting in a fixed train/test datasets would have resulted in a much smaller testing group, limiting the ability to test the models' performance. We have successfully used this approach in the past, when developing our segmentation models[270,273,279].

SPSS 27 was used for analysis.

2.5. Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local Ethics' Institutional Review Board.

3. Results

3.1. Baseline characteristics

A total of 334 patients were screened. After the exclusion criteria, a total of 250 measurements, from a total of 223 patients, were included (Table 16, Figure 33). Most lesions had an iFR > 0,89. There was a large imbalance between positive and negative iFR lesions in the right and circumflex coronary arteries subgroups. The difference was much less pronounced in the left anterior descending artery measurements (Table 17).

Table 16: Clinical characteristics of included patients.

Parameter	N +/- SD or N(%)
Age	68 ±11
Sex (male)	148 (66.37%)
Hypertension	180 (80,72%)
Diabetes mellitus	96 (43,05%)
Dyslipidemia	132 (59,19%)
Smoker (past or present)	89 (39,91%)
Chronic coronary syndrome	148 (66,37%)
Acute coronary syndrome	74 (33,18%)

Table 17: iFR results overall and stratified per artery. LAD – Left Anterior Descending artery. RCA – Right Coronary Artery; CX – Circumflex artery. SD – Standard Deviation.

	N (%)	iFR (mean ± SD)	iFR ≤ 0,89 (N / %)	iFR > 0,89 (N / %)
Total	250 (100%)	0,91 ± 0,006	65 (26,0%)	185 (74,0%)
LAD	129,0 (51,6%)	0,88 ± 0,009	55 (42,5%)	74 (57,4%)
RCA	76,0 (30,4%)	0,95 ± 0,009	5 (6,6%)	71 (93,4%)
CX	45,0 (18,0%)	0,96 ± 0,009	5 (33,33%)	40 (88,9%)

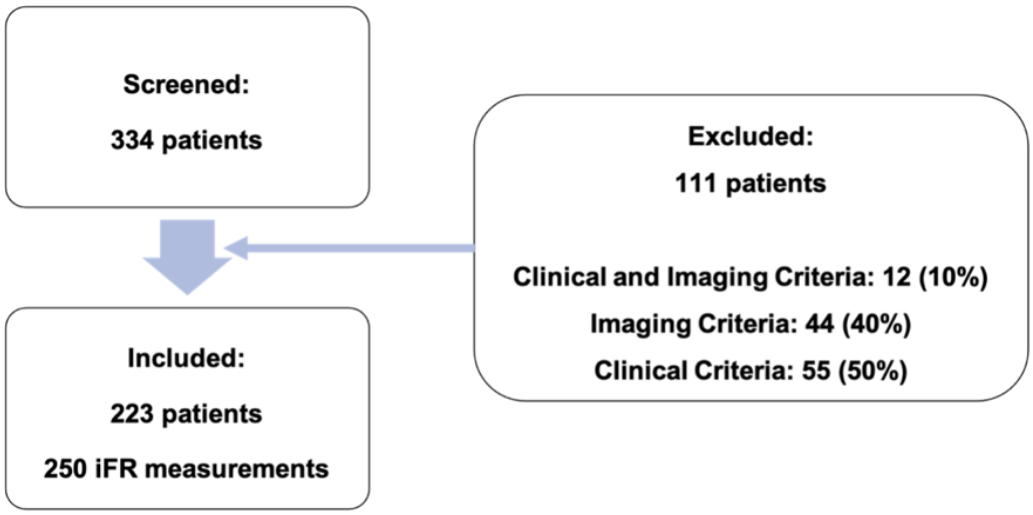


Figure 33: inclusion flowchart.

3.2. Physiology prediction of AI models

3.2.1. Overall Results

The difference between measured iFR classification and that of the AI models was not statistically significant only for model 1 (p = 0,063), whereas models 2 and 3 differed significantly (p < 0,001). Models 1 and 3 classified most lesions as negative, contrary to model 2, with only model 1 distributing classifications in similar proportions to those of the ground truth. Agreement was strongest for negative lesions. Details in Table 18.

All models performed with modest accuracy, with the worst results for model 2 and the best for model 3 (nearing 70%), followed closely by model 1. NPV was high for all models, performing close to 80 or 90%, in contrast with PPV. Models 2 and 3 had the highest sensitivity, whereas model 1 had the highest specificity. Details in Figure 34.

Table 18: iFR classification of lesions as measured vs as per each model prediction for all cases.

Classification (All Cases)		Model 1		Model 2		Model 3		Total
		≤ 0,89	> 0,89	≤ 0,89	> 0,89	≤ 0,89	> 0,89	
Measured iFR	≤ 0,89	24	41	53	12	48	17	65
	> 0,89	46	139	92	93	61	124	185
Total		70	180	145	105	109	141	250

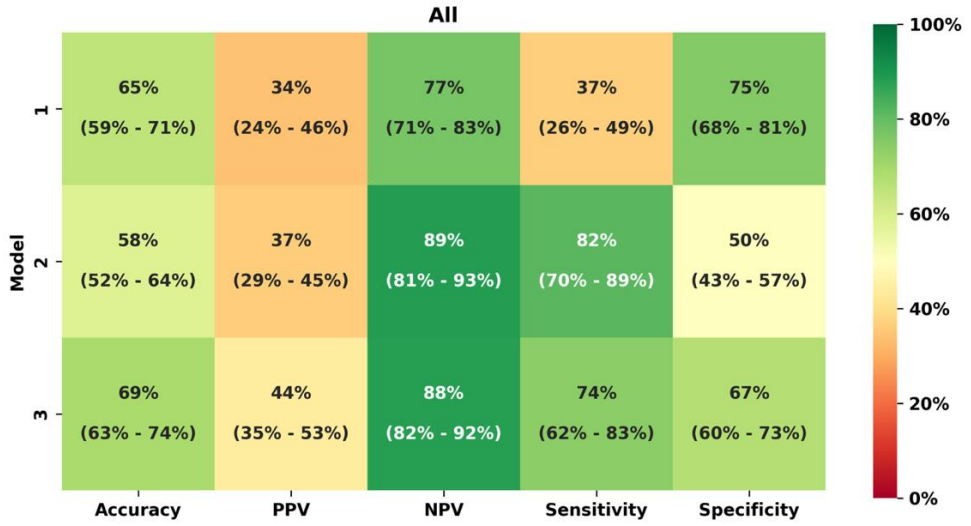


Figure 34: performance metrics (with 95% confidence interval) of each AI model regarding iFR binary classification using the measured iFR classification as reference, for all iFR measurements across all target vessels. Accuracy, Positive Predictive Value (PPV), Negative Predictive Value (NPV), Sensitivity and Specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values.

3.2.2. Left anterior descending (LAD) lesions

The models' classification of iFR as compared to the measured iFR was not statistically significant only for model 1 (p = 0,854), whereas models 2 and 3 differed significantly (p < 0,001). Models 1 and 2 classified most lesions as negative, where agreement was more common. Model 3 classified most lesions as positive. Details in Table 19.

Models 2 and 3 displayed modest accuracy, with better results for the latter, while model 1's performance barely surpassed 50%. The NPV and sensitivity was high or very high for model 2, nearing 90%, while model 3's performance was 78% for both. Model 1 only performed reasonably for specificity (68%). The PPV was superior to 50% for models 2 and 3. Details in Figure 35.

Table 19: iFR classification of lesions as measured vs as per each model prediction for left anterior descending (LAD) cases.

Classification (LAD)		Model 1		Model 2		Model 3		Total
		≤ 0,89	> 0,89	≤ 0,89	> 0,89	≤ 0,89	> 0,89	
Measured iFR	≤ 0,89	17	38	50	5	43	12	55
	> 0,89	24	50	43	31	32	42	74
Total		41	88	93	36	75	54	129

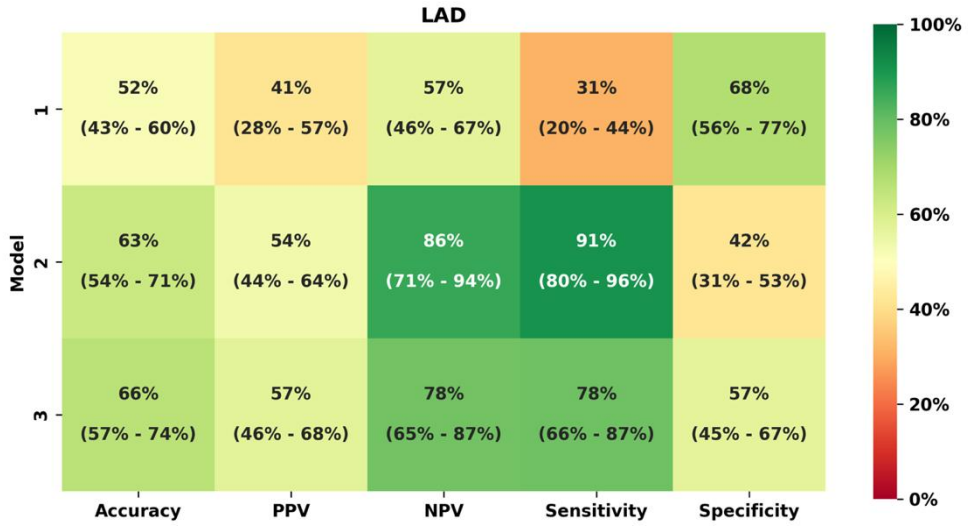


Figure 35: performance metrics (with 95% confidence interval) of each AI model regarding iFR binary classification prediction using the measured iFR classification as reference, for iFR measurements in the Left Anterior Descending Artery (LAD). Accuracy, Positive Predictive Value (PPV), Negative Predictive Value (NPV), Sensitivity and Specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values.

3.2.3. Right coronary artery (RCA) lesions

The iFR classification of model 1 differed significantly from the measured iFR (p = 0,005), whereas for models 2 and 3 there were no significant differences (p = 0,282 and 0,357, respectively). All models classified most lesions as negative (albeit in smaller proportion to the actual measurements distribution), especially model 1. Agreement was highest for negative lesions in all models. Details in Table 20.

Models 1 and 3 displayed the highest accuracy, especially the former (86%), contrasting with model 2. The NPV was always very high, with a maximum of 97% for model 1. On the opposite spectrum of performance, the PPV was very low for all models. Sensitivity was high for both models 1 and 3. Lastly, specificity was low for models 2 and 3, and modest for model 1. Details in Figure 36.

Table 20: iFR classification of lesions as measured vs as per each model prediction for right coronary artery (RCA) cases.

Classification (RCA)		Model 1		Model 2		Model 3		Total
		≤ 0,89	> 0,89	≤ 0,89	> 0,89	≤ 0,89	> 0,89	
Measured iFR	≤ 0,89	3	2	1	4	2	3	5
	> 0,89	9	62	29	42	14	57	71
Total		12	64	30	46	16	60	76

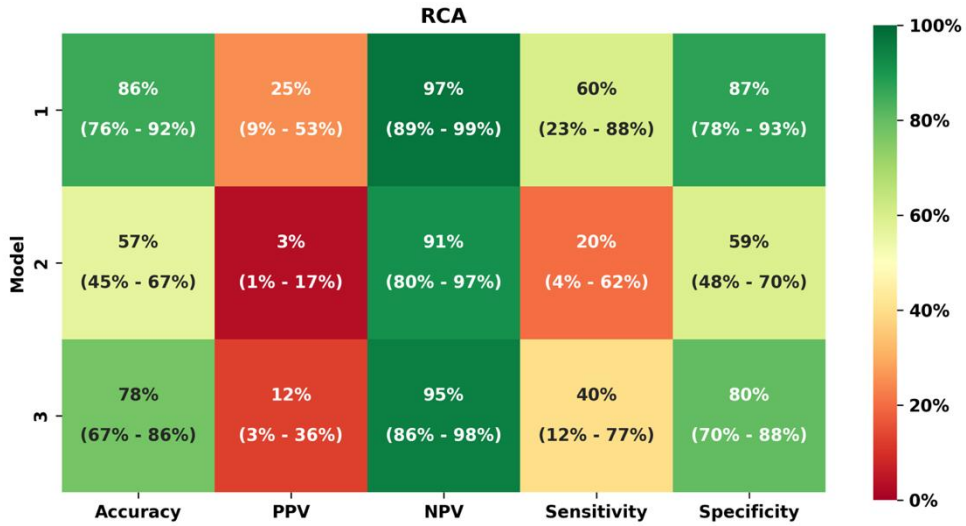


Figure 36: performance metrics (with 95% confidence interval) of each AI model regarding iFR binary classification prediction using the measured iFR classification as reference, for iFR measurements in the Right Coronary Artery (RCA). Accuracy, Positive Predictive Value (PPV), Negative Predictive Value (NPV), Sensitivity and Specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values.

3.2.4. Circumflex artery (CX) lesions

The Measured iFR classification differed significantly from the iFR classification of Model 1 (p = 0,039), but not models 2 and 3 (p < 0,673 and 0,333, respectively). All models classified most lesions as negative, but always in quite shorter proportion to the actual measurements' distribution. Agreement was highest for negative lesions in all models. Details in Table 21.

Model 1 scored higher in all metrics, with an accuracy of 69% a very high NPV (96%) and high specificity (80%). Sensitivity was 68% and the PPV was low (24%). Model 3 followed with similar, albeit inferior performance. Model 2 only scored high in NPV, at 87% Details in Figure 37.

Table 21: iFR classification of lesions as measured vs as per each model prediction for circumflex artery (Cx) cases.

Classification (Cx)		Model 1		Model 2		Model 3		Total
		≤ 0,89	> 0,89	≤ 0,89	> 0,89	≤ 0,89	> 0,89	
Measured iFR	≤ 0,89	4	1	2	3	3	2	5
	> 0,89	13	27	20	20	15	25	40
Total		17	28	22	23	18	27	45

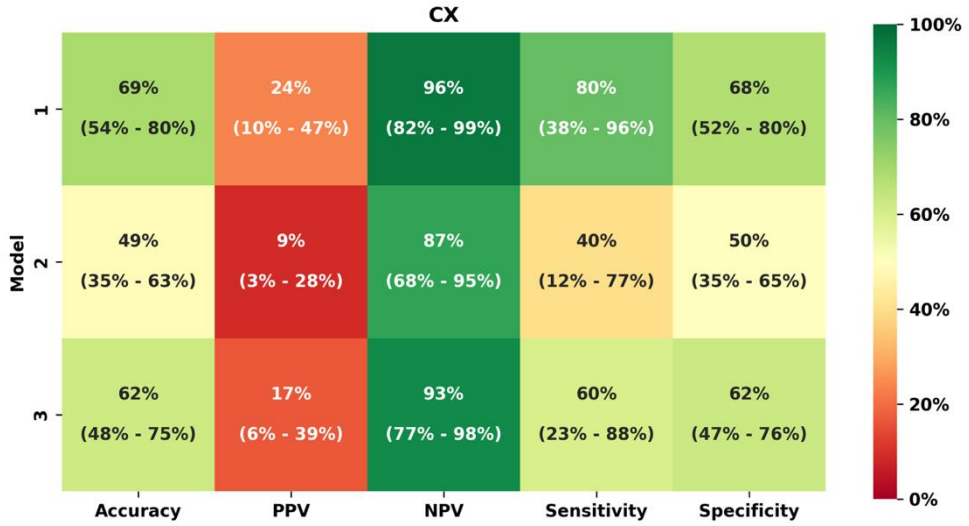


Figure 37: performance metrics (with 95% confidence interval) of each AI model regarding iFR binary classification prediction using the measured iFR classification as reference, for iFR measurements in the Circumflex Artery (Cx). Accuracy, Positive Predictive Value (PPV), Negative Predictive Value (NPV), Sensitivity and Specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values.

4. Discussion

4.1. Main findings – a proof of concept

In this study, we were able to develop AI models capable of binary iFR lesion classification. The accuracy of all models was modest, close to 70% for model 3. For both the LAD and CX, the best that any model could achieve was also just below 70%. However, for the RCA, an accuracy of 78% and 86% was achieved for models 3 and 1, respectively, which, in the latter case, is quite high.

The models displayed very different reliability in correctly classifying lesions as either positive (i.e. iFR ≤ 0,89) or negative (i.e. iFR > 0,89). Indeed, the PPV was low or very low for all models, likely because much fewer positive cases were available for training. Only for the LAD, where there was a larger number of positive lesions, were the models able to achieve modest performance, surpassing the 50% PPV mark. The NPV, however, was generally high or very high, ranging from 77 to 89% overall. For both the RCA and the CX, at least one model neared 100%.

At first glance, one might interpret these findings as a result because of the predominance of negative iFR cases, overwhelmingly so for the RCA and the CX. Therefore, the models could simply be producing a result based on the statistically higher likelihood of a negative result, rather than extrapolating from CAG images.

Notwithstanding, we believe several factors suggests otherwise. First, the models did not necessarily reproduce the distribution of positive/negative lesions of the true measurements, neither globally nor regarding specific vessels. For example, model 3, which achieved the highest accuracy overall, only classified lesions as negative in 56% of cases (vs 74% for real measurements). And the difference was even greater for the CX, where the models' classified lesions as negative ranged from 51% to 62% (vs a real result of 89%).

Furthermore, the models' NPV was also greater that the proportion of negative cases, reaching 86% for the LAD (vs 54% real negative cases) and very close to 100% for both the RCA and Cx (where the real proportion of negative cases was close to 90%). Thus, the number of iFR-negative cases provided enough training data for the models to correctly learn to classify a lesion as negative with high reliability.

Despite this, a significant number of true negative lesions were not identified, as evidence by modest specificity overall, except for models 1 and 3 in the case of RCA lesions, where higher performances were obtained.

No single model proved to be ideal. Model 1, followed by model 3, seemed to be the best option for the RCA and the Cx, whereas for the LAD both models 2 and 3 were a better fit. This suggests that finetuning a model for the target vessel may improve results.

In light of all of the above, we believe our models offer proof of concept that deep learning AI models derivation of coronary physiology data based on X-ray angiography alone is feasible.

4.2. Potential practical clinical implications

When faced with an either positive or negative result by any such software, the operator's main question is often whether the result is likely correct i.e how high is the PPV and/or NPV. Arguably, in the context of invasive coronary physiology, the NPV is of particular importance: a negative result enables the operator to conclude the procedure without engaging in further invasive manoeuvres (i.e deploying a guide catheter, wiring the target vessel, administering further drugs), as opposed to a positive result. The fact that the majority of measurements in invasive physiology are negative further strengthens this point[143,150,151,197,202,203,255]. For example, in this case series, 185 (74%) cases had an iFR > 0,89. Since the models classified lesions as negative in 42% to 72% of cases and the NPV approaches 89%, in the best-case scenario, arguably around one to two thirds of all measurements could have been avoided. For the RCA and CX, where the NPVs neared 100% and the proportion of cases classified as negative was even higher, the impact would have been even more significant.

Thus, while the models are by no means reliable enough for current clinical deployment, this somewhat simplified analysis illustrates the potential practical implications of this technology.

4.3. Analysing lesions on coronary angiography using AI – other studies

Two studies of FFR estimation from CAG using primarily AI methods were published. However, to the best of our knowledge, ours is the first published study to report fully automatic derivation of iFR from CAG using such methods.

Research Projects

Roguin *et al*[255] conducted a pilot feasibility study in a single center population consisting of 31 patients with predominantly LAD lesions (80%). They reported an accuracy of 90%, a NPV of 87%, a PPV of 94%, sensitivity of 88% and specificity of 93% when conducting a binary analysis of FFR ≤ 0,80, similarly to our approach. Their single model is able to derive an estimated FFR value, with an area under the curve of 0,91 and an r correlation coefficient of 0,71 ($p < 0,001$). The study does not report the exact AI methods. However, the binary approach to classify lesions, fully based on AI, using routine angiography projections (rather than predefined angulations), with a fully automatic method, is conceptually very similar to ours. While these results are impressive, with a seemingly superior performance compared to ours, the small sample size and single center nature of the study are limiting factors.

Cho *et al*[257] used a different approach. With a very large sample of 1501 lesions from a single center (predominantly LAD – 67%), the authors plotted the target vessel diameters together with clinical characteristics (age, sex, body surface area, and target segment) to binarily classify FFR measurements with a ≤ 0,80 threshold. An overall accuracy of 82%, a NPV of 84%, a PPV of 81%, sensitivity of 84% and specificity of 80% were reported in the test set, similarly to the external validation dataset of 79 patients. The main limitation of this study, from an AI research perspective, is that it does not provide a fully automated AI physiology estimation, as it requires manual segmentation and diameter calculations of the vessel with an external software, thereby rendering the process semi-automatic and somewhat time-consuming.

Another group is currently launching an initiative for fully automated AI-based PCI guidance interpretation, including FFR[258], but the project is underway and results are not yet available.

Lastly, one of the largest groups in iFR research has applied AI for interpreting iFR pullback curves and found non-inferiority to human performance[283]. While this is a very different task from the one we explored, it highlights how AI is also being applied to this important index which is so commonly used in clinical practice.

There has also been some exploration of the estimation of Coronary Flow Reserve (CFR) from CAG using AI methods, mainly to perform microcirculatory studies[284]. Other authors have tested the application of deep learning methods to classifying stenosis using Quantitative Coronary Analysis (QCA)[252] or automatically detecting significant coronary stenosis using bounding boxes[253,254], which could be of use given the operators' heterogeneity and tendency

for lesion overestimation when interpreting CAG, as we have recently demonstrated ourselves[282].

4.4. Physiology derivation from CAG images without a primary AI approach – other studies

The estimation of physiology from CAG images has been explored in recent years, with commercial software made recently available.

Most studies focused on FFR, using a ≤ 0,80 threshold. The FAST-FFR pivotal trial was a multicentre international study of 301 subjects with a predominance of LAD lesions (54,2%%). The correlation between estimated (FFRangio) and measured FFR was r = 0,80 (p < 0,001), with an accuracy of 92,2%, a NPV of 94,8%, a PPV of 89%, sensitivity of 93,5% and specificity of 91,2%[202]. A pooled analysis of 5 cohort studies yielded similar results[201].

The FAST I[203] and II[204] studies tested a similar approach. In the larger FAST II study[204] 334 patients from six centers were enrolled (with 66% LAD lesions). The correlation with invasive FFR was r = 0,74 (p < 0,001). Two clinical outcomes multicentre trials are ongoing to test the virtual (vFFR) vs invasive FFR approach[205,206].

The Quantitative Flow Ratio (QFR) is perhaps the most extensively "virtual" FFR index studied to date. After the encouraging results of the multicentre FAVOR pilot study (73 patients, 84 vessels, mostly LAD – 54,8%)[118], two further studies enrolled over 600 patients from Europe and Asia[197,198]. In the one with a large European cohort, values of 86,8%, 93%, 76,3%, 86,5% and 86,9% for accuracy, NPV, PPV, sensitivity and specificity were obtained, with a correlation of r = 0,83, p < 0,001[197]. A Chinese multicentric trial comparing a PCI QFR-guided strategy vs an invasive FFR-guided strategy yielded better outcomes for the QFR group[100].

The derivation of iFR from CAG has also been explored on the REVEAL iFR trial[208]. Published results are expected soon.

All of the above studies employ primarily non-AI methods to derive FFR from CAG, using a combination of 3D image reconstruction and computational fluid dynamics. They demonstrate that physiology can successfully be derived from CAG images alone and have a meaningful impact on clinical outcomes. Further ongoing research is likely to strengthen this

approach. While all of the above post-pilot studies yielded better performance results than our model, one thing seems ubiquitous: the NPV is always high, which is arguably, as mentioned above, of particular significance.

However, these approaches are not without disadvantages. A reasonable amount of manual input is necessary – marking the target vessel, proximal and distal points, defining "healthy" regions or correcting for imperfections in manual segmentation - as we have experienced ourselves when testing these platforms at our own catheterization laboratory. They are therefore semi-automatic, somewhat time consuming and may potentially not be as reliable in less experienced hands: the FAST II trial clearly illustrated this, as the performance at specific sites was lower than at the core lab – the authors reported an overall accuracy, NPV, PPV, specificity and sensitivity were 83% vs 90%, 85% vs 90%, 79% vs 90%, 71% vs 81% and 89% vs 95%, respectively[204].

Lastly, except for one[208], the above-mentioned methods require more than one projection (sometimes prespecified), which is probably the result of multiple factors. Indeed, the 3D nature of the coronary anatomy, the existence of energy losses in very distal segments (with resulting lower pressures), along with the limitations of CAG resolution and motion artifacts, all render a 3D approach more reliable, and are likely playing a role in hampering our models' accuracy, since it was based on a single two-dimensional (2D) frame. Thus, while our simplified approach may initially be perceived as advantageous, it was, in all likelihood, a limitation.

4.5. Limitations and future directions

As AI training is highly volume dependent, the dataset size was the most important limitation. This was especially relevant for cases with iFR ≤ 0,89 and cases pertaining to the CX. However, our relative distribution regarding target vessel and positive/negative cases is in agreement with previously published data[143,150,151,197,202,203,255]. As a result, obtaining a dataset with enough iFR positive cases for successful training, especially concerning the RCA and CX, will require a much larger dataset. The dataset size was also limiting with regards to the method employed for model testing. Using a classical train/test split of 80/20% would have resulted in a small testing dataset, reducing our ability to test the models' performance, particularly for subanalyzing performance per target vessel, given the naturally unbalanced characteristics of the dataset. We therefore preferred a cross-validation split in 10 subsets, as mentioned in the

methods section. While this is a common approach in the field of machine learning, it may be regarded as a limitation as well.

Some may view the extrapolation of iFR rather than FFR as a limitation, because FFR was directly compared to angiography in clinical outcomes trials[57,93,135], whereas iFR was only directly compared to FFR[150,151]. However, iFR has repeatedly been shown to be non-inferior to FFR and is today (together with other resting indexes), the default tool of epicardial physiology assessment in many labs due to its simplicity. As a result, the current amount of iFR measurements far outpaces those of FFR in our lab, and thus provides a much larger base for future training, improvement and validation.

Another limitation is the fact the model provides a binary classification, but not yet the iFR value itself. During preliminary testing, it was clear that determining the exact iFR value would require a much larger dataset, which was beyond the scope of a pivotal study.

The use of a single image (human optimized and labelled end-diastolic 2D frame) for the models' training, instead of a 3D reconstruction based on multiple 2D projections, was also a limitation, given the truly 3D nature of the coronary anatomy and lesions. The fact that the above-mentioned non-AI approaches obtained a superior performance with 3D reconstruction, rather than a 2D approach like ours, further supports this consideration.

The single center retrospective dataset is another limitation, as external validation will be required in the future. Notwithstanding, physiology results have been shown to be quite reproducible and thus the impact of this particular limitation is less likely to be of significance.

The concept of this study is thus exploratory and aimed at proof of concept. We aim to greatly enlarge the training dataset through multi-institutional collaboration (as we have done for our segmentation models[273,279]) which will be essential for performance improvements and external validation. We also aim to further improve our models using multiple projections and 3D reconstruction, which may likely enhance performance. The ultimate aim of deploying them in clinical practice – this may occur either as standalone solution, or by improving existing software.

5. Conclusion

We developed deep learning AI models capable of binary lesion classification using an iFR threshold of 0,89. While the overall accuracy of the models is not yet high enough for clinical deployment, the high negative predictive capacity of these models is of clinical significance and potential clinical application. This pivotal study therefore offers proof of concept for further development, with larger and multicentric training and validation datasets poised for the future. This approach has the potential to evolve to a standalone software aid in the cath lab, or further integration in existing, non-AI based software solutions, by streamlining workflows and/or improving their performance. This could prove to be of great value for patient management and to improve cath lab flows.

6. Acknowledgements

MNM was responsible for conceptualization, data/image gathering, processing and analysis and paper drafting. JLS was responsible for technical and AI tasks, data and image processing, model implementation and training. TR, BS and RMO were responsible for data gathering, processing and analysis. ALO was responsible for supervising the work of JLS, having directly taken part in the same tasks. FJP was responsible for supervising the work of MNM, having directly taken part in the same tasks.

All authors revised the paper critically for important intellectual content, gave final approval for its publication and agree to be accountable for all respects of its accuracy and integrity.

7. Data Availability

Detailed full-scale study data cannot currently be made publicly available due to limitations imposed by national data protection regulations, as this is a retrospective study and no informed consent was obtainable regarding this particular analysis. Both our research team and others in the national scientific community are working to develop a framework where such would be possible. However, independent replication of our segmentation models is possible, given that the detailed description of our experimentations and relevant code is publicly available[270,273,279,282].

CHAPTER 4: OVERALL DISCUSSION AND CONCLUSION

1. Overall Discussion

We successfully developed an AI model capable of accurate CAG segmentation and validated it with data from four different centers, across a wide range of stenosis severity, target vessel, operators and equipment.

Importantly, the assessment of the quality of these models was extensively analyzed by three complementary methods, which, to our knowledge, has never been done before.

First, standard metrics of overlap comparison were carried out, using the manual baseline annotations as reference. We took several steps to ensure the quality of these was as high as possible: a small group of annotators was selected, thereby minimizing variability, contrasting with large numbers of annotators as in previous works[247]; the annotators were cardiology physicians and therefore utmostly qualified in the interpretation of CAG images, whereas other groups have resorted to medical students[249,250] or have not disclosed the exact qualifications of the annotators[247]; the labelled images were systematically reviewed and corrected to minimize errors every step of the way; furthermore, the dataset used for the best performing model was the result of a combination of an initial AI model plus human annotations, thereby perfecting an already high quality training dataset. We believe all these steps, together with a carefully chosen CNN architecture[270], were paramount in developing a model which exceled in all metrics (Dice Score, Intersection Over Union, Accuracy, Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value), both on its own and when compared to previously published work, despite having been developed from a smaller dataset than most studies[243,244,246–251].

Secondly, we also assessed the quality of the segmentation by comparison with state-of-the art software, while specifically testing in stenosed segments, which is of particular clinical importance, given that these are the lesions operators focus on when interpreting CAG images and making decisions. Therefore, if segmentation as whole seemed appropriate but the models were inaccurate in lesion regions, they could not serve any clinical purpose. Once again, this contrasts with several previous works, where only the baseline annotations were used as reference and/or the metrics were assessed for the whole image [246,247,251].

Third, models may score quite highly in both of the above and still produce segmentation results that are flawed from a clinical perspective. For example, if metrics are globally high, but models were prone to errors in the main vessel segmentation (for example, missing a part

of a proximal LAD), the segmentation would not be fit for any clinical use. By developing a score that acknowledges that the relative weight of certain segmentation tasks is more important than others, we have integrated the eyes of interventionalists into the assessment of our model quality. Furthermore, the external validation of the score with operators from multiple centers, who were able to apply it easily and reproducibly, renders it scientifically ready for development and may be of use to other research groups.

We have shown that the visualization of AI automatic segmentation of CAG images renders its interpretation more reproducible across operators, while significantly reducing the tendency for lesion severity overestimation.

Operator heterogeneity in the visual estimation of lesion severity has been recognized for over 50 years, almost from the very beginning of the availability of CAG [71,72]. An optimistic perspective might consider that improvements in image quality and the appearance of fully digital systems might attenuate or abolish this tendency. Furthermore, one might also consider that generational differences across operators might be of relevance, as the use of physiology and intra-coronary imaging are quite popular among young operators, thus potentially rendering them more aware of their tendency for overestimation. The fact that multiple studies spanning 5 decades[85–90] have shown that these limitations persist suggest that DS visual estimation of CAG images is inherently prone to error and overestimation. By conducting such an analysis with four relatively young operators (with experiences in CAG image interpretation ranging from 8 to 15 years and ages around 40-years old) from four different centers with the latest generation imaging technology, we have once again confirmed this finding.

When assessing the same metric (DS by visual estimation) as compared with an objective measurement with validated software (QCA), the operators' assessment of DS was (1) very close to that of QCA, (2) with little inter-operator heterogeneity and (3) with a much smaller tendency for overestimation. Indeed, truly severe lesions continued to be classified as such, whereas intermediate lesions (and even some "truly" mild lesions) were often no longer estimated as severe. This is likely the result of a very clear demarcation of the transition between artery (white) and background (black) of segmented images, which when viewed in shades of grey in the original images can be quite difficult to appreciate, especially in stenosed regions.

These findings therefore highlight that automatic digital mapping systems of CAG images are likely warranted, as they render their interpretation more objective. Our immediate

and fully automatic segmentation model, which requires no input other than the image itself, therefore has the potential of reducing unwarranted revascularization, while potentially increasing the use of physiology and/or imaging and hopefully have a meaningful impact on clinical outcomes.

We successfully developed AI models capable of binary iFR lesion classification using a single bidimensional frame image. Despite the modest accuracy of all models, the high negative predictive value observed is of clinical relevance, as a reliable negative result effectively enables the operator to terminate the procedure without proceeding with further testing, either invasive or non-invasive. The fact that 60 to 70% of measurements in physiology are usually negative reinforces the relevance of this finding[143,150,151,197,202,203,255]. In the case-series herein explored, such models could have avoided invasive measurements in one to two thirds of cases, streamlining the procedure while reducing costs, as well as risk for the patient, considering that invasive physiology requires the use of guide-catheters, maneuvering guidewires into the target vessel and the administration of further drugs (nitrates and anticoagulation).

Another relevant issue is the fact that, to our knowledge, the derivation of iFR from CAG images has seldom been explored. Only a single study has addressed this issue, but while its rationale has been published[208], actual results have not. Indeed, currently published studies on CAG-based derivation of physiology have always focused on FFR[201–204], with one clinical trial already showing a positive impact on clinical outcomes[200], further highlighting the relevance of this sort of research. The preference for FFR is likely because this has been the most widely studied index to date and the only to have been directly studied vs angiography and medical therapy in clinical outcomes studies. However, it is well-recognized that no single index is perfect. Furthermore, iFR may be advantageous in tandem lesions, as it possibly better separates the individual contribution of each lesion to the overall ischemic burden, while predicting the physiology impact of PCI in a selected region, thereby potentially enabling more selective revascularization strategies. Moreover, while the usefulness of FFR has been made clear, its exact role of non-culprit lesions after ACS (especially STEMI) has been challenged[199]. If the the results of the ongoing iMODERN trial[160], which is comparing an iFR vs CMR based strategy for non-culprit lesions after STEMI, is not unfavorable to the iFR-arm, its role in clinical practice will become even more relevant. Thus, studying not only the derivation of FFR, but also that of resting indexes, namely iFR, is both scientifically and clinically pertinent.

Overall Discussion and Conclusion

Lastly, despite the successful derivation of FFR from CAG images, the potential role of AI in these systems is underexplored. Indeed, the vast majority of research has used on the derivation of FFR primarily from computational fluid dynamics and 3D-QCA reconstruction. Only two groups have used primarily AI-methods to derive physiology, both with FFR[255–257]. Thus, the fact that we employed solely AI methods, based on a single 2D frame and achieved significant results further highlights the potential role of AI in this field.

Our research has several limitations. Regarding the segmentation model, the dataset size was relatively small when compared to previous studies. Furthermore, the fact that we excluded potential sources of artifacts such as patients with rhythm devices or past cardiac surgery limits the application of our model to these subsets.

With regards to the assessment of the impact of segmentation in operators' assessment of lesions, the most relevant limitation is the relatively small dataset size, where intermediate (QCA DS 50-70%) lesions predominate. The relatively small number of operators and centers was also a limitation.

Lastly, considering the iFR classification models, there are a greater number of limitations. First, the dataset size limited the models' performance. This is especially relevant for iFR positive (i.e. ≤ 0,89) cases, namely for the RCA and Cx. The small availability of such cases severely limits any chance of training models with high positive predictive value. Secondly, the single center retrospective nature of this study means that validation, ideally including external multicentric datasets, is necessary. Another limitation is the fact that our models were trained only for iFR estimation, but not FFR, which may be of particular importance in regions of with very large myocardial mass, such as the proximal LAD, where discordance between both methods has been reported[175]. Despite the fact that the training dataset also comprised a number of FFR measurements (concomitant to iFR, for historical reasons), the increased simplicity of iFR has made it the default index at our catheterization laboratory. Thus, there was, is and will be a much larger base for future training and improvement with iFR than FFR. Lastly, our models are not currently capable of deriving the actual iFR value, but only produce a binary classification, which, in the context of ischemia, a continuous variable, is also a limitation.

2. Future perspectives and concluding remarks

We successfully produced new AI models capable of coronary angiography segmentation of excellent quality, demonstrating that their application can positively impact operators' ability to accurately classify the severity of coronary artery lesions. By additionally integrating physiology derivation directly from coronary images alone, the interpretation of coronary angiography becomes quickly and automatically more reproducible and objective. A proposed workflow for the modern interpretation of coronary angiography with digital tools based in our work depicted in Figure 38.

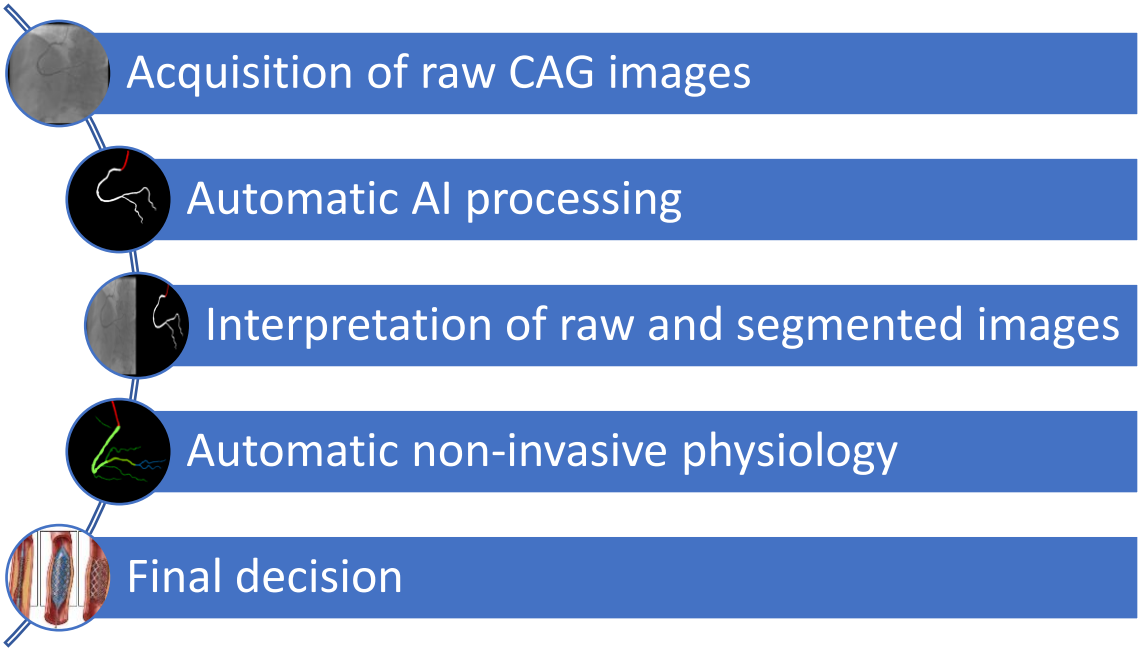


Figure 38: a proposed workflow for the modern interpretation of invasive coronary angiography.

This approach may potentially improve the selection of lesions that warrant revascularization versus those that do not, potentially improving clinical outcomes. However, given the above mentioned-limitations, it is not yet ready for clinical deployment, but the results herein presented offer an important proof-of-concept. In the future, we hope to enlarge the training dataset and include other centers. But whatever the outcome, given the fast pace of research in the field in the last few years, we believe our contribution has already been valuable and may, in the future, be applied either as a standalone solution or by improving existing software.

Overall Discussion and Conclusion

In my opinion, for over fifty years and since its inception, the primary method of asserting the severity of coronary artery disease based on invasive coronary angiography images - visual estimation - has hardly changed. It is time that this persistent conundrum is finally overcome. Digital mapping systems, with the help of AI, may finally bring the paradigm of CAG interpretation into the Digital Era that characterizes the 21st century. The research herein presented has been and will be valuable in this endeavor, and I believe such systems will soon become the norm in everyday catheterization laboratories around the world.

REFERENCES

1. Nowbar AN, Howard JP, Finegold JA, Asaria P, Francis DP. 2014 Global geographic analysis of mortality from ischaemic heart disease by country, age and income: Statistics from World Health Organisation and United Nations. *Int J Cardiol* 2014;**174**:293.
2. Ralapanawa U, Sivakanesan R. Epidemiology and the Magnitude of Coronary Artery Disease and Acute Coronary Syndrome: A Narrative Review. *J Epidemiol Glob Health* 2021;**11**:169.
3. Timmis A, Vardas P, Townsend N, Torbica A, Katus H, Smedt D De, Gale CP, Maggioni AP, Petersen SE, Huculeci R, Kazakiewicz D, Benito Rubio V de, Ignatiuk B, Raisi-Estabragh Z, Pawlak A, Karagiannidis E, Treskes R, Gaita D, Beltrame JF, McConnachie A, Bardinet I, Graham I, Flather M, Elliott P, Mossialos EA, Weidinger F, Achenbach S. European Society of Cardiology: cardiovascular disease statistics 2021. *Eur Heart J* 2022;**43**:716–799.
4. Forssmann W. Die Sondierung des Rechten Herzens. *Klin Wochenschr* 1929;**8**:2085–2087.
5. Seldinger SI. Catheter replacement of the needle in percutaneous arteriography; a new technique. *Acta radiol* 1953;**39**:368–376.
6. M Macedo M. The Portuguese School of Angiography. *Rev Port. Card* 1999;**18**:I-9-I–10.
7. Santos R dos, Lamas A, Pereira Caldas J. Arteriografia da aorta e dos vasos abdominais. *Med Contemp* 1929;**47**:93–97.
8. Macieira Coelho E. Eduardo Coelho, Pioneiro da Cardiologia. *Acta Med Port* 1995;**8**:57–61.
9. Berry D. Pioneers in Cardiology: Eduardo de Araújo Coelho, M.D., PhD. *Eur Heart J* 2010;**31**:760–761.
10. Sones FMJ, Shirey EK. Cine coronary arteriography. *Mod Concepts Cardiovasc Dis* 1962;**41**:735–738.
11. Judkins MP. Selective coronary arteriography. I. A percutaneous transfemoral technic. *Radiology* 1967;**89**:815–824.
12. Radner S. Thoracal aortography by catheterization from the radial artery; preliminary report of a new technique. *Acta radiol* 1948;**29**:178–180.
13. Vineberg A, Miller G. Treatment of Coronary Insufficiency. *Can Med Assoc J* 1951;**64**:204.

14. Goetz RH, Rohman M, Haller JD, Dee R, Rosenak SS. Internal mammary-coronary artery anastomosis: A Nonsuture Method Employing Tantalum Rings. *J Thorac Cardiovasc Surg* 1961;**41**:378–386.
15. Garrett HE, Dennis EW, DeBakey ME. Aortocoronary Bypass With Saphenous Vein Graft: Seven-Year Follow-Up. *JAMA* 1973;**223**:792–794.
16. Kolessov VI. Mammary artery-coronary artery anastomosis as method of treatment for angina pectoris. *J Thorac Cardiovasc Surg* 1967;**54**:535–544.
17. Green GE, Stertzer SH, Reppert EH. Coronary Arterial Bypass Grafts. *Ann Thorac Surg* 1968;**5**:443–450.
18. DOTTER CT, JUDKINS MP. Transluminal treatment of arteriosclerotic obstruction: Description of a new technique and a preliminary report of its application. *Circulation* 1964;**30**:654–670.
19. Zeitler E, Schoop W, Zahnow W. The treatment of occlusive arterial disease by transluminal catheter angioplasty. *Radiology* 1971;**99**:19–26.
20. Grüntzig AR, Senning Å, Siegenthaler WE. Nonoperative dilatation of coronary-artery stenosis: percutaneous transluminal coronary angioplasty. *N Engl J Med* 1979;**301**:61–68.
21. Simpson JB, Baim DS, Robert EW, Harrison DC. A new catheter system for coronary angioplasty. *Am J Cardiol* 1982;**49**:1216–1222.
22. Bonzel T, Wollschläger H, Just H. A new catheter system for the mechanical dilatation of coronary stenoses with exchangeable intracoronary catheters, fast flow of the contrast agent and improved control. *Biomed Tech (Berl)* 1986;**31**:195–200.
23. Rousseau H, Puel J, Joffre F, Sigwart U, Duboucher C, Imbert C, Knight C, Kropf L, Wallsten H. Self-expanding endovascular prosthesis: an experimental study. *Radiology* 1987;**164**:709–714.
24. Palmaz JC, Sibbitt RR, Reuter SR, Tio FO, Rice WJ. Expandable intraluminal graft: a preliminary study. Work in progress. *Radiology* 1985;**156**:73–77.
25. Serruys PW, Jaegere P de, Kiemeneij F, Macaya C, Rutsch W, Heyndrickx G, Emanuelsson H, Marco J, Legrand V, Materne P, Belardi J, Sigwart U, Colombo A, Goy JJ, Heuvel P van den, Delcan J, Morel M. A Comparison of Balloon-Expandable-Stent Implantation with Balloon Angioplasty in Patients with Coronary Artery Disease. *New Eng J Med* 1994;**331**:489–495.
26. Sousa JE, Costa MA, Abizaid AC, Rensing BJ, Abizaid AS, Tanajura LF, Kozuma K, Langenhove G Van, Sousa AGMR, Falotico R, Jaeger J, Popma JJ, Serruys PW.

Sustained Suppression of Neointimal Proliferation by Sirolimus-Eluting Stents. *Circulation* 2001;**104**:2007–2011.

27. Colombo A, Drzewiecki J, Banning A, Grube E, Hauptmann K, Silber S, Dudek D, Fort S, Schiele F, Zmudka K, Guagliumi G, Russell ME. Randomized study to assess the effectiveness of slow- and moderate-release polymer-based paclitaxel-eluting stents for coronary artery lesions. *Circulation* 2003;**108**:788–794.
28. Urban P, Meredith IT, Abizaid A, Pocock SJ, Carrié D, Naber C, Lipiecki J, Richardt G, Iñiguez A, Brunel P, Valdes-Chavarri M, Garot P, Talwar S, Berland J, Abdellaoui M, Eberli F, Oldroyd K, Zambahari R, Gregson J, Greene S, Stoll H-P, Morice M-C. Polymer-free Drug-Coated Coronary Stents in Patients at High Bleeding Risk. *New Eng J Med* 2015;**373**:2038–2047.
29. Valgimigli M, Patialiakas A, Thury A, McFadden E, Colangelo S, Campo G, Tebaldi M, Ungi I, Tondi S, Roffi M, Menozzi A, Cesare N de, Garbo R, Meliga E, Testa L, Gabriel HM, Airoldi F, Ferlini M, Liistro F, Dellavalle A, Vranckx P, Briguori C, ZEUS Investigators. Zotarolimus-Eluting Versus Bare-Metal Stents in Uncertain Drug-Eluting Stent Candidates. *J Am Coll Cardiol* 2015;**65**:805–815.
30. Watanabe H, Domei T, Morimoto T, Natsuaki M, Shiomi H, Toyota T, Ohya M, Suwa S, Takagi K, Nanasato M, Hata Y, Yagi M, Suematsu N, Yokomatsu T, Takamisawa I, Doi M, Noda T, Okayama H, Seino Y, Tada T, Sakamoto H, Hibi K, Abe M, Kawai K, Nakao K, Ando K, Tanabe K, Ikari Y, Hanaoka KI, Morino Y, Kozuma K, Kadota K, Furukawa Y, Nakagawa Y, Kimura T. Effect of 1-Month Dual Antiplatelet Therapy Followed by Clopidogrel vs 12-Month Dual Antiplatelet Therapy on Cardiovascular and Bleeding Events in Patients Receiving PCI. *JAMA* 2019;**321**:2414.
31. Valgimigli M, Frigoli E, Heg D, Tijssen J, Jüni P, Vranckx P, Ozaki Y, Morice M-C, Chevalier B, Onuma Y, Windecker S, Tonino PAL, Roffi M, Lesiak M, Mahfoud F, Bartunek J, Hildick-Smith D, Colombo A, Stanković G, Iñiguez A, Schultz C, Kornowski R, Ong PJL, Alasnag M, Rodriguez AE, Moschovitis A, Laanmets P, Donahue M, Leonardi S, Smits PC. Dual Antiplatelet Therapy after PCI in Patients at High Bleeding Risk. *New Eng J Med* 2021;**385**:1643–1655.
32. Valgimigli M, Wlodarczak A, Tölg R, Merkely B, Kelbæk H, Legutko J, Galli S, Godin M, Toth GG, Lhermusier T, Honton B, Dietrich PL, Stammen F, Ferdinande B, Silvain J, Capodanno D, Cayla G, Bioflow-DAPT Investigators. Biodegradable-Polymer or Durable-Polymer Stents in Patients at High Bleeding Risk: A Randomized, Open-Label Clinical Trial. *Circulation* 2023.

33. Windecker S, Latib A, Kedhi E, Kirtane AJ, Kandzari DE, Mehran R, Price MJ, Abizaid A, Simon DI, Worthley SG, Zaman A, Hudec M, Poliacikova P, Abdul Ghapar AK bin, Selvaraj K, Petrov I, Mylotte D, Pinar E, Moreno R, Fabbiocchi F, Pasupati S, Kim H-S, Aminian A, Tie C, Wlodarczak A, Hur S-H, Marx SO, Jankovic I, Brar S, Bousquette L, Liu M, Stone GW. Polymer-based or Polymer-free Stents in Patients at High Bleeding Risk. *N Engl J Med* 2020;**382**:1208–1218.

34. Neumann FJ, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, Byrne RA, Collet JP, Falk V, Head SJ, Jüni P, Kastrati A, Koller A, Kristensen SD, Niebauer J, Richter DJ, Seferovic PM, Sibbing D, Stefanini GG, Windecker S, Yadav R, Zembala MO, Wijns W, Glineur D, Aboyans V, Achenbach S, Agewall S, Andreotti F, Barbato E, Baumbach A, Brophy J, Bueno H, Calvert PA, Capodanno D, Davierwala PM, Delgado V, Dudek D, Freemantle N, Funck-Brentano C, Gaemperli O, Gielen S, Gilard M, Gorenek B, Haasenritter J, Haude M, Ibanez B, Iung B, Jeppsson A, Katritsis D, Knuuti J, Kolh P, Leite-Moreira A, Lund LH, Maisano F, Mehilli J, Metzler B, Montalescot G, Pagano D, Petronio AS, Piepoli MF, Popescu BA, Sádaba R, Shlyakhto E, Silber S, Simpson IA, Sparv D, Tavilla G, Thiele H, Tousek P, Belle E Van, Vranckx P, Witkowski A, Zamorano JL, Roffi M. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J* 2019;**40**:87–165.

35. Yock PG, Linker DT, Angelsen BAJ, Tech. Two-dimensional intravascular ultrasound: technical development and initial clinical experience. *J Am Soc Echocardiogr* 1989;**2**:296–304.

36. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA, Fujimoto JG. Optical coherence tomography. *Science* 1991;**254**:1178–1181.

37. Gao XF, Ge Z, Kong XQ, Kan J, Han L, Lu S, Tian NL, Lin S, Lu QH, Wang XY, Li QH, Liu ZZ, Chen Y, Qian XS, Wang J, Chai DY, Chen CH, Pan T, Ye F, Zhang JJ, Chen SL. 3-Year Outcomes of the ULTIMATE Trial Comparing Intravascular Ultrasound Versus Angiography-Guided Drug-Eluting Stent Implantation. *JACC Cardiovasc Interv* 2021;**14**:247–257.

38. Hong SJ, Mintz GS, Ahn CM, Kim JS, Kim BK, Ko YG, Kang TS, Kang WC, Kim YH, Hur SH, Hong BK, Choi D, Kwon H, Jang Y, Hong MK. Effect of Intravascular Ultrasound–Guided Drug-Eluting Stent Implantation: 5-Year Follow-Up of the IVUS-XPL Randomized Trial. *JACC Cardiovasc Interv* 2020;**13**:62–71.

39. Lee JM, Choi KH, Song Y Bin, Lee J-Y, Lee S-J, Lee SY, Kim SM, Yun KH, Cho JY, Kim CJ, Ahn H-S, Nam C-W, Yoon H-J, Park YH, Lee WS, Jeong J-O, Song PS, Doh J-H, Jo S-H, Yoon C-H, Kang MG, Koh J-S, Lee KY, Lim Y-H, Cho Y-H, Cho J-M, Jang WJ, Chun K-J, Hong D, Park TK, Yang JH, Choi S-H, Gwon H-C, Hahn J-Y. Intravascular Imaging–Guided or Angiography-Guided Complex PCI. *New Eng J Med* 2023;**388**:1668–1679.
40. Ali ZA, Galougahi KK, Maehara A, Shlofmitz RA, Fabbiocchi F, Guagliumi G, Alfonso F, Akasaka T, Matsumura M, Mintz GS, Ben-Yehuda O, Zhang Z, Rapoza RJ, West NEJ, Stone GW. Outcomes of optical coherence tomography compared with intravascular ultrasound and with angiography to guide coronary stent implantation: one-year results from the ILUMIEN III: OPTIMIZE PCI trial. *EuroIntervention* 2021;**16**:1085–1091.
41. Jones DA, Rathod KS, Koganti S, Hamshere S, Astroulakis Z, Lim P, Sirker A, O'Mahony C, Jain AK, Knight CJ, Dalby MC, Malik IS, Mathur A, Rakhit R, Lockie T, Redwood S, MacCarthy PA, Desilva R, Weerackody R, Wragg A, Smith EJ, Bourantas C V. Angiography Alone Versus Angiography Plus Optical Coherence Tomography to Guide Percutaneous Coronary Intervention: Outcomes From the Pan-London PCI Cohort. *JACC Cardiovasc Interv* 2018;**11**:1313–1321.
42. Holm NR, Andreasen LN, Neghabat O, Laanmets P, Kumsars I, Bennett J, Olsen NT, Odenstedt J, Hoffmann P, Dens J, Chowdhary S, O'Kane P, Bülow Rasmussen S-H, Heigert M, Havndrup O, Kuijk JP Van, Biscaglia S, Mogensen LJH, Henareh L, Burzotta F, H Eek C, Mylotte D, Llinas MS, Koltowski L, Knaapen P, Calic S, Witt N, Santos-Pardo I, Watkins S, Lønborg J, Kristensen AT, Jensen LO, Calais F, Cockburn J, McNeice A, Kajander OA, Heestermans T, Kische S, Eftekhari A, Spratt JC, Christiansen EH, OCTOBER Trial Group. OCT or Angiography Guidance for PCI in Complex Bifurcation Lesions. *N Engl J Med* 2023;**389**:1477-1487;.
43. Kataruka A, Maynard CC, Hira RS, Dean L, Dardas T, Gurm H, Brown J, Ring ME, Doll JA. Government Regulation and Percutaneous Coronary Intervention Volume, Access and Outcomes: Insights From the Washington State Cardiac Care Outcomes Assessment Program. *J Am Heart Assoc* 2022;**11**:25607.
44. Melly L, Torregrossa G, Lee T, Jansens JL, Puskas JD. Fifty years of coronary artery bypass grafting. *J Thorac Dis* 2018;**10**:1960–1967.
45. Barbato E, Noc M, Baumbach A, Dudek D, Bunc M, Skalidis E, Banning A, Legutko J, Witt N, Pan M, Tilsted HH, Nef H, Tarantini G, Kazakiewicz D, Huculeci R, Cook S,

Magdy A, Desmet W, Cayla G, Vinereanu D, Voskuil M, Goktekin O, Vardas P, Timmis A, Haude M. Mapping interventional cardiology in Europe: the European Association of Percutaneous Cardiovascular Interventions (EAPCI) Atlas Project. *Eur Heart J* 2020;**41**:2579–2588.

46. Pereira H, Campante Teles R, Costa M, Canas da Silva P, Gama Ribeiro V da, Brandão V, Martins D, Matias F, Pereira-Machado F, Baptista J, Farto e Abreu P, Santos R, Drummond A, Cyrne de Carvalho H, Calisto J, Silva JC, Pipa JL, Marques J, Sousa P, Fernandes R, Cruz Ferreira R, Ramos S, Oliveira E, Almeida M. Angioplastia primária em Portugal entre 2002-2013. Atividade segundo o Registo Nacional de Cardiologia de Intervenção. *Rev Port. Card* 2016;**35**:395–404.
47. Cannon CP, Weintraub WS, Demopoulos LA, Vicari R, Fery MG, Lakkis N, Neumann FJ, Robertson DH, DeLuca PT, DiBattiste PM, Gibson CM, Braunwald E. Comparison of Early Invasive and Conservative Strategies in Patients with Unstable Coronary Syndromes Treated with the Glycoprotein IIb/IIIa Inhibitor Tirofiban. *New Eng J Med* 2001;**2**:1127–1129.
48. Kofoed KF, Kelbæk H, Riis Hansen P, Torp-Pedersen C, Høfsten D, Kløvgaard L, Holmvang L, Helqvist S, Jørgensen E, Galatius S, Pedersen F, Bang L, Saunamaki K, Clemmensen P, Linde JJ, Heitmann M, Wendelboe Nielsen O, Raymond IE, Peter Kristiansen O, Hastrup Svendsen I, Bech J, Dominguez Vall-Lamora MH, Kragelund C, Fritz Hansen T, Dahlgaard Hove J, Jørgensen T, Fornitz GG, Steffensen R, Jurlander B, Abdulla J, Lyngbæk S, Elming H, Krohn Therkelsen S, Abildgaard U, Skov Jensen J, Gislason G, Køber L V., Engstrøm T. Early Versus Standard Care Invasive Examination and Treatment of Patients With Non-ST-Segment Elevation Acute Coronary Syndrome. *Circulation* 2018;**138**:2741–2750.
49. Mehta SR, Granger CB, Boden WE, Steg PG, Bassand J-P, Faxon DP, Afzal R, Chrolavicius S, Jolly SS, Widimsky P, Avezum A, Rupprecht H-J, Zhu J, Col J, Natarajan MK, Horsman C, Fox KAA, Yusuf S. Early versus delayed invasive intervention in acute coronary syndromes. *N Engl J Med* 2009;**360**:2165–2175.
50. Andersen HR, Nielsen TT, Rasmussen K, Thuesen L, Kelbaek H, Thayssen P, Abildgaard U, Pedersen F, Madsen JK, Grande P, Villadsen AB, Krusell LR, Haghfelt T, Lomholt P, Husted SE, Vigholt E, Kjaergard HK, Mortensen LS. A Comparison of Coronary Angioplasty with Fibrinolytic Therapy in Acute Myocardial Infarction. *New Eng J Med* 2003;**349**:733–742.

51. Mehta SR, Wood DA, Storey RF, Mehran R, Bainey KR, Nguyen H, Meeks B, Pasquale G Di, López-Sendón J, Faxon DP, Mauri L, Rao S V., Feldman L, Steg PG, Avezum Á, Sheth T, Pinilla-Echeverri N, Moreno R, Campo G, Wrigley B, Kedev S, Sutton A, Oliver R, Rodés-Cabau J, Stanković G, Welsh R, Lavi S, Cantor WJ, Wang J, Nakamya J, Bangdiwala SI, Cairns JA. Complete Revascularization with Multivessel PCI for Myocardial Infarction. *New Eng J Med* 2019;**381**:1411–1421.
52. Lee JM, Kim HK, Park KH, Choo EH, Kim CJ, Lee SH, Kim MC, Hong YJ, Ahn SG, Doh JH, Lee SY, Park SD, Lee HJ, Kang MG, Koh JS, Cho YK, Nam CW, Koo BK, Lee BK, Yun KH, Hong D, Joh HS, Choi KH, Park TK, Yang JH, Song Y Bin, Choi SH, Gwon HC, Hahn JY. Fractional flow reserve versus angiography-guided strategy in acute myocardial infarction with multivessel disease: a randomized trial. *Eur Heart J* 2023;**44**:473–484.
53. Biscaglia S, Guiducci V, Escaned J, Moreno R, Lanzilotti V, Santarelli A, Cerrato E, Sacchetta G, Jurado-Roman A, Menozzi A, Amat Santos I, Díez Gil JL, Ruozzi M, Barbierato M, Fileti L, Picchi A, Lodolini V, Biondi-Zoccai G, Maietti E, Pavasini R, Cimaglia P, Tumscitz C, Erriquez A, Penzo C, Colaiori I, Pignatelli G, Casella G, Iannopollo G, Menozzi M, Varbella F, Caretta G, Dudek D, Barbato E, Tebaldi M, Campo G. Complete or Culprit-Only PCI in Older Patients with Myocardial Infarction. *New Eng J Med* 2023;**389**:889–898.
54. Laforgia PL, Auguadro C, Bronzato S, Durante A. The Reduction of Mortality in Acute Myocardial Infarction: From Bed Rest to Future Directions. *Int J Prev Med* 2022;**13**:56.
55. Rajkumar CA, Foley MJ, Ahmed-Jushuf F, Nowbar AN, Simader FA, Davies JR, O'Kane PD, Haworth P, Routledge H, Kotecha T, Gamma R, Clesham G, Williams R, Din J, Nijjer SS, Curzen N, Ruparelia N, Sinha M, Dungu JN, Ganesananthan S, Khamis R, Mughal L, Kinnaird T, Petraco R, Spratt JC, Sen S, Sehmi J, Collier DJ, Sohaib A, Keeble TR, Cole GD, Howard JP, Francis DP, Shun-Shin MJ, Al-Lamee RK. A Placebo-Controlled Trial of Percutaneous Coronary Intervention for Stable Angina. *N Engl J Med* 2023;**389**:2319-2330
56. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ, Knudtson M, Dada M, Casperson P, Harris CL, Chaitman BR, Shaw L, Gosselin G, Nawaz S, Title LM, Gau G, Blaustein AS, Booth DC, Bates ER, Spertus JA, Berman DS, Mancini GBJ, Weintraub WS. Optimal Medical Therapy with or without PCI for Stable Coronary Disease. *New Eng J Med* 2007;**356**:1503–1516.

References

57. Bruyne B De, Pijls NHJ, Kalesan B, Barbato E, Tonino PAL, Piroth Z, Jagic N, Möbius-Winkler S, Rioufol G, Witt N, Kala P, MacCarthy P, Engström T, Oldroyd KG, Mavromatis K, Manoharan G, Verlee P, Frobert O, Curzen N, Johnson JB, Jüni P, Fearon WF. Fractional Flow Reserve–Guided PCI versus Medical Therapy in Stable Coronary Disease. *New Eng J Med* 2012;**367**:991–1001.

58. Byrne RA, Rossello X, Coughlan JJ, Barbato E, Berry C, Chieffo A, Claeys MJ, Dan G-A, Dweck MR, Galbraith M, Gilard M, Hinterbuchner L, Jankowska EA, Jüni P, Kimura T, Kunadian V, Leosdottir M, Lorusso R, Pedretti RFE, Rigopoulos AG, Rubini Gimenez M, Thiele H, Vranckx P, Wassmann S, Wenger NK, Ibanez B, Group ESD, Halvorsen S, James S, Abdelhamid M, Aboyans V, Marsan NA, Antoniou S, Asteggiano R, Bäck M, Capodanno D, Casado-Arroyo R, Cassese S, Čelutkienė J, Cikes M, Collet J-P, Ducrocq G, Falk V, Fauchier L, Geisler T, Gorog DA, Holmvang L, Jaarsma T, Jones HW, Køber L, Koskinas KC, Kotecha D, Krychtiuk KA, Landmesser U, Lazaros G, Lewis BS, Lindahl B, Linhart A, Løchen M-L, Mamas MA, McEvoy JW, Mihaylova B, Mindham R, Mueller C, Neubeck L, Niebauer J, Nielsen JC, Niessner A, Paradies V, Pasquet AA, Petersen SE, Prescott E, Rakisheva A, Rocca B, Rosano GMC, Sade LE, Schiele F, Siller-Matula JM, Sticherling C, Storey RF, Thielmann M, Vrints C, Windecker S, Wiseth R, Witkowski A, Amine Bouzid M El, Hayrapetyan H, Metzler B, Lancellotti P, Bajrić M, Karamfiloff K, Mitsis A, Ostadal P, Sørensen R, Elwasify T, Marandi T, Ryödi E, Collet J-P, Chukhrukidze A, Mehilli J, Davlouros P, Becker D, Guðmundsdóttir IJ, Crowley J, Abramowitz Y, Indolfi C, Sakhov O, Elezi S, Beishenkulov M, Erglis A, Moussallem N, Benlamin H, Dobilienė O, Degrell P, Balbi MM, Grosu A, Lakhal Z, Berg J ten, Pejkov H, Angel K, Witkowski A, Sousa Almeida M De, Chioncel O, Bertelli L, Stojkovic S, Studenčan M, Radšel P, Ferreiro JL, Ravn-Fischer A, Räber L, Marjeh MYB, Hassine M, Yildirir A, Parkhomenko A, Banning AP, Prescott E, James S, Arbelo E, Baigent C, Borger MA, Buccheri S, Ibanez B, Køber L, Koskinas KC, McEvoy JW, Mihaylova B, Mindham R, Neubeck L, Nielsen JC, Pasquet AA, Rakisheva A, Rocca B, Rossello X, Vaartjes I, Vrints C, Witkowski A, Zeppenfeld K. 2023 ESC Guidelines for the management of acute coronary syndromes: Developed by the task force on the management of acute coronary syndromes of the European Society of Cardiology (ESC). *Eur Heart J* 2023;**44**:3720–3826.

59. Bonow RO, Mann DL, Zipes DP, Libby P, eds. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*. 9th edition. 2012.

60. Austen WG, Edwards JE, Frye RL, Gensini GG, Gott VL, Griffith LS, McGoon DC, Murphy ML, Roe BB. A reporting system on patients evaluated for coronary artery disease. Report of the Ad Hoc Committee for Grading of Coronary Artery Disease, Council on Cardiovascular Surgery, American Heart Association. *Circulation* 1975;**51**:5–40.
61. CASS Investigators. Coronary artery surgery study (CASS): a randomized trial of coronary artery bypass surgery. Survival data. *Circulation* 1983;**68**:939–950.
62. Alderman EL, Stadius M. The angiographie definitions of the Bypass Angioplasty Revascularization Investigation. *Coron Artery Dis* 1992;**3**:1189–1208.
63. Scanlon PJ, Faxon DP, Audet AM, Carabello B, Dehmer GJ, Eagle KA, Legako RD, Leon DF, Murray JA, Nissen SE, Pepine CJ, Watson RM, Ritchie JL, Gibbons RJ, Cheitlin MD, Gardner TJ, Garson A, Russell RO, Ryan TJ, Smith SC. ACC/AHA guidelines for coronary angiography. A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines (Committee on Coronary Angiography). Developed in collaboration with the Society for Cardiac Angiography and Interventions. *J Am Coll Cardiol* 1999;**33**:1756–1824.
64. *The PCR-EAPCI Textbook (Online)*
65. Chesebro JH, Knatterud G, Roberts R, Borer J, Cohen LS, Dalen J, Dodge HT, Francis CK, Hillis D, Ludbrook P, Markis JE, Mueller H, Passamani ER, Powers ER, Rao AK, Robertson T, Ross A, Ryan TJ, Sobel BE, Willerson J, Williams DO, Zaret BL, Braunwald E. Thrombolysis in Myocardial Infarction (TIMI) Trial, Phase I: A comparison between intravenous tissue plasminogen activator and intravenous streptokinase. Clinical findings through hospital discharge. *Circulation* 1987;**76**:142–154.
66. Peter Rentrop K, Cohen M, Blanke H, Phillips RA. Changes in collateral channel filling immediately after controlled coronary artery occlusion by an angioplasty balloon in human subjects. *J Am Coll Cardiol* 1985;**5**:587–592.
67. Özcan C, Deleskog A, Schjerning Olsen AM, Nordahl Christensen H, Lock Hansen M, Hilmar Gislason G. Coronary artery disease severity and long-term cardiovascular risk in patients with myocardial infarction: a Danish nationwide register-based cohort study. *Eur Heart J Cardiovasc Pharmacother* 2018;**4**:25.
68. Reynolds HR, Shaw LJ, Min JK, Page CB, Berman DS, Chaitman BR, Picard MH, Kwong RY, O'Brien SM, Huang Z, Mark DB, Nath RK, Dwivedi SK, Smanio PEP, Stone PH, Held C, Keltai M, Bangalore S, Newman JD, Spertus JA, Stone GW, Maron

DJ, Hochman JS. Outcomes in the ISCHEMIA Trial Based on Coronary Artery Disease and Ischemia Severity. *Circulation* 2021;**144**:1024–1038.

69. Køber L, Thune JJ, Nielsen JC, Haarbo J, Videbæk L, Korup E, Jensen G, Hildebrandt P, Steffensen FH, Bruun NE, Eiskjær H, Brandes A, Thøgersen AM, Gustafsson F, Egstrup K, Videbæk R, Hassager C, Svendsen JH, Høfsten DE, Torp-Pedersen C, Pehrson S. Defibrillator Implantation in Patients with Nonischemic Systolic Heart Failure. *N Engl J Med* 2016;**375**:1221–1230.
70. Bardy GH, Lee KL, Mark DB, Poole JE, Packer DL, Boineau R, Domanski M, Troutman C, Anderson J, Johnson G, McNulty SE, Clapp-Channing N, Davidson-Ray LD, Fraulo ES, Fishbein DP, Luceri RM, Ip JH. Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure. *N Engl J Med* 2005;**352**:225–237.
71. Detre KM, Wright E, Murphy ML, Takaro T. Observer agreement in evaluating coronary angiograms. *Circulation* 1975;**52**:979–986.
72. DeRouen TA, Murray JA, Owen W. Variability in the analysis of coronary arteriograms. *Circulation* 1977;**55**:324–328.
73. Serruys PW, Reiber JHC, Wijns W, Brand M v.d., Kooijman CJ, Katen HJ ten, Hugenholtz PG. Assessment of percutaneous transluminal coronary angioplasty by quantitative coronary angiography: diameter versus densitometric area measurements. *Am J Cardiol* 1984;**54**:482–488.
74. Gronenschild E, Janssen J, Tijdens F. CAAS. II: A second generation system for off-line and on-line quantitative coronary angiography. *Cathet Cardiovasc Diagn* 1994;**33**:61–75.
75. Gensini GG, Kelly AE, Costa BC Da, Huntington PP. Quantitative angiography: the measurement of coronary vasomobility in the intact animal and man. *Chest* 1971;**60**:522–530.
76. Tu S, Koning G, Jukema W, Reiber JHC. Assessment of obstruction length and optimal viewing angle from biplane X-ray angiograms. *Int J Cardiovasc Imaging* 2010;**26**:5–17.
77. Ramcharitar S, Onuma Y, Aben J-P, Consten C, Weijers B, Morel M-A, Serruys P. A novel dedicated quantitative coronary analysis methodology for bifurcation lesions. *EuroIntervention* 2008;**3**:553–557.
78. Zwet PMJ van der, Reiber JHC. A new approach for the quantification of complex lesion morphology: the gradient field transform; basic principles and validation results. *J Am Coll Cardiol* 1994;**24**:216–224.

79. Grundeken MJ, Garcia-Garcia HM, Kumsars I, Lesiak M, Kayaert P, Dens J, Stella PR, Winter RJ de, Laak LL, Généreux P, Kaplan A V., Leon MB, Wykrzykowska JJ, Onuma Y, Serruys PW. Segmental comparison between a dedicated bifurcation stent and balloon angioplasty using intravascular ultrasound and three-dimensional quantitative coronary angiography: A subgroup analysis of the Tryton IDE randomized trial. *Catheter Cardiovasc Interv* 2017;**89**:E53–E63.
80. Girasis C, Schuurbiers JCH, Muramatsu T, Aben JP, Onuma Y, Soekhradj S, Morel MA, Geuns RJ Van, Wentzel JJ, Serruys PW. Advanced three-dimensional quantitative coronary angiographic assessment of bifurcation lesions: methodology and phantom validation. *EuroIntervention* 2013;**8**:1451–1460.
81. Tuinenburg JC, Koning G, Rareş A, Janssen JP, Lansky AJ, Reiber JHC. Dedicated bifurcation analysis: basic principles. *Int J Cardiovasc Imaging* 2011;**27**:167–174.
82. Lansky A, Tuinenburg J, Costa M, Maeng M, Koning G, Popma J, Cristea E, Gavit L, Costa R, Rares A, Es GA Van, Lefevre T, Reiber H, Louvard Y, Morice MC. Quantitative angiographic methods for bifurcation lesions: a consensus statement from the European Bifurcation Group. *Catheter Cardiovasc Interv* 2009;**73**:258–266.
83. Schuurbiers JCH, Lopez NG, Ligthart J, Gijsen FJH, Dijkstra J, Serruys PW, Steen AF Van der, Wentzel JJ. In vivo validation of CAAS QCA-3D coronary reconstruction using fusion of angiography and intravascular ultrasound (ANGUS). *Catheter Cardiovasc Interv* 2009;**73**:620–626.
84. Dvir D, Marom H, Guetta V, Kornowski R. Three-dimensional coronary reconstruction from routine single-plane coronary angiograms: in vivo quantitative validation. *Int J Cardiovasc Intervent* 2005;**7**:141–145.
85. Fischer JJ, Samady H, McPherson JA, Sarembock IJ, Powers ER, Gimple LW, Ragosta M. Comparison between visual assessment and quantitative angiography versus fractional flow reserve for native coronary narrowings of moderate severity. *Am J Cardiol* 2002;**90**:210–215.
86. Zhang H, Mu L, Hu S, Nallamothu BK, Lansky AJ, Xu B, Bouras G, Cohen DJ, Spertus JA, Masoudi FA, Curtis JP, Gao R, Ge J, Yang Y, Li J, Li X, Zheng X, Li Y, Krumholz HM, Jiang L, Wang J, Zhang H, Huang C, Ma L, Hu H, Kong X, Zhao R, Du W, Li H, Yang P, Feng Z, Mao C, Li B, Han Q, Duan L, Jin C, Guo Y, Chen X, Lin F, Wu Y, Tian Y, Li J, Zhu C, Dong Q, Pan H, Gao P, Sun L, Zhang W, Xie S, Zhuang Y, Li W, Ge Z, Wen H, Liu Q, Li Y, Han Y, Zhao P, Zhao W, Hong L, Li L, Yuan L, Li Y, Zhao P, Sun J, Chai R, Yang S, Li Y, Gao J, Zhang S, Yang Y, Wu G, Mao J, Zheng C, Bian

H, Chen B, Cao J, Yan H, Chen L, Liu Q, Chen L, Li B, Yang B, Li J, Wang J, Liu F, Meng X, Shao B, Liao Z, Liu Z, Jing W, Guo Z, Zhou C, Wang Y, Yu T, Li D, Luo Y, Zhu H, Lin J, Chai D, Zhao W, Zheng Y, Liu Z, Zhou W, Zou L, Jin Y, Zhang X, Zhang X, Li Z, Liu Y, Yu Q, Xing Y, Wu W, Feng J, Fa X, Zhang L, Sun L, Liu L, Qu P, Wang H, Lou D, Yuan D, Wang D, Shen X, Jiang J, Huang H, Zeng J, Wu M, Zhou Y, Zhang H, Jun L, Hu M, Zhang L, Guo Y, Sun H. Comparison of Physician Visual Assessment With Quantitative Coronary Angiography in Assessment of Stenosis Severity in China. *JAMA Intern Med* 2018;**178**:239–247.

87. Beauman GJ, Vogel RA. Accuracy of individual and panel visual interpretations of coronary arteriograms; Implications for clinical decisions. *J Am Coll Cardiol* 1990;**16**:108–113.
88. Nallamothu BK, Spertus JA, Lansky AJ, Cohen DJ, Jones PG, Kureshi F, Dehmer GJ, Drozda JP, Walsh MN, Brush JE, Koenig GC, Waites TF, Gantt DS, Kichura G, Chazal RA, O'Brien PK, Valentine CM, Rumsfeld JS, Reiber JHC, Elmore JG, Krumholz RA, Weaver WD, Krumholz HM. Comparison of clinical interpretation with visual assessment and quantitative coronary angiography in patients undergoing percutaneous coronary intervention in contemporary practice: the Assessing Angiography (A2) project. *Circulation* 2013;**127**:1793–1800.
89. Adjedj J, Xaplanteris P, Toth G, Ferrara A, Pellicano M, Ciccarelli G, Floré V, Barbato E, Bruyne B De. Visual and Quantitative Assessment of Coronary Stenoses at Angiography Versus Fractional Flow Reserve: The Impact of Risk Factors. *Circ Cardiovasc Imaging* 2017;**10**.
90. Toth GG, Toth B, Johnson NP, Vroey F De, Serafino L Di, Pyxaras S, Rusinaru D, Gioia G Di, Pellicano M, Barbato E, Mieghem C Van, Heyndrickx GR, Bruyne B De, Wijns W. Revascularization decisions in patients with stable angina and intermediate lesions: results of the international survey on interventional strategy. *Circ Cardiovasc Interv* 2014;**7**:751–759.
91. Frye RL, Alderman EL, Andrews K, Bost J, Bourassa M, Chaitman BR, Detre K, Faxon DP, Follmann D, Hlatky M, Jones RH, Kelsey SF, Rogers WJ, Rosen AD, Schaff H, Sellers MA, Sopko G, Tyrrell KS, Williams DO. Comparison of Coronary Bypass Surgery with Angioplasty in Patients with Multivessel Disease. *New Eng J Med* 1996;**335**:217.
92. Serruys PW, Morice M-C, Kappetein AP, Colombo A, Holmes DR, Mack MJ, Ståhle E, Feldman TE, Brand M van den, Bass EJ, Dyck N Van, Leadley K, Dawkins KD, Mohr

FW. Percutaneous Coronary Intervention versus Coronary-Artery Bypass Grafting for Severe Coronary Artery Disease. *New Eng J Med* 2009;**360**:961–972.

93. Tonino PAL, Bruyne B De, Pijls NHJ, Siebert U, Ikeno F, `t Veer M van, Klauss V, Manoharan G, Engstrøm T, Oldroyd KG, Lee PN Ver, MacCarthy PA, Fearon WF. Fractional Flow Reserve versus Angiography for Guiding Percutaneous Coronary Intervention. *New Eng J Med* 2009;**360**:213–224.
94. Velazquez EJ, Lee KL, Deja MA, Jain A, Sopko G, Marchenko A, Ali IS, Pohost G, Gradinac S, Abraham WT, Yii M, Prabhakaran D, Szwed H, Ferrazzi P, Petrie MC, O'Connor CM, Panchavinnin P, She L, Bonow RO, Rankin GR, Jones RH, Rouleau J-L. Coronary-Artery Bypass Surgery in Patients with Left Ventricular Dysfunction. *New Eng J Med* 2011;**364**:1607–1616.
95. Maron DJ, Hochman JS, Reynolds HR, Bangalore S, O'Brien SM, Boden WE, Chaitman BR, Senior R, López-Sendón J, Alexander KP, Lopes RD, Shaw LJ, Berger JS, Newman JD, Sidhu MS, Goodman SG, Ruzyllo W, Gosselin G, Maggioni AP, White HD, Bhargava B, Min JK, Mancini GBJ, Berman DS, Picard MH, Kwong RY, Ali ZA, Mark DB, Spertus JA, Krishnan MN, Elghamaz A, Moorthy N, Hueb WA, Demkow M, Mavromatis K, Bockeria O, Peteiro J, Miller TD, Szwed H, Doerr R, Keltai M, Selvanayagam JB, Steg PG, Held C, Kohsaka S, Mavromichalis S, Kirby R, Jeffries NO, Frank E. Harrell Jr, Rockhold FW, Broderick S, T. Bruce Ferguson Jr, Williams DO, Harrington RA, Stone GW, Rosenberg Y. Initial Invasive or Conservative Strategy for Stable Coronary Disease. *New Eng J Med* 2020;**382**:1395–1407.
96. Wald DS, Morris JK, Wald NJ, Chase AJ, Edwards RJ, Hughes LO, Berry C, Oldroyd KG. Randomized Trial of Preventive Angioplasty in Myocardial Infarction. *New Eng J Med* 2013;**369**:1115–1123.
97. Engstrøm T, Kelbæk H, Helqvist S, Høfsten DE, Kløvgaard L, Holmvang L, Jørgensen E, Pedersen F, Saunamäki K, Clemmensen P, Backer O De, Ravkilde J, Tilsted HH, Villadsen AB, Aarøe J, Jensen SE, Raungaard B, Køber L. Complete revascularisation versus treatment of the culprit lesion only in patients with ST-segment elevation myocardial infarction and multivessel disease (DANAMI-3 - PRIMULTI): An open-label, randomised controlled trial. *The Lancet* 2015;**386**:665–671.
98. Smits PC, Abdel-Wahab M, Neumann F-J, Boxma-de Klerk BM, Lunde K, Schotborgh CE, Piroth Z, Horak D, Wlodarczak A, Ong PJ, Hambrecht R, Angerås O, Richardt G, Omerovic E. Fractional Flow Reserve–Guided Multivessel Angioplasty in Myocardial Infarction. *New Eng J Med* 2017;**376**:1234–1244.

References

99. Gershlick AH, Khan JN, Kelly DJ, Greenwood JP, Sasikaran T, Curzen N, Blackman DJ, Dalby M, Fairbrother KL, Banya W, Wang D, Flather M, Hetherington SL, Kelion AD, Talwar S, Gunning M, Hall R, Swanton H, McCann GP. Randomized trial of complete versus lesion-only revascularization in patients undergoing primary percutaneous coronary intervention for STEMI and multivessel disease: the CvLPRIT trial. *J Am Coll Cardiol* 2015;**65**:963–972.
100. Xu B, Tu S, Song L, Jin Z, Yu B, Fu G, Zhou Y, Wang J, Chen Y, Pu J, Chen L, Qu X, Yang J, Liu X, Guo L, Shen C, Zhang Y, Zhang Q, Pan H, Fu X, Liu J, Zhao Y, Escaned J, Wang Y, Fearon WF, Dou K, Kirtane AJ, Wu Y, Serruys PW, Yang W, Wijns W, Guan C, Leon MB, Qiao S, Stone GW. Angiographic quantitative flow ratio-guided coronary intervention (FAVOR III China): a multicentre, randomised, sham-controlled trial. *The Lancet* 2021;**398**:2149–2159.
101. Perera D, Clayton T, O'Kane PD, Greenwood JP, Weerackody R, Ryan M, Morgan HP, Dodd M, Evans R, Canter R, Arnold S, Dixon LJ, Edwards RJ, Silva K De, Spratt JC, Conway D, Cotton J, McEntegart M, Chiribiri A, Saramago P, Gershlick A, Shah AM, Clark AL, Petrie MC. Percutaneous Revascularization for Ischemic Left Ventricular Dysfunction. *New Eng J Med* 2022;**387**(15):1351-1360.
102. Virani SS, Newby LK, Arnold S V., Bittner V, Brewer LPC, Demeter SH, Dixon DL, Fearon WF, Hess B, Johnson HM, Kazi DS, Kolte D, Kumbhani DJ, Lofaso J, Mahtta D, Mark DB, Minissian M, Navar AM, Patel AR, Piano MR, Rodriguez F, Talbot AW, Taqueti VR, Thomas RJ, Diepen S Van, Wiggins B, Williams MS. 2023 AHA/ACC/ACCP/ASPC/NLA/PCNA Guideline for the Management of Patients With Chronic Coronary Disease: A Report of the American Heart Association/American College of Cardiology Joint Committee on Clinical Practice Guidelines. *Circulation* 2023;**148**:E9–E119.
103. Lawton JS, Tamis-Holland JE, Bangalore S, Bates ER, Beckie TM, Bischoff JM, Bittl JA, Cohen MG, Dimaio JM, Don CW, Fremes SE, Gaudino MF, Goldberger ZD, Grant MC, Jaswal JB, Kurlansky PA, Mehran R, Metkus TS, Nnacheta LC, Rao S V., Sellke FW, Sharma G, Yong CM, Zwischenberger BA. 2021 ACC/AHA/SCAI Guideline for Coronary Artery Revascularization: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2022;**145**:E18–E114.
104. Vahanian A, Beyersdorf F, Praz F, Milojevic M, Baldus S, Bauersachs J, Capodanno D, Conradi L, Bonis M De, Paulis R De, Delgado V, Freemantle N, Gilard M, Haugaa KH,

Jeppsson A, Jüni P, Pierard L, Prendergast BD, Sádaba JR, Tribouilloy C, Wojakowski W, Group ESD, Neumann F-J, Myers P, Abdelhamid M, Achenbach S, Asteggiano R, Barili F, Borger MA, Carrel T, Collet J-P, Foldager D, Habib G, Hassager C, Irs A, Iung B, Jahangiri M, Katus HA, Koskinas KC, Massberg S, Mueller CE, Nielsen JC, Pibarot P, Rakisheva A, Roffi M, Rubboli A, Shlyakhto E, Siepe M, Sitges M, Sondergaard L, Sousa-Uva M, Tarantini G, Zamorano JL, Praz F, Milojevic M, Baldus S, Bauersachs J, Capodanno D, Conradi L, Bonis M De, Paulis R De, Delgado V, Freemantle N, Gilard M, Haugaa KH, Jeppsson A, Jüni P, Pierard L, Prendergast BD, Sádaba JR, Tribouilloy C, Wojakowski W. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. Developed by the Task Force for the management of valvular heart disease of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS). *Eur Heart J* 2022;**43**:561–632.

105. Tonino PAL, Fearon WF, Bruyne B De, Oldroyd KG, Leesar MA, Lee PN Ver, MacCarthy PA, van't Veer M, Pijls NHJ. Angiographic Versus Functional Severity of Coronary Artery Stenoses in the FAME Study: Fractional Flow Reserve Versus Angiography in Multivessel Evaluation. *J Am Coll Cardiol* 2010;**55**:2816–2821.
106. Garcia-Garcia HM, McFadden EP, Farb A, Mehran R, Stone GW, Spertus J, Onuma Y, Morel MA, Es GA Van, Zuckerman B, Fearon WF, Taggart D, Kappetein AP, Krucoff MW, Vranckx P, Windecker S, Cutlip D, Serruys PW. Standardized end point definitions for coronary intervention trials: The academic research consortium-2 consensus document. *Circulation* 2018;**137**:2635–2650.
107. Suzuki N, Asano T, Nakazawa G, Aoki J, Tanabe K, Hibi K, Ikari Y, Kozuma K. Clinical expert consensus document on quantitative coronary angiography from the Japanese Association of Cardiovascular Intervention and Therapeutics. *Cardiovasc Interv Ther* 2020;**35**:105.
108. Gutiérrez-Chico JL, Serruys PW, Girasis C, Garg S, Onuma Y, Brugaletta S, García-García H, Es GA Van, Regar E. Quantitative multi-modality imaging analysis of a fully bioresorbable stent: a head-to-head comparison between QCA, IVUS and OCT. *Int J Cardiovasc Imaging* 2012;**28**:467–478.
109. Bruining N, Tanimoto S, Otsuka M, Weustink A, Ligthart J, Winter S De, Mieghem C Van, Nieman K, Feyter PJ De, Domburg RT Van, Serruys PW. Quantitative multi-modality imaging analysis of a bioabsorbable poly-L-lactic acid stent design in the acute phase: a comparison between 2- and 3D-QCA, QCU and QMSCT-CA. *EuroIntervention* 2008;**4**:285–291.

References

110. Abizaid AS, Mintz GS, Abizaid A, Mehran R, Lansky AJ, Pichard AD, Satler LF, Wu H, Kent KM, Leon MB. One-year follow-up after intravascular ultrasound assessment of moderate left main coronary artery disease in patients with ambiguous angiograms. *J Am Coll Cardiol* 1999;**34**:707–715.
111. Ozaki Y, Violaris AG, Kobayashi T, Keane D, Camenzind E, Mario C Di, Feyter P De, Roelandt JRTC, Serruys PW. Comparison of coronary luminal quantification obtained from intracoronary ultrasound and both geometric and videodensitometric quantitative angiography before and after balloon angioplasty and directional atherectomy. *Circulation* 1997;**96**:491–499.
112. Goto K, Mintz GS, Litherland C, Lansky AJ, Weisz G, McPherson JA, Bruyne B De, Serruys PW, Stone GW, Maehara A. Lumen Measurements From Quantitative Coronary Angiography and IVUS: A PROSPECT Substudy. *JACC Cardiovasc Imaging* 2016;**9**:1011–1013.
113. Voros S, Rinehart S, Vazquez-Figueroa JG, Kalynych A, Karmpaliotis D, Qian Z, Joshi PH, Anderson H, Murrieta L, Wilmer C, Carlson H, Ballard W, Brown C. Prospective, head-to-head comparison of quantitative coronary angiography, quantitative computed tomography angiography, and intravascular ultrasound for the prediction of hemodynamic significance in intermediate and severe lesions, using fractional flow reserve as reference standard (from the ATLANTA I and II Study). *Am J Cardiol* 2014;**113**:23–29.
114. Naganuma T, Latib A, Costopoulos C, Takagi K, Naim C, Sato K, Miyazaki T, Kawaguchi M, Panoulas VF, Basavarajaiah S, Figini F, Chieffo A, Montorfano M, Carlino M, Colombo A. The role of intravascular ultrasound and quantitative angiography in the functional assessment of intermediate coronary lesions: correlation with fractional flow reserve. *Cardiovasc Revasc Med* 2014;**15**:3–7.
115. Yong ASC, Ng ACC, Brieger D, Lowe HC, Ng MKC, Kritharides L. Three-dimensional and two-dimensional quantitative coronary angiography, and their prediction of reduced fractional flow reserve. *Eur Heart J* 2011;**32**:345–353.
116. Nijjer SS, Dewaard GA, Sen S, Hoef TP Van De, Petraco R, Echavarría-Pinto M, Lavieren MA Van, Meuwissen M, Danad I, Knaapen P, Escaned J, Piek JJ, Davies JE, Royen N Van. Coronary pressure and flow relationships in humans: phasic analysis of normal and pathological vessels and the implications for stenosis assessment: a report from the Iberian-Dutch-English (IDEAL) collaborators. *Eur Heart J* 2016;**37**:2069–2080.

117. Tang X, Dai N, Zhang BC, Cai H, Huo Y, Yang M, Jiang Y, Duan S, Shen J, Zhu M, Xu Y, Ge J. Comparison of 2D-QCA, 3D-QCA and coronary angiography derived FFR in predicting myocardial ischemia assessed by CZT-SPECT MPI. *J Nucl Cardiol* 2023;**30**:1973–1982.
118. Tu S, Westra J, Yang J, Birgelen C von, Ferrara A, Pellicano M, Nef H, Tebaldi M, Murasato Y, Lansky A, Barbato E, Heijden LC van der, Reiber JHC, Holm NR, Wijns W. Diagnostic Accuracy of Fast Computational Approaches to Derive Fractional Flow Reserve From Diagnostic Coronary Angiography: The International Multicenter FAVOR Pilot Study. *JACC Cardiovasc Interv* 2016;**9**:2024–2035.
119. Shah R, Yow E, Jones WS, Kohl LP, Kosinski AS, Hoffmann U, Lee KL, Fordyce CB, Mark DB, Lowe A, Douglas PS, Patel MR. Comparison of visual assessment of coronary stenosis with independent quantitative coronary angiography: Findings from the PROMISE trial. *Am Heart J* 2017;**184**:1.
120. Quantitative coronary angiography-guidance versus intravascular ultrasound-guidance for drug-eluting stent implantation (GUIDE-DES trial): A multicenter, non-inferiority, randomised-controlled trial. *Presented at ESC Congress 2023*.
121. Topol EJ. Coronary angioplasty for acute myocardial infarction. *Ann Intern Med* 1988;**109**:970–980.
122. Ryan TJ, Bauman WB, Kennedy JW, Kereiakes DJ, King SB, McCallister BD, Smith SC, Ullyot DJ. Guidelines for percutaneous transluminal coronary angioplasty: A report of the American College of Cardiology/American Heart Association Task Force on the Assesment of Diagnostic and Therapeutic Cardiovascular Procedures (Committee on Percutaneous Transluminal Coronary Angioplasty). *J Am Coll Cardiol* 1993;**22**:2033–2054.
123. Leaman DM, Brower RW, Meester GT, Serruys P, Brand M van den. Coronary artery atherosclerosis: severity of the disease, severity of angina pectoris and compromised left ventricular function. *Circulation* 1981;**63**:285–292.
124. Holm NR, Mäkikallio T, Lindsay MM, Spence MS, Erglis A, Menown IBA, Trovik T, Kellerth T, Kalinauskas G, Mogensen LJH, Nielsen PH, Niemelä M, Lassen JF, Oldroyd K, Berg G, Stradins P, Walsh SJ, Graham ANJ, Endresen PC, Fröbert O, Trivedi U, Anttila V, Hildick-Smith D, Thuesen L, Christiansen EH, Lindsay M, Eskola M, Romppanen H, Kellerth T, Jensen LO, Linder RBA, Pentikainen M, Hervold A, Banning A, Zaman A, Cotton J, Eriksen E, Margus S, Mogensen LJH, Kervinen K, Berg G, Hanratty CG, Kumsars I, Steigen TK, Graham AN, Corbascio M, Kajander O,

Hartikainen J, Anttila V. Percutaneous coronary angioplasty versus coronary artery bypass grafting in the treatment of unprotected left main stenosis: updated 5-year outcomes from the randomised, non-inferiority NOBLE trial. *The Lancet* 2020;**395**:191–199.

125. Stone GW, Kappetein AP, Sabik JF, Pocock SJ, Morice M-C, Puskas J, Kandzari DE, Karmpaliotis D, Brown WM, Lembo NJ, Banning A, Merkely B, Horkay F, Boonstra PW, Boven AJ van, Ungi I, Bogáts G, Mansour S, Noiseux N, Sabaté M, Pomar J, Hickey M, Gershlick A, Buszman PE, Bochenek A, Schampaert E, Pagé P, Modolo R, Gregson J, Simonton CA, Mehran R, Kosmidou I, Généreux P, Crowley A, Dressler O, Serruys PW. Five-Year Outcomes after PCI or CABG for Left Main Coronary Disease. *New Eng J Med* 2019;**381**:1820–1830.
126. Wijns W, Serruys PW, Reiber JHC, Brand M van den, Simoons ML, Kooijman CJ, Balakumaran K, Hugenholtz PG. Quantitative angiography of the left anterior descending coronary artery: correlations with pressure gradient and results of exercise thallium scintigraphy. *Circulation* 1985;**71**:273–279.
127. Serruys PW, Wijns W, Reiber JH, Feyter P de, Brand M van den, Piscione F, Hugenholtz PG. Values and limitations of transstenotic pressure gradients measured during percutaneous coronary angioplasty *Herz* 1985;**103**:337-42*:*
128. Emanuelsson H, Dohnal M, Lamm C, Tenerz L. Initial experiences with a miniaturized pressure transducer during coronary angioplasty. *Cathet Cardiovasc Diagn* 1991;**24**:137–143.
129. Mario C Di, Feyter PJ De, Slager CJ, Jaegere P De, Roelandt JRTC, Serruys PW. Intracoronary blood flow velocity and transstenotic pressure gradient using sensor-tip pressure and doppler guidewires: A new technology for the assessment of stenosis severity in the catheterization laboratory. *Cathet Cardiovasc Diagn* 1993;**28**:311–319.
130. Serruys PW, Mario C Di, Meneveau N, Jaegere P de, Strikwerda S, Feyter PJ de, Emanuelsson H. Intracoronary pressure and flow velocity with sensor-tip guidewires: a new methodologic approach for assessment of coronary hemodynamics before and after coronary interventions. *Am J Cardiol* 1993;**71**.
131. Pijls NHJ, Son JAM Van, Kirkeeide RL, Bruyne B De, Gould KL. Experimental basis of determining maximum coronary, myocardial, and collateral blood flow by pressure measurements for assessing functional stenosis severity before and after percutaneous transluminal coronary angioplasty. *Circulation* 1993;**87**:1354–1367.

132. Götberg M, Cook CM, Sen S, Nijjer S, Escaned J, Davies JE. The Evolving Future of Instantaneous Wave-Free Ratio and Fractional Flow Reserve. *J Am Coll Cardiol* 2017;**70**:1379-1402 .

133. Toth GG, Bruyne B De, Rusinaru D, Gioia G Di, Bartunek J, Pellicano M, Vanderheyden M, Adjedj J, Wijns W, Pijls NHJ, Barbato E. Impact of Right Atrial Pressure on Fractional Flow Reserve Measurements: Comparison of Fractional Flow Reserve and Myocardial Fractional Flow Reserve in 1,600 Coronary Stenoses. *JACC Cardiovasc Interv* 2016;**9**:453–459.

134. Bruyne B De, Bartunek J, Sys SU, Heyndrickx GR. Relation between myocardial fractional flow reserve calculated from coronary pressure measurements and exercise-induced myocardial ischemia. *Circulation* 1995;**92**:39–46.

135. Bech GJW, Bruyne B De, Pijls NHJ, Muinck ED de, Hoorntje JCA, Escaned J, Stella PR, Boersma E, Bartunek J, Koolen JJ, Wijns W. Fractional Flow Reserve to Determine the Appropriateness of Angioplasty in Moderate Coronary Stenosis. *Circulation* 2001;**103**:2928–2934.

136. Baptista SB, Raposo L, Santos L, Ramos R, Calé R, Jorge E, Machado C, Costa M, Infante de Oliveira E, Costa J, Pipa J, Fonseca N, Guardado J, Silva B, Sousa M-J, Silva JC, Rodrigues A, Seca L, Fernandes R. Impact of Routine Fractional Flow Reserve Evaluation During Coronary Angiography on Management Strategy and Clinical Outcome. *Circ Cardiovasc Interv* 2016;**9**:e003288.

137. Xaplanteris P, Fournier S, Pijls NHJ, Fearon WF, Barbato E, Tonino PAL, Engstrøm T, Kääb S, Dambrink J-H, Rioufol G, Toth GG, Piroth Z, Witt N, Fröbert O, Kala P, Linke A, Jagic N, Mates M, Mavromatis K, Samady H, Irimpen A, Oldroyd K, Campo G, Rothenbühler M, Jüni P, Bruyne B De. Five-Year Outcomes with PCI Guided by Fractional Flow Reserve. *New Eng J Med* 2018;**379**:250–259.

138. Rioufol G, Dérimay F, Roubille F, Perret T, Motreff P, Angoulvant D, Cottin Y, Meunier L, Cetran L, Cayla G, Harbaoui B, Wiedemann JY, Belle É Van, Pouillot C, Noirclerc N, Morelle JF, Soto FX, Caussin C, Bertrand B, Lefèvre T, Dupouy P, Lesault PF, Albert F, Barthelemy O, Koning R, Leborgne L, Barnay P, Chapon P, Armero S, Lafont A, Piot C, Amaz C, Vaz B, Benyahya L, Varillon Y, Ovize M, Mewton N, Finet G, Fournier A, Jarry G, Leleu F, Malaquin D, Mirode A, Belle L, Mangin L, Hirsch J Lou, Metge M, Pansiery M, Soto FX, Boge A, HadjHamou K, Miliani I, Molins G, Mourot S, Pelletier M, Ressencourt O, Schaad F, Coste P, Chasseriaud W, Poustis P, Morelle JF, Demicheli T, Range G, Thuaire C, Barber-Chamoux N, Combaret N, Malclès G, Souteyrand G,

Buffet P, Gudjonvick A, L'Huillier I, Lorgis L, Richard C, Baronne-Rochette G, Bouvaist H, Marlière S, Ormezzano O, Vanzetto G, Trouillet C, Valy Y, VanBelle E, Bauters C, Delhaye C, Lemesle G, Rihani R, Graux P, Lemahieu JM, Besnard C, Courand PY, Dauphin R, Lantelme P, Caignault JR, Dubreuil O, Ranc S, Ritz B, Bergerot C, Bochaton T, Bonnefoy-Cudraz E, Bresson D, Dementhon J, Derimay F, Green L, Prieur C, Sanchez I, Zouaghi O, Arméro S, Ben-Amer H, Chevalier B, Garot P, Hovasse T, Louvard Y, Morice MC, Tavolaro O, Unterseeh T, Cung DTT, Macia JC, Levy G, Roth O, Jacquemin L, Cornillet L, Ledermann B, Schmutz L, Karam N, Rahal S, Amabile N, Girard P, Veugeois A, Barthélémy O, Collet JP, Montalescot G, Berland J, Godin M, Landolff Q, Zoghlami B, Bougrini K, Geyer C, Glanenapp J, Mascarel P, Rambaud G, ViFane R, Desveaux B, Ivanes F, Pacouret G, Quilliet LE, SaintEtienne C, Bretelle C, Champin S. Fractional Flow Reserve to Guide Treatment of Patients With Multivessel Coronary Artery Disease. *J Am Coll Cardiol* 2021;**78**:1875–1885.

139. Davies JE, Whinnett ZI, Francis DP, Manisty CH, Aguado-Sierra J, Willson K, Foale RA, Malik IS, Hughes AD, Parker KH, Mayet J. Evidence of a Dominant Backward-Propagating "Suction" Wave Responsible for Diastolic Coronary Filling in Humans, Attenuated in Left Ventricular Hypertrophy. *Circulation* 2006;**113**:1768–1778.
140. Nijjer SS, Sen S, Petraco R, Davies JE. Advances in coronary physiology. *Circ J* 2015;**79**:1172–1184.
141. Sen S, Escaned J, Malik IS, Mikhail GW, Foale RA, Mila R, Tarkin J, Petraco R, Broyd C, Jabbour R, Sethi A, Baker CS, Bellamy M, Al-Bustami M, Hackett D, Khan M, Lefroy D, Parker KH, Hughes AD, Francis DP, Mario C Di, Mayet J, Davies JE. Development and Validation of a New Adenosine-Independent Index of Stenosis Severity From Coronary Wave–Intensity Analysis: Results of the ADVISE (ADenosine Vasodilator Independent Stenosis Evaluation) Study. *J Am Coll Cardiol* 2012;**59**:1392–1402.
142. Petraco R, Escaned J, Sen S, Nijjer S, Asrress KN, Echavarria-Pinto M, Lockie T, Khawaja MZ, Cuevas C, Foin N, Broyd C, Foale RA, Hadjiloizou N, Malik IS, Mikhail GW, Sethi A, Kaprielian R, Baker CS, Lefroy D, Bellamy M, Al-Bustami M, Khan MA, Hughes AD, Francis DP, Mayet J, Mario C Di, Redwood S, Davies JE. Classification performance of instantaneous wave-free ratio (iFR) and fractional flow reserve in a clinical population of intermediate coronary stenoses: results of the ADVISE registry. *EuroIntervention* 2013;**9**:91–101.

143. Nobre Menezes M, Francisco ARG, Carrilho Ferreira P, Jorge C, Torres D, Cardoso P, Duarte JA, Marques da Costa J, Infante de Oliveira E, Pinto FJ, Canas da Silva P. Comparative analysis of fractional flow reserve and instantaneous wave-free ratio: Results of a five-year registry. *Rev Port. Card* 2018;**37**:511–520.

144. Jeremias A, Maehara A, Généreux P, Asrress KN, Berry C, Bruyne B De, Davies JE, Escaned J, Fearon WF, Gould KL, Johnson NP, Kirtane AJ, Koo BK, Marques KM, Nijjer S, Oldroyd KG, Petraco R, Piek JJ, Pijls NH, Redwood S, Siebes M, Spaan JAE, Van'T Veer M, Mintz GS, Stone GW. Multicenter Core Laboratory Comparison of the Instantaneous Wave-Free Ratio and Resting Pd/Pa With Fractional Flow Reserve: The RESOLVE Study. *J Am Coll Cardiol* 2014;**63**:1253–1261.

145. Johnson NP, Kirkeeide RL, Asrress KN, Fearon WF, Lockie T, Marques KMJ, Pyxaras SA, Rolandi MC, 'T Veer M Van, Bruyne B De, Piek JJ, Pijls NHJ, Redwood S, Siebes M, Spaan JAE, Gould KL. Does the Instantaneous Wave-Free Ratio Approximate the Fractional Flow Reserve? *J Am Coll Cardiol* 2013;**61**:1428–1435.

146. Sen S, Asrress KN, Nijjer S, Petraco R, Malik IS, Foale RA, Mikhail GW, Foin N, Broyd C, Hadjiloizou N, Sethi A, Al-Bustami M, Hackett D, Khan MA, Khawaja MZ, Baker CS, Bellamy M, Parker KH, Hughes AD, Francis DP, Mayet J, Mario C Di, Escaned J, Redwood S, Davies JE. Diagnostic Classification of the Instantaneous Wave-Free Ratio Is Equivalent to Fractional Flow Reserve and Is Not Improved With Adenosine Administration: Results of CLARIFY (Classification Accuracy of Pressure-Only Ratios Against Indices Using Flow Study). *J Am Coll Cardiol* 2013;**61**:1409–1420.

147. Hoef TP Van De, Meuwissen M, Escaned J, Sen S, Petraco R, Lavieren MA Van, Echavarria-Pinto M, Nolte F, Nijjer S, Chamuleau SAJ, Voskuil M, Eck-Smit BLF Van, Verberne HJ, Henriques JPS, Koch KT, Winter RJ De, Spaan JAE, Siebes M, Tijssen JGP, Davies JE, Piek JJ. Head-to-head comparison of basal stenosis resistance index, instantaneous wave-free ratio, and fractional flow reserve: diagnostic accuracy for stenosis-specific myocardial ischaemia. *EuroIntervention* 2015;**11**:914–925.

148. Waard G de, Danad I, Cunha RP da, Teunissen P, Hoef T van de, Raijmakers P, Lammertsma A, Davies J, Knaapen P, Royen N Van. Hyperemic FFR and baseline ifr have an equivalent diagnostic accuracy when compared to myocardial blood flow quantified by H215O pet perfusion imaging. *J Am Coll Cardiol* 2014;**63**:A1692.

149. Escaned J, Echavarría-Pinto M, Garcia-Garcia HM, Hoef TP Van De, Vries T De, Kaul P, Raveendran G, Altman JD, Kurz HI, Brechtken J, Tulli M, Birgelen C Von, Schneider JE, Khashaba AA, Jeremias A, Baucum J, Moreno R, Meuwissen M, Mishkel G, Geuns

RJ Van, Levite H, Lopez-Palop R, Mayhew M, Serruys PW, Samady H, Piek JJ, Lerman A. Prospective Assessment of the Diagnostic Accuracy of Instantaneous Wave-Free Ratio to Assess Coronary Stenosis Relevance: Results of ADVISE II International, Multicenter Study (ADenosine Vasodilator Independent Stenosis Evaluation II). *JACC Cardiovasc Interv* 2015;**8**:824–833.

150. Davies JE, Sen S, Dehbi H-M, Al-Lamee R, Petraco R, Nijjer SS, Bhindi R, Lehman SJ, Walters D, Sapontis J, Janssens L, Vrints CJ, Khashaba A, Laine M, Belle E Van, Krackhardt F, Bojara W, Going O, Härle T, Indolfi C, Niccoli G, Ribichini F, Tanaka N, Yokoi H, Takashima H, Kikuta Y, Erglis A, Vinhas H, Canas Silva P, Baptista SB, Alghamdi A, Hellig F, Koo B-K, Nam C-W, Shin E-S, Doh J-H, Brugaletta S, Alegria-Barrero E, Meuwissen M, Piek JJ, Royen N van, Sezer M, Mario C Di, Gerber RT, Malik IS, Sharp ASP, Talwar S, Tang K, Samady H, Altman J, Seto AH, Singh J, Jeremias A, Matsuo H, Kharbanda RK, Patel MR, Serruys P, Escaned J. Use of the Instantaneous Wave-free Ratio or Fractional Flow Reserve in PCI. *New Eng J Med* 2017;**376**:1824–1834.
151. Götberg M, Christiansen EH, Gudmundsdottir IJ, Sandhall L, Danielewicz M, Jakobsen L, Olsson S-E, Öhagen P, Olsson H, Omerovic E, Calais F, Lindroos P, Maeng M, Tödt T, Venetsanos D, James SK, Kåregren A, Nilsson M, Carlsson J, Hauer D, Jensen J, Karlsson A-C, Panayi G, Erlinge D, Fröbert O. Instantaneous Wave-free Ratio versus Fractional Flow Reserve to Guide PCI. *New Eng J Med* 2017;**376**:1813–1823.
152. Escaned J, Ryan N, Mejía-Rentería H, Cook CM, Dehbi HM, Alegria-Barrero E, Alghamdi A, Al-Lamee R, Altman J, Ambrosia A, Baptista SB, Bertilsson M, Bhindi R, Birgander M, Bojara W, Brugaletta S, Buller C, Calais F, Silva PC, Carlsson J, Christiansen EH, Danielewicz M, Mario C Di, Doh JH, Erglis A, Erlinge D, Gerber RT, Going O, Gudmundsdottir I, Härle T, Hauer D, Hellig F, Indolfi C, Jakobsen L, Janssens L, Jensen J, Jeremias A, Kåregren A, Karlsson AC, Kharbanda RK, Khashaba A, Kikuta Y, Krackhardt F, Koo BK, Koul S, Laine M, Lehman SJ, Lindroos P, Malik IS, Maeng M, Matsuo H, Meuwissen M, Nam CW, Niccoli G, Nijjer SS, Olsson H, Olsson SE, Omerovic E, Panayi G, Petraco R, Piek JJ, Ribichini F, Samady H, Samuels B, Sandhall L, Sapontis J, Sen S, Seto AH, Sezer M, Sharp ASP, Shin ES, Singh J, Takashima H, Talwar S, Tanaka N, Tang K, Belle E Van, Royen N van, Varenhorst C, Vinhas H, Vrints CJ, Walters D, Yokoi H, Fröbert O, Patel MR, Serruys P, Davies JE, Götberg M. Safety of the Deferral of Coronary Revascularization on the Basis of Instantaneous

Wave-Free Ratio and Fractional Flow Reserve Measurements in Stable Coronary Artery Disease and Acute Coronary Syndromes. *JACC Cardiovasc Interv* 2018;**11**:1437–1449.

153. Götberg M, Berntorp K, Rylance R, Christiansen EH, Yndigegn T, Gudmundsdottir IJ, Koul S, Sandhall L, Danielewicz M, Jakobsen L, Olsson SE, Olsson H, Omerovic E, Calais F, Lindroos P, Maeng M, Venetsanos D, James SK, Kåregren A, Carlsson J, Jensen J, Karlsson AC, Erlinge D, Fröbert O. 5-Year Outcomes of PCI Guided by Measurement of Instantaneous Wave-Free Ratio Versus Fractional Flow Reserve. *J Am Coll Cardiol* 2022;**79**:965–974.
154. Berntorp K, Rylance R, Yndigegn T, Koul S, Fröbert O, Christiansen EH, Erlinge D, Götberg M. Clinical Outcome of Revascularization Deferral With Instantaneous Wave-Free Ratio and Fractional Flow Reserve: A 5-Year Follow-Up Substudy From the iFR-SWEDEHEART Trial. *J Am Heart Assoc* 2023;**12**:28423.
155. Yndigegn T, Koul S, Rylance R, Berntorp K, Mohammad MA, Omerovic E, Sarno G, Linder R, Fröbert O, Jensen J, Schiopu A, Erlinge D, Götberg M. Long-term Safety of Revascularization Deferral Based on Instantaneous Wave-Free Ratio or Fractional Flow Reserve. *J Soc Cardiovascular Angiog Interv* 2023;**2**:101046.
156. DEFINE-FLAIR: five-year follow-up. *Presented at EuroPCR 2023*.
157. Berry C, McClure JD, Oldroyd KG. Coronary revascularization guided by instantaneous wave-free ratio compared with fractional flow reserve: pooled 5-year mortality in the DEFINE-FLAIR and iFR-SWEDEHEART trials. *Eur Heart J* 2023;**44**:4388–4390.
158. Eftekhari A, Holck EN, Westra J, Olsen NT, Bruun NH, Jensen LO, Engstrøm T, Christiansen EH. Instantaneous wave free ratio vs. fractional flow reserve and 5-year mortality: iFR SWEDEHEART and DEFINE FLAIR. *Eur Heart J* 2023;**44**:4376–4384.
159. Lee JM, Hwang D, Park J, Tong Y, Koo BK. Physiologic mechanism of discordance between instantaneous wave-free ratio and fractional flow reserve: Insight from 13N-ammonium positron emission tomography. *Int J Cardiol* 2017;**243**:91–94.
160. Beijnink CWH, Thim T, Heijden DiJ Van Der, Klem I, Al-Lamee R, Vos JL, Koop Y, DIjkgraaf MGW, Beijk MAM, Kim RJ, Davies J, Raposo L, Baptista SB, Escaned J, Piek JJ, Maeng M, Royen N Van, Nijveldt R. Instantaneous wave-free ratio guided multivessel revascularisation during percutaneous coronary intervention for acute myocardial infarction: study protocol of the randomised controlled iMODERN trial. *BMJ Open* 2021;**11**: e044035
161. Patel MR, Jeremias A, Maehara A, Matsumura M, Zhang Z, Schneider J, Tang K, Talwar S, Marques K, Shammas NW, Gruberg L, Seto A, Samady H, Sharp ASP, Ali

ZA, Mintz G, Davies J, Stone GW. 1-Year Outcomes of Blinded Physiological Assessment of Residual Ischemia After Successful PCI: DEFINE PCI Trial. *JACC Cardiovasc Interv* 2022;**15**:52–61.

162. Bommel RJ van, Masdjedi K, Diletti R, Lemmert ME, Zandvoort L van, Wilschut J, Zijlstra F, Jaegere P de, Daemen J, Mieghem NM van. Routine Fractional Flow Reserve Measurement After Percutaneous Coronary Intervention. *Circ Cardiovasc Interv* 2019;**12**:e007428.
163. Collison D, Didagelos M, Aetesam-Ur-Rahman M, Copt S, McDade R, McCartney P, Ford TJ, McClure J, Lindsay M, Shaukat A, Rocchiccioli P, Brogan R, Watkins S, McEntegart M, Good R, Robertson K, O'Boyle P, Davie A, Khan A, Hood S, Eteiba H, Berry C, Oldroyd KG. Post-stenting fractional flow reserve vs coronary angiography for optimization of percutaneous coronary intervention (TARGET-FFR). *Eur Heart J* 2021;**42**:4656–4668.
164. Pijls NHJ, Bruyne B De, Bech GJW, Liistro F, Heyndrickx GR, Bonnier HJRM, Koolen JJ. Coronary pressure measurement to assess the hemodynamic significance of serial stenoses within one coronary artery: validation in humans. *Circulation* 2000;**102**:2371–2377.
165. Bruyne B De, Pijls NHJ, Heyndrickx GR, Hodeige D, Kirkeeide R, Gould KL. Pressure-derived fractional flow reserve to assess serial epicardial stenoses: theoretical basis and animal validation. *Circulation* 2000;**101**:1840–1847.
166. Nijjer SS, Sen S, Petraco R, Mayet J, Francis DP, Davies JER. The Instantaneous wave-Free Ratio (iFR) pullback: a novel innovation using baseline physiology to optimise coronary angioplasty in tandem lesions. *Cardiovasc Revasc Med* 2015;**16**:167–171.
167. Daniels D V., Van'T Veer M, Pijls NHJ, Horst A Van Der, Yong AS, Bruyne B De, Fearon WF. The impact of downstream coronary stenoses on fractional flow reserve assessment of intermediate left main disease. *JACC Cardiovasc Interv* 2012;**5**:1021–1025.
168. Munhoz D, Collet C, Mizukami T, Yong A, Leone AM, Eftekhari A, Ko B, Costa BR da, Berry C, Collison D, Perera D, Christiansen EH, Rivero F, Zimmermann FM, Ando H, Matsuo H, Nakayama M, Escaned J, Sonck J, Sakai K, Adjedj J, Desta L, Nunen LX van, West NEJ, Fournier S, Storozhenko T, Amano T, Engstrøm T, Johnson T, Shinke T, Biscaglia S, Fearon WF, Ali Z, Bruyne B De, Johnson NP. Rationale and design of the pullback pressure gradient (PPG) global registry. *Am Heart J* 2023;**265**:170–179.

169. Collet C, Sonck J, Vandeloo B, Mizukami T, Roosens B, Lochy S, Argacha JF, Schoors D, Colaiori I, Gioia G Di, Kodeboina M, Suzuki H, 't Veer M Van, Bartunek J, Barbato E, Cosyns B, Bruyne B De. Measurement of Hyperemic Pullback Pressure Gradients to Characterize Patterns of Coronary Atherosclerosis. *J Am Coll Cardiol* 2019;**74**:1772–1784.

170. Sonck J, Mizukami T, Johnson NP, Nagumo S, Gallinoro E, Candreva A, Mileva N, Munhoz D, Shinke T, Svanerud J, Barbato E, Bruyne B De, Collet C. Development, validation, and reproducibility of the pullback pressure gradient (PPG) derived from manual fractional flow reserve pullbacks. *Catheter Cardiovasc Interv* 2022;**99**:1518–1525.

171. Nijjer SS, Sen S, Petraco R, Escaned J, Echavarria-Pinto M, Broyd C, Al-Lamee R, Foin N, Foale RA, Malik IS, Mikhail GW, Sethi AS, Al-Bustami M, Kaprielian RR, Khan MA, Baker CS, Bellamy MF, Hughes AD, Mayet J, Francis DP, Mario C Di, Davies JER. Pre-Angioplasty Instantaneous Wave-Free Ratio Pullback Provides Virtual Intervention and Predicts Hemodynamic Outcome for Serial Lesions and Diffuse Coronary Artery Disease. *JACC Cardiovasc Interv* 2014;**7**:1386–1396.

172. Kikuta Y, Cook CM, Sharp ASP, Salinas P, Kawase Y, Shiono Y, Giavarini A, Nakayama M, Rosa S De, Sen S, Nijjer SS, Al-Lamee R, Petraco R, Malik IS, Mikhail GW, Kaprielian RR, Wijntjens GWM, Mori S, Hagikura A, Mates M, Mizuno A, Hellig F, Lee K, Janssens L, Horie K, Mohdnazri S, Herrera R, Krackhardt F, Yamawaki M, Davies J, Takebayashi H, Keeble T, Haruta S, Ribichini F, Indolfi C, Mayet J, Francis DP, Piek JJ, Mario C Di, Escaned J, Matsuo H, Davies JE. Pre-Angioplasty Instantaneous Wave-Free Ratio Pullback Predicts Hemodynamic Outcome In Humans With Coronary Artery Disease: Primary Results of the International Multicenter iFR GRADIENT Registry. *JACC Cardiovasc Interv* 2018;**11**:757–767.

173. Modi BN, Rahman H, Ryan M, Ellis H, Pavlidis A, Redwood S, Clapp B, Chowienczyk P, Perera D. Comparison of fractional flow reserve, instantaneous wave-free ratio and a novel technique for assessing coronary arteries with serial lesions. *EuroIntervention* 2020;**16**:577–583.

174. Cook CM, Ahmad Y, Shun-Shin MJ, Nijjer S, Petraco R, Al-Lamee R, Mayet J, Francis DP, Sen S, Davies JE. Quantification of the effect of pressure wire drift on the diagnostic performance of fractional flow reserve, instantaneous Wave-Free Ratio, and Whole-Cycle Pd/Pa. *Circ Cardiovasc Interv* 2016;**9**: e002988.

References

175. Dérimay F, Johnson NP, Zimmermann FM, Adjedj J, Witt N, Hennigan B, Koo BK, Barbato E, Esposito G, Trimarco B, Rioufol G, Park SJ, Baptista SB, Chrysant GS, Leone AM, Jeremias A, Berry C, Bruyne B De, Oldroyd KG, Pijls NHJ, Fearon WF. Predictive factors of discordance between the instantaneous wave-free ratio and fractional flow reserve. *Catheter Cardiovasc Interv* 2019;**94**:356–363.
176. Ahmad Y, Götberg M, Cook C, Howard JP, Malik I, Mikhail G, Frame A, Petraco R, Rajkumar C, Demir O, Iglesias JF, Bhindi R, Koul S, Hadjiloizou N, Gerber R, Ramrakha P, Ruparelia N, Sutaria N, Kanaganayagam G, Ariff B, Fertleman M, Anderson J, Chukwuemeka A, Francis D, Mayet J, Serruys P, Davies J, Sen S. Coronary Hemodynamics in Patients With Severe Aortic Stenosis and Coronary Artery Disease Undergoing Transcatheter Aortic Valve Replacement: Implications for Clinical Indices of Coronary Stenosis Severity. *JACC Cardiovasc Interv* 2018;**11**:2019–2031.
177. Pesarini G, Scarsini R, Zivelonghi C, Piccoli A, Gambaro A, Gottin L, Rossi A, Ferrero V, Vassanelli C, Ribichini F. Functional Assessment of Coronary Artery Disease in Patients Undergoing Transcatheter Aortic Valve Implantation: Influence of Pressure Overload on the Evaluation of Lesions Severity. *Circ Cardiovasc Interv* 2016;**9**: e004088-
178. Stundl A, Shamekhi J, Bernhardt S, Starke M, Al-Kassou B, Weber M, Sedaghat A, Treede H, Grube E, Nickenig G, Werner N, Sinning JM. Fractional flow reserve in patients with coronary artery disease undergoing TAVI: a prospective analysis. *Clin Res Cardiol* 2020;**109**:746–754.
179. Ahmad Y, Vendrik J, Eftekhari A, Howard JP, Cook C, Rajkumar C, Malik I, Mikhail G, Ruparelia N, Hadjiloizou N, Nijjer S, Al-Lamee R, Petraco R, Warisawa T, Wijntjens GWM, Koch KT, Hoef T Van De, Waard G De, Echavarria-Pinto M, Frame A, Sutaria N, Kanaganayagam G, Ariff B, Anderson J, Chukwuemeka A, Fertleman M, Koul S, Iglesias JF, Francis D, Mayet J, Serruys P, Davies J, Escaned J, Royen N Van, Götberg M, Juhl Terkelsen C, Høj Christiansen E, Piek JJ, Baan J, Sen S. Determining the Predominant Lesion in Patients With Severe Aortic Stenosis and Coronary Stenoses: A Multicenter Study Using Intracoronary Pressure and Flow. *Circ Cardiovasc Interv* 2019;**12**: e008263.
180. Gioia G Di, Bruyne B De, Pellicano M, Bartunek J, Colaiori I, Fiordelisi A, Canciello G, Xaplanteris P, Fournier S, Katbeh A, Franco D, Kodeboina M, Morisco C, Praet F Van, Casselman F, Degrieck I, Stockman B, Vanderheyden M, Barbato E. Fractional flow reserve in patients with reduced ejection fraction. *Eur Heart J* 2020;**41**:1665–1672.

181. Arashi H, Yamaguchi J, Ri T, Otsuki H, Nakao M, Kamishima K, Jujo K, Minami Y, Ogawa H, Hagiwara N. The impact of tissue Doppler index E/e' ratio on instantaneous wave-free ratio. *J Cardiol* 2018;**71**:237–243.

182. Petraco R, Park JJ, Sen S, Nijjer SS, Malik IS, Echavarriá-Pinto M, Asrress KN, Nam CW, Maciás E, Foale RA, Sethi A, Mikhail GW, Kaprielian R, Baker CS, Lefroy D, Bellamy M, Al-Bustami M, Khan MA, Gonzalo N, Hughes AD, Francis DP, Mayet J, Mario C Di, Redwood S, Escaned J, Koo BK, Davies JE. Hybrid iFR-FFR decision-making strategy: Implications for enhancing universal adoption of physiology-guided coronary revascularisation. *EuroIntervention* 2013;**8**:1157–1165.

183. Tebaldi M, Biscaglia S, Fineschi M, Musumeci G, Marchese A, Leone AM, Rossi ML, Stefanini G, Maione A, Menozzi A, Tarantino F, Lodolini V, Gallo F, Barbato E, Tarantini G, Campo G. Evolving Routine Standards in Invasive Hemodynamic Assessment of Coronary Stenosis: The Nationwide Italian SICI-GISE Cross-Sectional ERIS Study. *JACC Cardiovasc Interv* 2018;**11**:1482–1491.

184. Svanerud J, Ahn JM, Jeremias A, 'T Veer M Van, Gore A, Maehara A, Crowley A, Pijls NHJ, Bruyne B De, Johnson NP, Hennigan B, Watkins S, Berry C, Oldroyd KG, Park SJ, Ali ZA. Validation of a novel non-hyperaemic index of coronary artery stenosis severity: the Resting Full-cycle Ratio (VALIDATE RFR) study. *EuroIntervention* 2018;**14**:806–814.

185. Kumar G, Desai R, Gore A, Rahim H, Maehara A, Matsumura M, Kirtane A, Jeremias A, Ali Z. Real world validation of the nonhyperemic index of coronary artery stenosis severity-Resting full-cycle ratio-RE-VALIDATE. *Catheter Cardiovasc Interv* 2020;**96**:E53–E58.

186. Ligthart J, Masdjedi K, Witberg K, Mastik F, Zandvoort L Van, Lemmert ME, Wilschut J, Diletti R, Jaegere P De, Zijlstra F, Kardys I, Mieghem NM Van, Daemen J. Validation of resting diastolic pressure ratio calculated by a novel algorithm and its correlation with distal coronary artery pressure to aortic pressure, instantaneous wave-free ratio, and fractional flow reserve: The DPR study. *Circ Cardiovasc Interv* 2018;**11**: e006911.

187. Johnson NP, Li W, Chen X, Hennigan B, Watkins S, Berry C, Fearon WF, Oldroyd KG. Diastolic pressure ratio: new approach and validation vs. the instantaneous wave-free ratio. *Eur Heart J* 2019;**40**:2585.

188. Leone AM, Martin-Reyes R, Baptista SB, Amabile N, Raposo L, Franco Pelaez JA, Trani C, Cialdella P, Basile E, Zimbardo G, Burzotta F, Porto I, Aurigemma C, Rebuzzi AG, Faustino M, Niccoli G, Abreu PF, Slama MS, Spagnoli V, Arrieta MT, Amat Santos

IJ, La Torre Hernandez JM De, Palop RL, Crea F. The Multi-center Evaluation of the Accuracy of the Contrast MEdium INduced Pd/Pa RaTiO in Predicting FFR (MEMENTO-FFR) Study. *EuroIntervention* 2016;**12**:708–715.

189. Pijls NHJ, Bruyne B De, Smith L, Aarnoudse W, Barbato E, Bartunek J, Bech GJW, Vosse F Van De. Coronary Thermodilution to Assess Flow Reserve. *Circulation* 2002;**105**:2482–2486.
190. Everaars H, Waard GA de, Driessen RS, Danad I, Ven PM van de, Raijmakers PG, Lammertsma AA, Rossum AC van, Knaapen P, Royen N van. Doppler Flow Velocity and Thermodilution to Assess Coronary Flow Reserve: A Head-to-Head Comparison With [15O]H2O PET. *JACC Cardiovasc Interv* 2018;**11**:2044–2054.
191. Hoef TP Van De, Lavieren MA Van, Damman P, Delewi R, Piek MA, Chamuleau SAJ, Voskuil M, Henriques JPS, Koch KT, Winter RJ De, Spaan JAE, Siebes M, Tijssen JGP, Meuwissen M, Piek JJ. Physiological basis and long-term clinical outcome of discordance between fractional flow reserve and coronary flow velocity reserve in coronary stenoses of intermediate severity. *Circ Cardiovasc Interv* 2014;**7**:301–311.
192. Lee JM, Jung JH, Hwang D, Park J, Fan Y, Na SH, Doh JH, Nam CW, Shin ES, Koo BK. Coronary Flow Reserve and Microcirculatory Resistance in Patients With Intermediate Coronary Stenosis. *J Am Coll Cardiol* 2016;**67**:1158–1169.
193. Usui E, Murai T, Kanaji Y, Hoshino M, Yamaguchi M, Hada M, Hamaya R, Kanno Y, Lee T, Yonetsu T, Kakuta T. Clinical significance of concordance or discordance between fractional flow reserve and coronary flow reserve for coronary physiological indices, microvascular resistance, and prognosis after elective percutaneous coronary intervention. *EuroIntervention* 2018;**14**:798–805.
194. Fearon WF, Balsam LB, Farouque HMO, Robbins RC, Fitzgerald PJ, Yock PG, Yeung AC. Novel index for invasively assessing the coronary microcirculation. *Circulation* 2003;**107**:3129–3132.
195. Fearon WF, Kobayashi Y. Invasive Assessment of the Coronary Microvasculature: The Index of Microcirculatory Resistance. *Circ Cardiovasc Interv* 2017;**10**.
196. Tu S, Barbato E, Köszegi Z, Yang J, Sun Z, Holm NR, Tar B, Li Y, Rusinaru D, Wijns W, Reiber JHC. Fractional Flow Reserve Calculation From 3-Dimensional Quantitative Coronary Angiography and TIMI Frame Count: A Fast Computer Model to Quantify the Functional Significance of Moderately Obstructed Coronary Arteries. *JACC Cardiovasc Interv* 2014;**7**:768–777.

197. Westra J, Andersen BK, Campo G, Matsuo H, Koltowski L, Eftekhari A, Liu T, Serafino L Di, Girolamo D Di, Escaned J, Nef H, Naber C, Barbierato M, Tu S, Neghabat O, Madsen M, Tebaldi M, Tanigaki T, Kochman J, Somi S, Esposito G, Mercone G, Mejia-Renteria H, Ronco F, Bøtker HE, Wijns W, Christiansen EH, Holm NR. Diagnostic Performance of In-Procedure Angiography-Derived Quantitative Flow Reserve Compared to Pressure-Derived Fractional Flow Reserve: The FAVOR II Europe-Japan Study. *J Am Heart Assoc* 2018;**7**.

198. Xu B, Tu S, Qiao S, Qu X, Chen Y, Yang J, Guo L, Sun Z, Li Z, Tian F, Fang W, Chen J, Li W, Guan C, Holm NR, Wijns W, Hu S. Diagnostic Accuracy of Angiography-Based Quantitative Flow Ratio Measurements for Online Assessment of Coronary Stenosis. *J Am Coll Cardiol* 2017;**70**:3077–3087.

199. Puymirat E, Cayla G, Simon T, Steg PG, Montalescot G, Durand-Zaleski I, Bras A le, Gallet R, Khalife K, Morelle J-F, Motreff P, Lemesle G, Dillinger J-G, Lhermusier T, Silvain J, Roule V, Labèque J-N, Rangé G, Ducrocq G, Cottin Y, Blanchard D, Charles Nelson A, Bruyne B De, Chatellier G, Danchin N. Multivessel PCI Guided by FFR or Angiography for Myocardial Infarction. *New Eng J Med* 2021;**385**:297–308.

200. Andersen BK, Sejr-Hansen M, Westra J, Campo G, Efterkhari A, Tu S, Escaned J, Koltowski L, Stähli B, Erglis A, Jaruševičius G, Žiubrytė G, Råmunddal T, Liu T, Wijns W, Landmesser U, Maillard L, Matsuo H, Christiansen EH, Holm NR. Quantitative flow ratio versus fractional flow reserve for guiding percutaneous coronary intervention: design and rationale of the randomised FAVOR III Europe Japan trial. *EuroIntervention* 2023;**18**:E1357–E1364.

201. Witberg G, Bruyne B De, Fearon WF, Achenbach S, Engstrom T, Matsuo H, Kornowski R. Diagnostic Performance of Angiogram-Derived Fractional Flow Reserve: A Pooled Analysis of 5 Prospective Cohort Studies. *Cardiovasc Interv* 2020;**13**:488–497.

202. Fearon WF, Achenbach S, Engstrom T, Assali A, Shlofmitz R, Jeremias A, Fournier S, Kirtane AJ, Kornowski R, Greenberg G, Jubeh R, Kolansky DM, McAndrew T, Dressler O, Maehara A, Matsumura M, Leon MB, Bruyne B De, FAST-FFR Study Investigators. Accuracy of Fractional Flow Reserve Derived From Coronary Angiography. *Circulation* 2019;**139**:477–484.

203. Masdjedi K, Zandvoort LJC van, Balbi MM, Gijsen FJH, Ligthart JMR, Rutten MCM, Lemmert ME, Wilschut J, Diletti R, Jaegere P de, Zijlstra F, Mieghem NM van, Daemen J. Validation of a three-dimensional quantitative coronary angiography-based software

to calculate fractional flow reserve: the FAST study. *EuroIntervention* 2020;**16**:591–599.

204. Masdjedi K, Tanaka N, Belle E Van, Porouchani S, Linke A, Woitek FJ, Bartorelli AL, Ali ZA, Dekker WK den, Wilschut J, Diletti R, Zijlstra F, Boersma E, Mieghem NM Van, Spitzer E, Daemen J. Vessel fractional flow reserve (vFFR) for the assessment of stenosis severity: the FAST II study. *EuroIntervention* 2022;**17**:1498.
205. Study Details | Comparison of Vessel-FFR Versus FFR in Intermediate Coronary Stenoses | ClinicalTrials.govhttps://www.clinicaltrials.gov/study/NCT03497637 (5 December 2023)
206. Scoccia A, Byrne RA, Banning AP, Landmesser U, Belle E Van, Amat-Santos IJ, Sabaté M, Tijssen JGP, Spitzer E, Daemen J. Fractional flow reserve or 3D-quantitative-coronary-angiography based vessel-FFR guided revascularization. Rationale and study design of the prospective randomized fast III trial. *Am Heart J* 2023;**260**:1–8.
207. Wang L, Travieso A, Hoeven N van der, Leeuwen MAH van, Janssens G, Mejía-Rentería H, Jerónimo A, Gonzalo N, Nijveldt R, Royen N van, Escaned J. Improved Nonculprit Stenosis Assessment in Patients With ST-Segment Elevation Myocardial Infarction Using Quantitative Flow Ratio. *Cardiovascular Interventions* 2023;**16**:1828–1830.
208. Ono M, Serruys PW, Patel MR, Escaned J, Akasaka T, Lavieren MA van, Haase C, Grass M, Kogame N, Hara H, Kawashima H, Wykrzykowska JJ, Piek JJ, Garg S, O'Leary N, Inderbitzen B, Onuma Y. A prospective multicenter validation study for a novel angiography-derived physiological assessment software: Rationale and design of the radiographic imaging validation and evaluation for Angio-iFR (ReVEAL iFR) study. *Am Heart J* 2021;**239**:19–26.
209. Turing AM. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 1937;**s2-42**:230–265.
210. Turing AM. Computing machinery and intelligence. *Mind* 1950;**LIX**:433–460.
211. Marvao A de, Dawes TJ, Howard JP, O'Regan DP. Artificial intelligence and the cardiologist: what you need to know for 2020. *Heart* 2020:heartjnl-2019-316033.
212. Sardar P, Abbott JD, Kundu A, Aronow HD, Granada JF, Giri J. Impact of Artificial Intelligence on Interventional Cardiology: From Decision-Making Aid to Advanced Interventional Procedure Assistance. *JACC Cardiovasc Interv* 2019;**12**:1293–1303.

213. Kuwaiti A Al, Nazer K, Al-Reedy A, Al-Shehri S, Al-Muhanna A, Subbarayalu AV, Muhanna D Al, Al-Muhanna FA. A Review of the Role of Artificial Intelligence in Healthcare. *J Pers Med* 2023;**13**.
214. Topol E. *Deep medicine: how artificial intelligence can make healthcare human again*. 2019.
215. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;**323**:533–536.
216. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention Is All You Need. *Adv Neural Inf Process Syst* 2017;**2017-December**:5999–6009.
217. Makimoto H, Kohro T. Adopting artificial intelligence in cardiovascular medicine: a scoping review. *Hypertension Research 2023* 2023:1–15.
218. Schepart A, Burton A, Durkin L, Fuller A, Charap E, Bhambri R, Ahmad FS. Artificial intelligence–enabled tools in cardiovascular medicine: A survey of current use, perceptions, and challenges. *Cardiovasc Digit Health J* 2023;**4**:101–110.
219. Attia ZI, Harmon DM, Behr ER, Friedman PA. Application of artificial intelligence to the electrocardiogram. *Eur Heart J* 2021;**42**:4717–4730.
220. Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol* 2021;**18**:465–478.
221. Kwon JM, Jo YY, Lee SY, Kim KH. Artificial intelligence using electrocardiography: Strengths and pitfalls. *Eur Heart J* 2021;**42**:2896–2898.
222. Chang SN, Tseng YH, Chen JJ, Chiu FC, Tsai CF, Hwang JJ, Wang YC, Tsai CT. An artificial intelligence-enabled ECG algorithm for identifying ventricular premature contraction during sinus rhythm. *Eur J Med Res* 2022;**27**:289.
223. Valente Silva B, Marques J, Nobre Menezes M, Oliveira AL, Pinto FJ. Artificial intelligence-based diagnosis of acute pulmonary embolism: Development of a machine learning model using 12-lead electrocardiogram. *Rev Port. Card* 2023;**42**:643–651.
224. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. *J Am Coll Cardiol* 2016;**68**:2287–2295.
225. Asch FM, Poilvert N, Abraham T, Jankowski M, Cleve J, Adams M, Romano N, Hong H, Mor-Avi V, Martin RP, Lang RM. Automated Echocardiographic Quantification of Left Ventricular Ejection Fraction Without Volume Measurements Using a Machine

Learning Algorithm Mimicking a Human Expert. *Circ Cardiovasc Imaging* 2019;**12**:e009303.

226. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, Heidenreich PA, Harrington RA, Liang DH, Ashley EA, Zou JY. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* 2020;**580**:252–256.
227. He B, Kwan AC, Cho JH, Yuan N, Pollick C, Shiota T, Ebinger J, Bello NA, Wei J, Josan K, Duffy G, Jujjavarapu M, Siegel R, Cheng S, Zou JY, Ouyang D. Blinded, randomized trial of sonographer versus AI cardiac function assessment. *Nature 2023 616:7957* 2023;**616**:520–524.
228. Moghaddasi H, Nourian S. Automatic assessment of mitral regurgitation severity based on extensive textural features on 2D echocardiography videos. *Comput Biol Med* 2016;**73**:47–55.
229. Playford D, Bordin E, Mohamad R, Stewart S, Strange G. Enhanced Diagnosis of Severe Aortic Stenosis Using Artificial Intelligence: A Proof-of-Concept Study of 530,871 Echocardiograms. *JACC Cardiovasc Imaging* 2020;**13**:1087–1090.
230. Martin SS, Assen M van, Rapaka S, Hudson HT, Fischer AM, Varga-Szemes A, Sahbaee P, Schwemmer C, Gulsun MA, Cimen S, Sharma P, Vogl TJ, Schoepf UJ. Evaluation of a Deep Learning–Based Automated CT Coronary Artery Calcium Scoring Algorithm. *JACC Cardiovasc Imaging* 2020;**13**:524–526.
231. Zreik M, Hamersvelt RW Van, Wolterink JM, Leiner T, Viergever MA, Išgum I. A Recurrent CNN for Automatic Detection and Classification of Coronary Artery Plaque and Stenosis in Coronary CT Angiography. *IEEE Trans Med Imaging* 2019;**38**:1588–1598.
232. Gaur S, Øvrehus KA, Dey D, Leipsic J, Bøtker HE, Jensen JM, Narula J, Ahmadi A, Achenbach S, Ko BS, Christiansen EH, Kaltoft AK, Berman DS, Bezerra H, Lassen JF, Nørgaard BL. Coronary plaque quantification and fractional flow reserve by coronary computed tomography angiography identify ischaemia-causing lesions. *Eur Heart J* 2016;**37**:1220–1227.
233. Dey D, Gaur S, Ovrehus KA, Slomka PJ, Betancur J, Goeller M, Hell MM, Gransar H, Berman DS, Achenbach S, Botker HE, Jensen JM, Lassen JF, Norgaard BL. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: A multicentre study. *Eur Radiol* 2018;**28**:2655–2664.
234. Hamersvelt RW van, Zreik M, Voskuil M, Viergever MA, Išgum I, Leiner T. Deep learning analysis of left ventricular myocardium in CT angiographic intermediate-

degree coronary stenosis improves the diagnostic accuracy for identification of functionally significant stenosis. *Eur Radiol* 2019;**29**:2350–2359.

235. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal* 2017;**35**:159–171.
236. Fahmy AS, Rausch J, Neisius U, Chan RH, Maron MS, Appelbaum E, Menze B, Nezafat R. Automated Cardiac MR Scar Quantification in Hypertrophic Cardiomyopathy Using Deep Convolutional Neural Networks. *JACC Cardiovasc Imaging* 2018;**11**:1917–1918.
237. Baessler B, Luecke C, Lurz J, Klingel K, Das A, Roeder M Von, Waha-Thiele S De, Besler C, Rommel KP, Maintz D, Gutberlet M, Thiele H, Lurz P. Cardiac MRI and texture analysis of myocardial T1 and T2 maps in myocarditis with acute versus chronic symptoms of heart failure. *Radiology* 2019;**292**:608–617.
238. Inan OT, Baran Pouyan M, Javaid AQ, Dowling S, Etemadi M, Dorier A, Heller JA, Bicen AO, Roy S, Marco T De, Klein L. Novel Wearable Seismocardiography and Machine Learning Algorithms Can Assess Clinical Status of Heart Failure Patients. *Circ Heart Fail* 2018;**11**:e004313.
239. Ahmad T, Lund LH, Rao P, Ghosh R, Warier P, Vaccaro B, Dahlström U, O'Connor CM, Michael Felker G, Desai NR. Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients. *J Am Heart Assoc* 2018;**7**.
240. Koulaouzidis G, Jadczyk T, Iakovidis DK, Koulaouzidis A, Bisnaire M, Charisopoulou D. Artificial Intelligence in Cardiology—A Narrative Review of Current Status. *J Clin Med* 2022;**11**.
241. Ali ZA, Landmesser U, Maehara A, Matsumura M, Shlofmitz RA, Guagliumi G, Price MJ, Hill JM, Akasaka T, Prati F, Bezerra HG, Wijns W, Leistner D, Canova P, Alfonso F, Fabbiocchi F, Dogan O, McGreevy RJ, McNutt RW, Nie H, Buccola J, West NEJ, Stone GW. Optical Coherence Tomography–Guided versus Angiography-Guided PCI. *New Eng J Med* 2023;**389**:1466–1476.
242. Subhan S, Malik J, Haq A ul, Qadeer MS, Zaidi SMJ, Orooj F, Zaman H, Mehmoodi A, Majeedi U. Role of Artificial Intelligence and Machine Learning in Interventional Cardiology. *Curr Probl Cardiol* 2023;**48**:101698.
243. Yang S, Kweon J, Kim Y-H. Major Vessel Segmentation on X-ray Coronary Angiography using Deep Networks with a Novel Penalty Loss Function. *Proceedings of Machine Learning Research-Under Review*.

References

244. Jun TJ, Kweon J, Kim YH, Kim D. T-Net: Nested encoder–decoder architecture for the main vessel segmentation in coronary angiography. *Neural Networks* 2020;**128**:216–233.
245. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. Unet++: A nested u-net architecture for medical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2018;**11045 LNCS**:3–11.
246. Xian Z, Wang X, Yan S, Yang D, Chen J, Peng C. Main Coronary Vessel Segmentation Using Deep Learning in Smart Medical. *Math Probl Eng* 2020;**2020**.
247. Du T, Xie L, Zhang H, Liu X, Wang X, Chen D, Xu Y, Sun Z, Zhou W, Song L, Guan C, Lansky A, Xu B. Training and validation of a deep learning architecture for the automatic analysis of coronary angiography. *EuroIntervention* 2021;**17**:32–40.
248. Yang S, Kweon J, Roh J-H, Lee J-H, Kang H, Park L-J, Kim DJ, Yang H, Hur J, Kang D-Y, Lee PH, Ahn J-M, Kang S-J, Park D-W, Lee S-W, Kim Y-H, Lee CW, Park S-W, Park S-J. Deep learning segmentation of major vessels in X-ray coronary angiography. *Scientific Reports 2019 9:1* 2019;**9**:1–11.
249. Wang L, Liang D, Yin X, Qiu J, Yang Z, Xing J, Dong J, Ma Z. Coronary artery segmentation in angiographic videos utilizing spatial-temporal information. *BMC Medical Imaging 2020 20:1* 2020;**20**:1–10.
250. Liang D, Qiu J, Wang L, Yin X, Xing J, Yang Z, Dong J, Ma Z. Coronary angiography video segmentation method for assisting cardiovascular disease interventional treatment. *BMC Medical Imaging 2020 20:1* 2020;**20**:1–8.
251. Gao Z, Wang L, Soroushmehr R, Wood A, Gryak J, Nallamothu B, Najarian K. Vessel segmentation for X-ray coronary angiography using ensemble methods with deep learning and filter-based features. *BMC Med Imaging* 2022;**22**:1–17.
252. Liu X, Wang X, Chen D, Zhang H. Automatic Quantitative Coronary Analysis Based on Deep Learning. *Applied Sciences 2023, Vol 13, Page 2975* 2023;**13**:2975.
253. Avram R, Olgin JE, Ahmed Z, Verreault-Julien L, Wan A, Barrios J, Abreau S, Wan D, Gonzalez JE, Tardif J-C, So DY, Soni K, Tison GH. CathAI: fully automated coronary angiography interpretation and stenosis estimation. *npj Digital Medicine 2023 6:1* 2023;**6**:1–12.
254. Freitas SA, Zeiser FA, Costa CA Da, O. Ramos G De. DeepCADD: A Deep Learning Architecture for Automatic Detection of Coronary Artery Disease. *Proceedings of the International Joint Conference on Neural Networks* 2022;**2022-July**.

255. Roguin A, Abu Dogosh A, Feld Y, Konigstein M, Lerman A, Koifman E. Early Feasibility of Automated Artificial Intelligence Angiography Based Fractional Flow Reserve Estimation. *American Journal of Cardiology* 2021;**139**:8–14.

256. Ben-Assa E, Abu Salman A, Cafri C, Roguin A, Hellou E, Koifman E, Feld Y, Lev E, Sheinman G, Harari E, Abu Dogosh A, Beyar R, Garcia-Garcia HM, Davies J, Ben-Yehuda O. Performance of a novel artificial intelligence software developed to derive coronary fractional flow reserve values from diagnostic angiograms. *Coron Artery Dis* 2023;**34**:533–541.

257. Cho H, Lee JG, Kang SJ, Kim WJ, Choi SY, Ko J, Min HS, Choi GH, Kang DY, Lee PH, Ahn JM, Park DW, Lee SW, Kim YH, Lee CW, Park SW, Park SJ. Angiography-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. *J Am Heart Assoc* 2019;**8**: e011685.

258. Ploscaru V, Popa-Fotea NM, Calmac L, Itu LM, Mihai C, Bataila V, Dragoescu B, Puiu A, Cojocaru C, Costin MA, Scafa-Udriste A. Artificial intelligence and cloud based platform for fully automated PCI guidance from coronary angiography-study protocol. *PLoS One* 2022;**17**:e0274296.

259. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiade M, Bonow RO, Huang CC, Deo RC. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 2015;**131**:269–279.

260. Somashekhar SP, Sepúlveda M-J, Norden AD, Rauthan A, Arun K, Patil P, Ethadka RY, Kumar RC. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. *J Clin Oncol* 2017;**35**:8527–8527.

261. Fialho I, Beringuilho M, Madeira D, Ferreira JB, Faria D, Ferreira H, Roque D, Santos MB, Morais C, Gil V, Augusto JB. Acute myocardial infarction on YouTube – is it all fake news? *Rev Port. Card (English edition)* 2021;**40**:815–825.

262. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Med* 2019**;25**: 65–69

263. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N, Lukaschuk E, Sanghvi MM, Zemrak F, Fung K, Paiva JM, Carapella V, Kim YJ, Suzuki H, Kainz B, Matthews PM, Petersen SE, Piechnik SK, Neubauer S, Glocker B, Rueckert D. Automated cardiovascular magnetic resonance image analysis with fully

convolutional networks 08 Information and Computing Sciences 0801 Artificial Intelligence and Image Processing. *Journal of Cardiovasc Magn Reson* 2018;**20**:65.

264. Ali W Ben, Pesaranghader A, Avram R, Overtchouk P, Perrin N, Laffite S, Cartier R, Ibrahim R, Modine T, Hussin JG. Implementing Machine Learning in Interventional Cardiology: The Benefits Are Worth the Trouble. *Front Cardiovasc Med* 2021;**0**:1775.
265. Gonzalez & Woods, Digital Image Processing, 4th Edition | Pearsonhttps://www.pearson.com/us/higher-education/program/Gonzalez-Digital-Image-Processing-4th-Edition/PGM241219.html (2 September 2021)
266. Zhu X, Cheng Z, Wang S, Chen X, Lu G. Coronary angiography image segmentation based on PSPNet. *Comput Methods Programs Biomed* 2021;**200**:105897.
267. Sim HW, Ananthakrishna R, Chan SP, Low AF, Lee CH, Chan MY, Tay EL, Loh PH, Chan KH, Tan HC LJP. Treatment of Very Small De Novo Coronary Artery Disease With 2.0 mm Drug-Coated Balloons Showed 1-Year Clinical Outcome Comparable With 2.0 mm Drug-Eluting Stents. *J Invasive Cardiol. 2018 Jul;30(7):256-261.*
268. LC van der H, MM K, PW D, AR S, M H, MM L, GC L, MG S, CJ D, C von B. Small-vessel treatment with contemporary newer-generation drug-eluting coronary stents in all-comers: Insights from 2-year DUTCH PEERS (TWENTE II) randomized trial. *Am Heart J* 2016;**176**:28–35.
269. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2015;**9351**:234–241.
270. Silva JL, Menezes MN, Rodrigues T, Silva B, Pinto FJ, Oliveira AL. Encoder-Decoder Architectures for Clinically Relevant Coronary Artery Segmentation. *arXiv:210611447 [eessIV]* 2021.
271. Tan M, Le Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *36th International Conference on Machine Learning, ICML 2019* 2019;**2019-June**:10691–10700.
272. Sudre CH, Li W, Vercauteren T, Ourselin S, Cardoso MJ. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. *Deep learning in medical image analysis and multimodal learning for clinical decision support : Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, held in conjunction with MICCAI 2017 Quebec City, QC,.* 2017;**2017**:240.

273. Nobre Menezes M, Lourenço-Silva J, Silva B, Rodrigues O, Francisco ARG, Carrilho Ferreira P, Oliveira AL, Pinto FJ. Development of deep learning segmentation models for coronary X-ray angiography: Quality assessment by a new global segmentation score and comparison with human performance. *Rev Port. Card* 2022;**41**:1011–1021.
274. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat* 2001;**29**:1189–1232.
275. Zhou ZH, Feng J. Deep Forest. *Natl Sci Rev* 2017;**6**:74–86.
276. Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal Loss for Dense Object Detection. *Proceedings of the IEEE International Conference on Computer Vision* 2017;**2017-October**:2999–3007.
277. M'hiri F, Duong L, Desrosiers C, Dahdah N, Miró J, Cheriet M. Automatic evaluation of vessel diameter variation from 2D X-ray angiography. *Int J Comput Assist Radiol Surg* 2017;**12**:1867–1876.
278. Pyxaras SA, Harmel EK. Artificial intelligence for the assessment of coronary artery disease- The dawn of a new era? *Int J Cardiol* 2021;**343**:3–4.
279. Nobre Menezes M, Silva JL, Silva B, Rodrigues T, Guerreiro C, Guedes JP, Santos MO, Oliveira AL, Pinto FJ. Coronary X-ray angiography segmentation using Artificial Intelligence: a multicentric validation study of a deep learning model. *Int J Cardiovasc Imaging 2023* 2023:1–12.
280. Raposo L, Gonçalves M, Roque D, Gonçalves PA, Magno P, Brito J, Leal S, Madeira S, Santos M, Teles RC, e Abreu PF, Almeida M, Morais C, Mendes M, Baptista SB. Adoption and patterns of use of invasive physiological assessment of coronary artery disease in a large cohort of 40821 real-world procedures over a 12-year period. *Rev Port. Card* 2021;**40**:771–781.
281. Koo B-K, Hu X, Kang J, Zhang J, Jiang J, Hahn J-Y, Nam C-W, Doh J-H, Lee B-K, Kim W, Huang J, Jiang F, Zhou H, Chen P, Tang L, Jiang W, Chen X, He W, Ahn S-G, Yoon M-H, Kim U, Lee J-M, Hwang D, Ki Y-J, Shin E-S, Kim H-S, Tahk S-J, Wang J. Fractional Flow Reserve or Intravascular Ultrasonography to Guide PCI. *N Engl J Med* 2022;**387**:779–789.
282. Nobre Menezes M, Silva B, Silva JL, Rodrigues T, Marques JS, Guerreiro C, Guedes JP, Oliveira-Santos M, Oliveira AL, Pinto FJ. Segmentation of X-ray coronary angiography with an artificial intelligence deep learning model: Impact in operator visual assessment of coronary stenosis severity. *Catheter Cardiovasc Interv* 2023;**102**:631–640.

References

283. Cook CM, Warisawa T, Howard JP, Keeble TR, Iglesias JF, Schampaert E, Bhindi R, Ambrosia A, Matsuo H, Nishina H, Kikuta Y, Shiono Y, Nakayama M, Doi S, Takai M, Goto S, Yakuta Y, Karube K, Akashi YJ, Clesham GJ, Kelly PA, Davies JR, Karamasis G V., Kawase Y, Robinson NM, Sharp ASP, Escaned J, Davies JE. Algorithmic Versus Expert Human Interpretation of Instantaneous Wave-Free Ratio Coronary Pressure-Wire Pull Back Data. *JACC Cardiovasc Interv* 2019;**12**:1315–1324.
284. Zhao Q, Li C, Chu M, Gutiérrez-Chico JL, Tu S. Angiography-based coronary flow reserve: The feasibility of automatic computation by artificial intelligence. *Cardiol J* 2023;**30**:369–378.

APPENDIX

List of publications and Facsimile of published articles included in the Thesis

Revista Portuguesa de Cardiologia 41 (2022) 1011–1021

Portuguese Society of CARDIOLOGY

Revista Portuguesa de **Cardiologia**

Portuguese Journal of ***Cardiology***

www.revportcardiol.org

ORIGINAL ARTICLE

Development of deep learning segmentation models for coronary X-ray angiography: Quality assessment by a new global segmentation score and comparison with human performance

Miguel Nobre Menezes[a,b,*], **João Lourenço-Silva**[c], **Beatriz Silva**[a,b], **Tiago Rodrigues**[a,b], **Ana Rita G. Francisco**[a,b], **Pedro Carrilho Ferreira**[a,b], **Arlindo L. Oliveira**[c], **Fausto J. Pinto**[a,b]

[a] *Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal*
[b] *Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Lisboa, Portugal*
[c] *INESC-ID/Instituto Superior Técnico, University of Lisbon, Portugal*

Received 27 January 2022; accepted 3 April 2022
Available online 26 May 2022

KEYWORDS
Deep learning;
Artificial intelligence;
Machine learning;
Coronary angiography;
Coronary artery disease;
Percutaneous coronary intervention

Abstract
Introduction and objectives: Although automatic artificial intelligence (AI) coronary angiography (CAG) segmentation is arguably the first step toward future clinical application, it is underexplored. We aimed to (1) develop AI models for CAG segmentation and (2) assess the results using similarity scores and a set of criteria defined by expert physicians.
Methods: Patients undergoing CAG were randomly selected in a retrospective study at a single center. Per incidence, an ideal frame was segmented, forming a baseline human dataset (BH), used for training a baseline AI model (BAI). Enhanced human segmentation (EH) was created by combining the best of both. An enhanced AI model (EAI) was trained using the EH. Results were assessed by experts using 11 weighted criteria, combined into a Global Segmentation Score (GSS: 0–100 points). Generalized Dice Score (GDS) and Dice Similarity Coefficient (DSC) were also used for AI models assessment.
Results: 1664 processed images were generated. GSS for BH, EH, BAI and EAI were 96.9+/-5.7; 98.9+/-3.1; 86.1+/-10.1 and 90+/-7.6, respectively (95% confidence interval, $p<0.001$ for both paired and global differences). The GDS for the BAI and EAI was 0.9234±0.0361 and 0.9348±0.0284, respectively. The DSC for the coronary tree was 0.8904±0.0464 and 0.9134±0.0410 for the BAI and EAI, respectively. The EAI outperformed the BAI in all coronary segmentation tasks, but performed less well in some catheter segmentation tasks.

* Corresponding author.
E-mail address: mnmenezes.gm@gmail.com (M. Nobre Menezes).

https://doi.org/10.1016/j.repc.2022.04.001

Appendix

Conclusions: We successfully developed AI models capable of CAG segmentation, with good performance as assessed by all scores.

PALAVRAS-CHAVE
Aprendizagem profunda;
Inteligência artifical;
Aprendizagem de máquinas;
Coronariografia;
Doença coronária;
Intervenção coronária percutânea

Desenvolvimento de modelos de *deep learning* para segmentação de coronariografias: aferição de qualidade por um novo modelo de segmentação global e comparação com desempenho humano

Resumo
Introdução e objetivos: A segmentação automática de coronariografia (CRG) por inteligência artificial (IA) encontra-se pouco explorada na literatura médica. Os objetivos do presente estudo são (1) desenvolver modelos de IA para segmentação de CRG e (2) aferir os resultados por *scores* de similaridade e critérios definidos por peritos.
Métodos: Doentes submetidos a CRG foram retrospetivamente selecionados aleatoriamente num centro. Por incidência, segmentou-se um *frame* ideal, formando uma segmentação humana basal (HB), usada para treinar um modelo de IA basal (IAB). Da combinação de ambos acrescentou-se uma segmentação humana aperfeiçoada (HA), utilizada para treinar um modelo de IA aperfeiçoado (IAA). Os resultados foram aferidos com 11 critérios balanceados definidos por peritos, combinados num *Score de Segmentação Global* (SSC – 0–100 pontos). O *Score de Dice Generalizado* (SDG) e *Score de Dice de Similaridade* (SDS) aplicaram-se aos modelos de IA.
Resultados: Geraram-se 1664 imagens processadas. Os SCC para a HB, HA, IAB e IAA foram 96,9+/-5,7; 98,9+/-3,1; 86,1+/-10,1 e 90+/-7,6, respetivamente (IC 95%, p<0,001 - diferenças globais e emparelhadas). O SDG para o IAB e IAA foi 0,9234±0,0361 e 0,9348±0,0284, respetivamente. O SDS foi 0,8904±0,0464 e 0,9134±0,0410 para o IAB e IAA, respetivamente. O IAA exibiu superior desempenho ao IAB para as todas tarefas de segmentação coronária, mas não para todas as de cateter.
Conclusões: Desenvolvemos modelos de IA de segmentação automática de CRG, com bom desempenho de acordo com aferição por todos os *scores*.

Introduction

Artificial intelligence (AI) has shown great potential in medicine, in applications such as predictive data analysis,[1] decision making support[2] or even medical education/awareness improvement,[3] and especially in image analysis. Several publications have demonstrated impressive results with regards to electrocardiogram,[4] echocardiography,[5,6] or magnetic resonance imaging.[7,8]

The use of AI in Interventional Cardiology (IC) is, however, still a vastly underexplored field. Its application to coronary angiography (CAG) has been explored in very few medical or biology publication.[9–12] There are, nonetheless, many possibilities,[13] ranging from automatic anatomical identification, stenosis analysis, lesion subset characterization and perhaps even physiological index derivation. Regardless of the task, arguably the first step in applying AI to CAG is separating and identifying relevant information – the coronary tree – from non-relevant information (bones, other structures). This task is called segmentation.[14]

In this paper, we explore the development of AI models for automatic coronary artery segmentation from CAG, and assess the results from a clinical perspective, using a new set of criteria and score clinically defined by a panel of Interventional Cardiologists.

Methods

Dataset selection

Inclusion criteria

We retrospectively and randomly included patients who had undergone CAG and invasive physiology assessment (fractional flow reserve and/or other indexes) during the procedure at a single center (tertiary university hospital).

These patients have at least intermediate lesions in one or more vessels. Around one third usually undergo revascularization due to the severity of their disease.[15,16] Therefore, a dataset focusing on these patients comprises a wide spectrum of obstructive coronary artery disease in a relatively balanced way.

Revista Portuguesa de Cardiologia 41 (2022) 1011–1021

Exclusion criteria

We excluded cases where any of the following applied:

1) Major occluded vessels (acute or chronic)
2) Poor image quality
3) Less than two orthogonal views in the left coronary artery (LCA) - one caudal and one cranial - or absence of at least one left oblique (LAO) view - either cranial or simple - in the right coronary artery (RCA)
4) Patients with previous cardiac surgery, cardiac devices or other sources of potential artifact.

Image selection

A single best frame was selected for each diagnostic angulation incidence in each patient.

Dataset size

The dataset size was the result of a trade-off between two opposing criteria: dimension large enough for successful training of a deep convolutional neural network, estimated from published data[9,12,17,18] vs. expected time required to complete the annotation. We estimated the latter based on a short period of annotation testing prior to formal dataset creation. The trade-off pointed to a training set size of roughly 400.

We then randomly and consecutively selected patients until a total of at least 400 annotated images were obtained.

Baseline annotation process

Baseline human dataset images were annotated by two senior Cardiology Fellows (TR/BS) previously trained in CAG interpretation, under the supervision of an Interventional Cardiologist (MNM), who also annotated. Images were periodically reviewed and perfected by all three. This meant that any initial heterogeneity between annotators was corrected by consensus. The small size of the team was aimed at reducing heterogeneity, as we noticed during the preparatory phase that some operators tended to annotate too much (Supplementary figure 1, Appendix A), while others did the opposite (Supplementary figure 2, Appendix A).

Both the catheter (labeled red) and the coronaries (labeled white) were to be segmented.

The coronary tree was to be fully segmented up to branches of approximately 2 mm in caliper at their origin (as the vessel became smaller, it was to be segmented until discernible), using the catheter as reference (without formal measurements – eyeball appreciation was used). There were several reasons for this: (1) when performing percutaneous coronary intervention, vessels <2 mm are usually approached conservatively, as the risk of target lesion failure increases significantly[19,20]; (2) human annotation is cumbersome – segmenting every single vessel would increase the risk of errors significantly; (3) including very small vessels might increase the chances of artifacts from bone or other structures when training and applying AI models.

Baseline artificial intelligence model training

We performed segmentation using an encoder-decoder fully convolutional neural network based on the U-Net,[21] commonly used in medical image segmentation. As the name suggests, these neural networks are composed of an encoder, responsible for extracting image features, and a decoder, which processes those features to produce segmentation masks. To derive the best approach for this task, we conducted a comparative study of encoder and decoder architectures, which resulted in the proposal of the EfficientUNet++, a computationally efficient and high-performing decoder architecture[22] that, in this work, we combine with an EfficientNet-B5 encoder[23] (Figure 1).

To ensure fair evaluation, it was necessary to guarantee that each model was tested on data that it had not seen during training. Therefore, the dataset was split, at the patient level, into 13 subsets of approximately 32 angiograms each. Each subset segmentation was performed using a neural network trained exclusively on the remaining data. This enabled the assessment of the segmentation results for the entire cohort, as the usual splitting into a training and testing dataset would have yielded a much smaller group of images for result assessment.

The training hyperparameters, including the number of training epochs and the learning rate decay schedule, were set on the first train-test split, using one of the 12 training data subsets for validation. The selected values were then used on every other train-test split, and to train the model on the whole training set of the first split.

Enhanced human model

The results of the baseline AI training were reviewed by the annotating team, without any formal grading, which would be performed subsequently (see below). For each image, both human and AI segmentation were compared with the original. Each annotation was then perfected using a mixture of the best of baseline human segmentation and baseline AI, with additional de novo manual segmentation as needed.

Enhanced artificial intelligence model

The neural network architecture and training procedure were identical for both the baseline and enhanced AI model (Figure 1). The sole difference was the dataset. The baseline AI model was trained using the baseline human annotations, whereas the enhanced AI model was trained using enhanced human annotations.

Figure 2 outlines the development stages.

Performance assessment

Non-medical metrics

AI models were assessed using the Dice Similarity Coefficient (DSC) and Generalized Dice Score (GDS), measures of the overlap between segmentations. Given two segmentations, the DSC has a value between 0: no overlap and 1: total overlap, corresponding to the ratio between the area of their intersection and the sum of their areas. GDS[24] is a weighted

Appendix

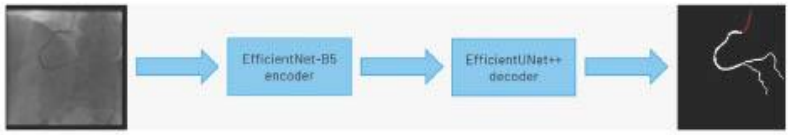


Figure 1 Segmentation model composed of an EfficientNet-B5 encoder and an EfficientUNet++ decoder.

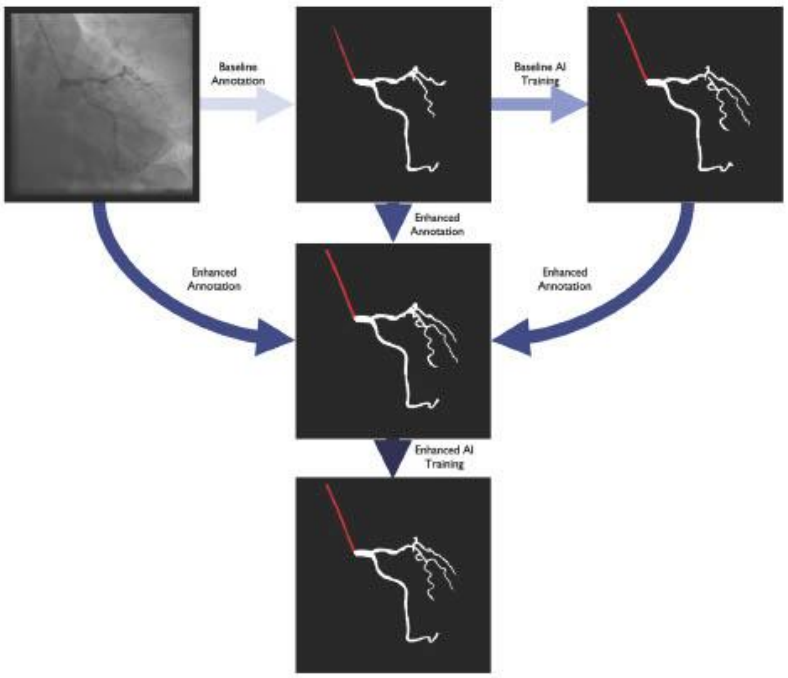


Figure 2 Annotation and training process.

sum of each class's DSC that attributes the same importance to all classes, regardless of their frequency. While DSC and GDS alone do not reflect clinical usefulness, they are helpful and entirely objective metrics that enable a simple comparison between models.

Clinical performance criteria

The DSC objectively assesses model performance. However, it does not provide a medically meaningful impression of whether segmentation is appropriate. Also, because the DCS can only be calculated based on previously annotated images, it cannot be applied to new, unannotated datasets in the future. To overcome these limitations, we created a set of criteria to assess performance as interpreted by expert physicians.

The following 11 criteria are as objectively defined as possible and were analyzed for each image. Each was independently met or not. A ''perfect'' example is shown in Figure 3. Supplementary Figures 3 to 13 (Appendix A) show error examples for each.

1) Catheter segmentation:
   a. **Main segmentation**: The distal part of the catheter (i.e. the closest discernible portion to the coronary artery in the ascending aorta) is correctly segmented and labeled (supplementary figure 3, Appendix A). If minor gaps are present, this criterion should be scored as met.
   b. **Gaps** (minor) are absent (supplementary figure 4, Appendix A).
   c. **Catheter thickness** is accurate, by visual appreciation (supplementary figure 5, Appendix A).
   d. **Location**: if parts of the catheter far from the coronary ostia (ascending and/or descending aorta) are segmented, there are no major gaps or artifacts (supplementary 6, Appendix A).
2) Vessel segmentation:
   a. **Main vessels** are correctly segmented and labeled. For the RCA, this includes the segments from the ostium to the crux (supplementary figure 7, Appendix A). For the LCA, this includes the segments from the left main ostium to the visually discernible distal segments of the left anterior descending or the circumflex (or most

Revista Portuguesa de Cardiologia 41 (2022) 1011–1021

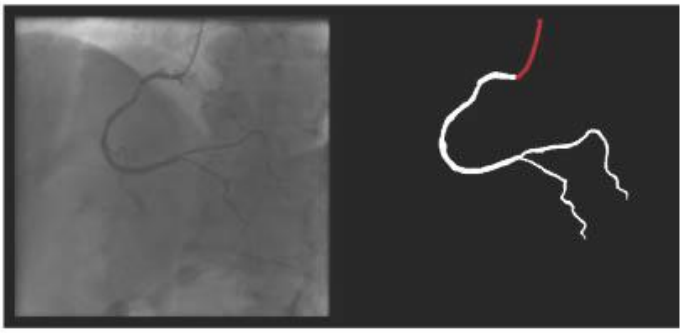


Figure 3 A segmentation case fulfilling all 11 criteria.

important obtuse marginal branch), depending on incidence. Branches are excluded from this criterion. If minor gaps are present, this criterion should be scored as met.

b. **Branch segmentation**: branches with a luminal diameter of at least approximately 2 mm (using the catheter size as reference) are correctly segmented and labeled (supplementary figure 8, Appendix A). Size is estimated by visual appreciation. If minor gaps are present, this criterion should be scored as met.
c. **Main vessel gaps** (minor) are absent (supplementary figure 9, Appendix A).
d. **Branch gaps** (minor) are absent (supplementary figure 10, Appendix A).
e. Catheter to artery **transition**: correct labeling of the catheter tip vs. coronary artery origin (supplementary figure 11, Appendix A).

3) Artifacts

a. **Coronary**: No non-coronary structures are labeled as part of the coronary tree (supplementary figure 12, Appendix A).
b. **Catheter**: No non-catheter structures are incorrectly labeled as part of the catheter (supplementary figure 13, Appendix A).

The criteria for these two artifacts are not applicable to the small catheter-artery transition area.

To provide an objective assessment, these criteria were scored by a panel of three Interventional Cardiologists (MNM, ARF, PCF), of whom two (ARF, PCF) took no part in any stage of the annotation/training process. Discrepancies were solved by agreement. All images were graded across all groups: baseline human segmentation, enhanced human segmentation, baseline AI and enhanced AI. During the grading process, the image group was blinded.

Lastly, because the abovementioned criteria are not equally important, a Global Segmentation Score (GSS, 1.5 to 100 points) was devised, taking into account the relevance of each criterion as defined by the three experts (Table 1). The panel was also asked to select which of the two AI models was preferred for each image, regardless of the final score.

Statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. To assess the association between qualitative (categorical) variables the Chi-Square test was used. To assess differences in quantitative variables we used

Table 1 scoring metrics for application of the Global Segmentation Score.

Criteria	Catheter vs. Coronary Relative Weight	Individual Criteria Relative Weight	Points
Main vessel segmentation	70%	40%	28.0
Main vessel gaps		10%	7.0
Catheter to artery transition		15%	10.5
Branch segmentation		20%	14.0
BranchGaps		5%	3.5
Coronary artifacts		10%	7.0
Catheter segmentation	30%	40%	12.0
Catheter gaps		10%	3.0
Catheter artifacts		15%	4.5
Catheter location		5%	1.5
Catheter thickness		30%	9.0
Total			100

Appendix

Table 2 Baseline clinical characteristics of patients from whom images were analyzed.

Factor	N+/-SD or N(%)
Age	67+/-11
Sex (male)	54 (78%)
Hypertension	56 (81.2%)
Diabetes mellitus	27 (39.1%)
Dyslipidemia	39 (56.5%)
Smoker (past or present)	26 (37.7%)
Chronic coronary syndromes	50 (72.5%)
Acute coronary syndrome	19 (27.5%)
Revascularization during/after CAG	21 (30.4%)

CAG: coronary angiography.

Table 3 Generalized Dice Score and class-wise Dice Similarity Coefficient obtained by the baseline and enhanced AI models. Results presented as mean ± standard deviation.

	BAI	EAI
GDS	0.9234±0.0361	0.9348±0.0284
Artery DSC	0.8904±0.0464	0.9134±0.0410
Catheter DSC	0.7526±0.1998	0.7975±0.1836

BAI: baseline AI model; GDS: Generalized Dice Score; EAI; enhanced AI model.

the Mann-Whitney test (two independent groups) or the Kruskal-Wallis test (multiple independent groups). A $p<0.05$ was used for statistical significance, except for multiple group comparisons, where we used a $p<0.01$. IBM SPSS Statistics 27 was used for statistical analysis.

Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local ethics committee.

Results

Baseline dataset

We included 416 images from 69 patients (Table 2). With two human and two AI datasets, 1664 processed images were generated.

Performance assessment

Non-medical metrics

Results are outlined in Table 3. These scores indicate that enhanced AI was generally superior to baseline AI. Segmentation performance was good and consistent across arteries, as indicated by the high mean and low standard deviation of the DSC. For the catheter, performance was lower and much less consistent.

Clinical performance

Overall performance – individual criteria assessment (Supplementary Table 1, Appendix A).

Coronary segmentation

The main vessels were correctly segmented in almost all cases across groups. Minor gaps occurred rarely in the baseline human segmentation and both AI models, although there was a small but non-significant improvement with the enhanced AI vs. baseline AI.

Branch segmentation was also correct almost always in all groups, albeit less so than main vessel segmentation. There was a small, yet significant, improvement with the enhanced AI vs. baseline AI.

Minor branch gaps were quite common, revealing very significant differences between AI and human models. While enhanced AI performed numerically better than baseline AI, it still produced small gaps in nearly two thirds of cases.

Coronary artifacts were very uncommon in human annotations and were usually minor imperfections in catheter/coronary crossovers. They were common and usually minor in both AI models, although there was a very significant improvement with the enhanced AI vs. baseline AI (14.4% vs. 25.7%).

Catheter/artery transition

Baseline human segmentation failed in 12% of cases and enhanced human segmentation missed 3.8%. Baseline AI produced a higher error rate (19.7%), but enhanced AI was numerically more often correct than baseline human segmentation, sometimes correctly identifying the transition where humans failed (Figure 4).

Catheter segmentation

Baseline human segmentation produced thickness imperfections (usually mildly engorged catheter) in 13.9% of cases, but otherwise, segmentation was almost always correct regarding other criteria. Baseline AI produced low error rates in main body segmentation. However, artifacts, usually quite minor and in the vicinity of coronary segments, occurred very frequently (41.1%). Another common error was catheter thickness (36.3%), often resulting in an overestimation of catheter size.

Enhanced human segmentation significantly improved on thickness issues, although imperfections persisted in 6.2% of cases.

Enhanced AI produced better results than the baseline AI model for catheter thickness (correct in 96.4%), also surpassing both human models (although the difference was not statistically significant when compared to the enhanced human segmentation). However, the performance of the enhanced AI otherwise decreased in all other criteria, especially regarding minor gaps, which became much more common (3.1% in the baseline AI model to 23.3%). Even main body segmentation was significantly affected, although successful in the vast majority of cases (86.5%). Despite this,

Revista Portuguesa de Cardiologia 41 (2022) 1011–1021

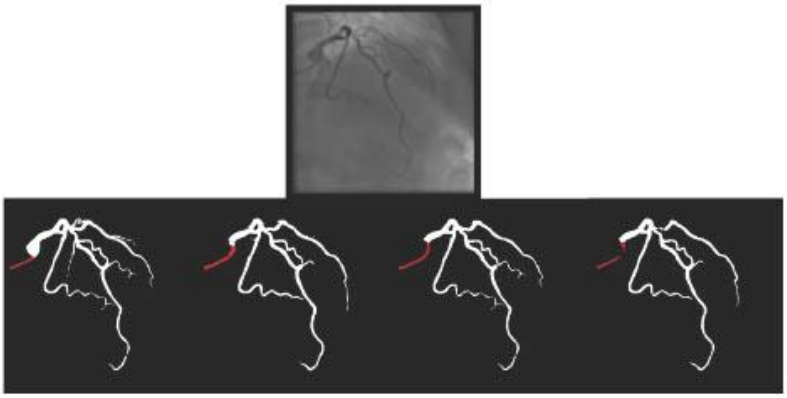


Figure 4 (left to right): The first human segmentation incorrectly labels contrast backflow as coronary. The baseline AI model improves on the human segmentation but is still not perfect. The enhanced human model segments the transition perfectly. The enhanced AI model is hampered in catheter segmentation but identifies the transition correctly.

Table 4 Performance by group according to Global Segmentation Score (significance at p<0.05 for paired differences and p<0.01 for multiple comparisons).

GSS	Group				p-value						
	BH	EH	BAI	EAI	Between all*	BH vs EH**	BAI vs EAI**	BH vs BAI**	EH vs EAI**	BH vs EAI**	EH vs BAI**
Mean+/- SD	96.9+/- 5.7	98.9+/- 3.1	86.1+/- 10.1	90+/-7.6	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Median (IQR)	100 (9)	100(0)	87.5 (9)	92 (9.5)							

BAI: baseline AI model; BH: baseline human model; EAI: enhanced AI model; EH: enhanced human model; GSS: Global Segmentation Score; IQR: interquartile range; SD: standard deviation.
* Kruskal-Wallis Test.
** Mann-Whitney Test.

in most failures catheter identification was still possible, as major gaps often occurred distally in areas of contrast backflow. There was a slight numerical worsening in artifact and location issues in enhanced AI vs. baseline AI.

Overall performance – Global Segmentation Score assessment and expert preference (Table 4)

Human models outperformed AI models. Enhanced models surpassed baseline models. The difference was statistically significant for all comparisons. GSS was very high for both AI models; the enhanced AI reached an average of 90 points.

With regards to expert preference, the enhanced AI model was preferred in 300 (72%) cases, the baseline AI model in 100 (24%) and in 16 (4%) cases no AI model was preferred.

Performance according to coronary artery – individual criteria assessment (Supplementary Table 2, Appendix A)

There was a trend toward better performance in the RCA, both regarding human and AI groups. The most notable and statistically significant differences occurred in catheter transition (regarding both AI models and the baseline human segmentation) and catheter segmentation (both AI models performed better in the RCA). Branch gaps were quite less frequent in the RCA with the enhanced AI model. Other differences, even if statistically significant, were very small.

Performance by coronary artery – Global Segmentation Score assessment (Supplementary Table 3, Appendix A)

All models scored very high for both arteries. There were very minor statistically significant differences for the baseline AI model only, favoring RCA segmentation.

Considering expert preference:

- RCA: Enhanced AI was preferred in 109 (68.6%) cases, the baseline AI was preferred in 43 (27%) and in 7 (4.4%) cases no AI model was preferred.
- LCA: Enhanced AI was preferred in 191 (74.3%) cases, the baseline AI was preferred in 57 (22.2%) and in 9 (3.5%) cases no AI was preferred.

Appendix

Performance according to angulation incidence – individual criteria assessment (Supplementary Tables 4 and 5, Appendix A)

Given the large amount of data, there being no significant differences in the vast majority of cases and for the sake of readability, only statistically significant differences are shown in the tables. Overall, the impact of incidences on model performance was limited, and affected almost exclusively the AI models.

Performance according to angulation incidence – Global Segmentation Score assessment (Supplementary Tables 6 and 7, Appendix A)

Differences were minor and only statistically significant for human performance in less common incidences (PA views for the LCA and PA cranial for the RCA).

Discussion

Overall considerations

Baseline human segmentation was generally correct. Catheter/coronary transition and catheter thickness errors were the most common. Poor individualization due to contrast backflow, catheter curves and human fatigue all likely contributed.

Enhanced human segmentation was nearly perfect. Mild transition issues remained, highlighting the difficulty of the task. As this model was actually a combination of the best of baseline human segmentation and baseline AI, it also demonstrates how AI can help improve human performance. Even these slight human imperfections highlight the need for rigorous quality control during and after the final results, rather than assuming human annotation is a ''perfect'' ground truth. This an inherent limitation to the annotation of medical images, as the sheer amount of cumbersome work is error prone.

Baseline AI performed CAG segmentation successfully yet was affected by the same two issues of the baseline human segmentation – transition and catheter thickness. The effort to correct these when developing the enhanced AI was fruitful in the case of transition but produced mixed results for catheter thickness. Impact on transition performance was impressive, as, at times, the enhanced AI even achieved correct assessments where humans failed (Figure 4). However, it seems the gain in catheter thickness accuracy was offset by losses in other catheter segmentation tasks. Lastly, every aspect of coronary segmentation improved in the enhanced AI, which performed better than baseline AI. The differences between the two AI models also highlight how relatively small differences in the ground truth can impact relevantly on AI training.

It may seem surprising that catheter segmentation was less successful than coronary segmentation. However, while intuitively one may think that catheter segmentation is an easier task and therefore the results would have been better for this task, from a machine learning perspective that is not the case. In particular, segmentation performance is highly dependent on the frequency of each class. Rarer classes, or ones that occupy smaller areas, are interpreted by the model as being less likely to appear. Furthermore, during training, the lower the number of pixels belonging to a particular class, the lower the penalty for segmenting that class incorrectly. Even though we used a loss function designed to mitigate this phenomenon, the poorer segmentation of less common classes (the catheter, in this case) is still evident in the results.

Right coronary artery segmentation was easier than LCA, however the differences were quite small and there were fewer than expected, considering its greater anatomical simplicity. Angulations also had a relatively small impact both on human and AI performance and small observed differences may be attributed to specific issues that are more common in certain incidences: contrast backflow (less problematic in PA or RAO caudal); coronary/catheter crossovers (such as spider or extreme RAO cranial – Figure 5); proximity of bone (such as RCA LAO views); smaller samples of some incidences, such as PA cranial; uncommon catheter pathways, such as the femoral approach, which sometimes produces a central vertical outline.

Globally, both AI models achieved a very high DSC, with higher performance in artery segmentation than in catheter segmentation, supporting the results of qualitative clinical assessment. When factors are weighed up based on their perceived relevance – as assessed by GSS – both performed very well. The enhanced AI scored an average of 90 points, meaning it provided 90% of what experts deemed most relevant when viewing a CAG. By all measures, the enhanced AI was the better model. However, the fact that differences between the two AI models were not large and that the enhanced AI was preferred in most, but not all cases, highlights the difficulty in improving an already good performance.

Other studies with artificial intelligence applied to coronary angiography segmentation/interpretation

Few studies regarding coronary artery segmentation based on AI technologies have been published in medical/biology journals to date. Yang et al.[12] successfully developed AI models capable of segmenting CAG. Their dataset was larger (3302 images/2042 patients) and was also annotated by two expert physicians. Different incidences were also used. They also focused exclusively on segmenting specific segments of major vessels with at least mild (>30%) stenotic lesions. Neither the branches nor the catheter were segmented, leading to a much simpler problem than the one addressed in this article.

Two other works[9,10] from the same baseline dataset, also developed AI-based CAG segmentation. Their dataset was also larger (4904 images from 170 videos). However, the annotations were performed by medical students and no details are provided regarding patient subset, target vessel or incidence.

Very recently, Du et al.[11] published the results of a broad study. They focused on two tasks: CAG segmentation and special lesion morphology identification (calcium, thrombus, among others). For the former task, which overlaps with ours, they used a very large dataset of 13 373 images

Revista Portuguesa de Cardiologia 41 (2022) 1011–1021

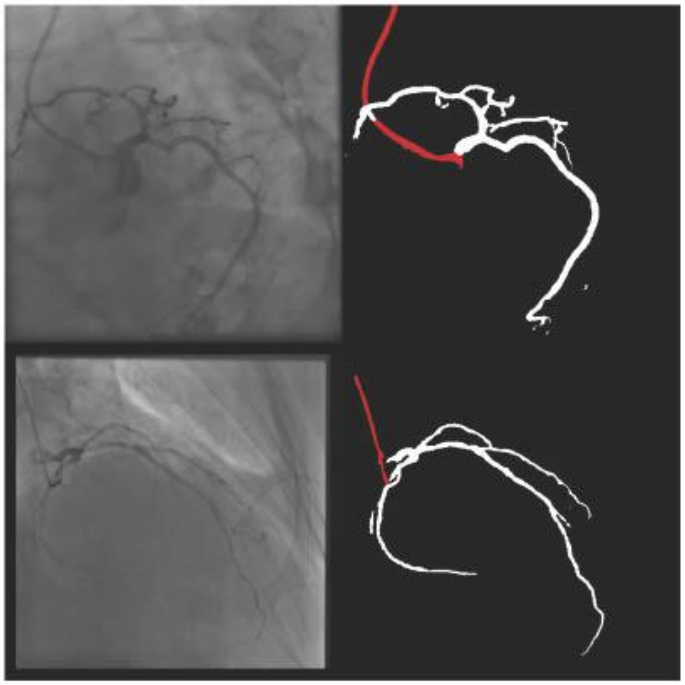


Figure 5 Crossovers in spider (above) and extreme RAO cranial (below) views generating artifacts.

distributed across ten incidences (six LCA and four RCA), annotated by ten qualified analysts. This was an all-comers study, rather than focusing on patient subsets. They too annotated catheter/arteries and additionally marked different coronary segments. Their model is impressive as judged by the presented images, as they even distinguished between contrast backflow, catheter and coronary. However, they did not specify the exact criteria for segmenting the coronary tree and their exact metrics make it difficult to assess exactly how their models performed in detail regarding segmentation.

While all the abovementioned groups have worked with datasets larger than ours, our study has several unique features: (1) there was medical rationale for vessel size segmentation; (2) results were assessed from a set of criteria defined by experts, capturing the quality of the segmentation from an Interventional Cardiologist's eyes; (3) human annotations were also graded, rather than assuming a perfect human ground truth; (4) specific segmentation tasks were appraised individually, enabling insights into strengths and weaknesses of AI and human models alike; (5) results were also considered globally with the GSS, by factoring the relevance of each criterion, enabling a broad, simple appreciation of the results. Furthermore, the ability to perform high-quality segmentation in a system trained using less data provides relevant evidence that more advanced AI systems can be effectively applied even in situations where the available data are limited.

Limitations

This is a single center retrospective dataset, involving a single image per projection and a smaller sample size than some previously published manuscripts. The images come from the same angiography devices (Siemens Artis) and thus we have not yet tested our models on images obtained from other equipment or image settings.

We have not yet conducted formal assessment on how well the models perform in segmenting specific degrees of stenosis severity. Our models are also yet to be tested for specific vessel disease types (calcium, thrombus), clinical settings (chronic total occlusion, ST-elevation myocardial infarction).

We have not yet assessed the performance of AI models on an external validation cohort. There are several reasons for this. We aimed to compare AI and human results in detail first and assess the exact performance of AI models for each segmentation task. A validation dataset would comprise a new set of images, which would not undergo human segmentation, thus impeding comparison with human performance. Also, validation implies that a metric be available for comparing results. Because the Dice methods require a ground truth human annotation for comparison, and the GSS was developed and applied for the first time for this paper, we felt a suitable metric was not yet available for performing validation prior to the current analysis. In addition, AI models are continuously and dynamically improving.

Appendix

As we are currently working on further testing and enhancing current AI models (view Future direction and implications section below), we felt performing external validation at this stage was premature.

The exclusion of cardiac devices/cardiac surgery and other foreign objects renders our models not yet applicable to such cases. We did not, however, exclude cases with previously implanted stents.

Lastly, focusing specifically on patients undergoing invasive physiology assessment may have created bias, limiting a broader application of the models to other patient subsets.

We are currently working to address all these issues in future research.

Future direction and implications

Coronary angiography segmentation in itself is not end objective but rather an essential milestone for developing AI systems capable of CAG analysis and interpretation. These results should, therefore, be regarded as a first step, rather than a final deployment tool. While not yet ready for immediate clinical application, the results of both AI models are already relevant, providing a framework that can be built upon in the future.

Further steps include testing the models for stenosed segments, which will be critical for clinical application. In the future, we aim to test our models with a validation cohort using new angiograms. Sub-segmentation, automatic anatomical identification and physiology are also areas for future research.

We will also strengthen the capabilities of our models further by broadening our training base to other patient and lesion subsets, focusing on particular issues where there is still room for improvement, as identified by our uniquely detailed analysis.

Our results also provide insight into which human tasks are most challenging, which may be of use to other researchers.

Global Segmentation Score is the first of its kind for assessing the quality of segmentations in CAG. By providing a reasonably objective and quantitative clinical measurement, it can be used as a benchmark for comparing and validating results across research groups.

Lastly, while conventional segmentation software does exist, it is not without limitations, and only by developing AI systems can we compare and improve both in the future. The potential implications of AI for Interventional Cardiology are immense, and we envisage a catherization lab of the future where all of these insights render the human eye more objective, thus improving patient care.

Conclusions

We successfully developed two AI models capable of good quality automatic CAG segmentation, as assessed by GDS, DSC and the GSS. From an expert's perspective, the latter and its individual criteria provided a feasible, reasonably objective and quantifiable way of assessing the results.

The enhanced AI model outperformed the baseline AI model in coronary segmentation tasks as well as globally. With regards to catheter segmentation tasks, the enhanced AI model improved on the task of catheter thickness, but performed less well in other catheter segmentation tasks. Both human segmentations were superior to both AI models, but only the enhanced human segmentation, built by combining the best of baseline human segmentation and baseline AI, achieved a near perfect GSS.

These results provide a relevant framework for building upon, potentially leading to future clinical application.

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgments

MNM was responsible for conceptualization, data gathering, processing and analysis, interpretating the results and paper drafting. JLS was responsible for technical and AI tasks, data and image processing, model implementation and training. TR and BS were responsible for data gathering, processing and analysis. ARF and PCF were responsible for data analysis and results interpretation. ALO was responsible for supervising the work of JLS, and participated directly in the same tasks. FJP was responsible for supervising the work of MNM, and participated directly in the same tasks.

All authors revised the paper critically for important intellectual content, gave final approval for its publication and agreed to be accountable for all respects of its accuracy and integrity.

Appendix A. Supplementary material

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.repc.2022.04.001.

References

1. Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. Circulation. 2015;131:269–79.
2. Somashekhar SP, Sepúlveda M-J, Norden AD, et al. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. J Clin Oncol. 2017;35 suppl.:8527.
3. Fialho I, Beringuilho M, Madeira D, et al. Acute myocardial infarction on YouTube – is it all fake news? Rev Port Cardiol (English Ed). 2021;40:815–25.
4. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med. 2019:65–9.
5. Narula S, Shameer K, Salem Omar AM, et al. Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. J Am Coll Cardiol. 2016;68:2287–95.
6. Asch FM, Poilvert N, Abraham T, et al. Automated echocardiographic quantification of left ventricular ejection fraction without volume measurements using a machine learning algorithm mimicking a human expert. Circ Cardiovasc Imaging. 2019;12:e009303.
7. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the

heart from cardiac cine magnetic resonance. Med Image Anal. 2017;35:159–71.
8. Bai W, Sinclair M, Tarroni G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks 08 Information and Computing Sciences 0801 Artificial Intelligence and Image Processing. J Cardiovasc Magn Reson. 2018;20:65.
9. Wang L, Liang D, Yin X, et al. Coronary artery segmentation in angiographic videos utilizing spatial-temporal information. BMC Med Imaging. 2020;20:1–10.
10. Liang D, Qiu J, Wang L, et al. Coronary angiography video segmentation method for assisting cardiovascular disease interventional treatment. BMC Med Imaging. 2020;20:1–8.
11. Du T, Xie L, Zhang H, et al. Training and validation of a deep learning architecture for the automatic analysis of coronary angiography. EuroIntervention. 2021;17:32–40.
12. Yang S, Kweon J, Roh JH, et al. Deep learning segmentation of major vessels in X-ray coronary angiography. Sci Rep. 2019;9:1–11.
13. Ben Ali W, Pesaranghader A, Avram R, et al. Implementing machine learning in interventional cardiology: the benefits are worth the trouble. Front Cardiovasc Med. 2021:1775.
14. Gonzalez & Woods, Digital image processing, 4th ed. | Pearson [Internet]. Available from: https://www.pearson.com/us/higher-education/program/Gonzalez-Digital-Image-Processing-4th-Edition/PGM241219.html [cited 02.09.21].
15. Davies JE, Sen S, Dehbi H-M, et al. Use of the instantaneous wave-free ratio or fractional flow reserve in PCI. N Engl J Med. 2017;376:1824–34.
16. Götberg M, Christiansen EH, Gudmundsdottir IJ, et al. Instantaneous wave-free ratio versus fractional flow reserve to guide PCI. N Engl J Med. 2017;376:1813–23.
17. Zhu X, Cheng Z, Wang S, et al. Coronary angiography image segmentation based on PSPNet. Comput Methods Programs Biomed. 2021;200:105897.
18. Jun TJ, Kweon J, Kim YH, et al. et: Nested encoder–decoder architecture for the main vessel segmentation in coronary angiography. Neural Networks. 2020;128:216–33.
19. Sim HW, Ananthakrishna R, Chan SP, et al. Treatment of very small de novo coronary artery disease with 2.0 mm drug-coated balloons showed 1-year clinical outcome comparable with 2.0 mm drug-eluting stents [Internet]. J Invasive Cardiol. 2018;30:256–61. Available from: https://www.hmpgloballearningnetwork.com/site/jic/articles/treatment-very-small-de-novo-coronary-artery-disease-20-mm-drug-coated-balloons-showed-1 [cited 29.08.21].
20. van der LCH, MM K, PW D, et al. Small-vessel treatment with contemporary newer-generation drug-eluting coronary stents in all-comers: Insights from 2-year DUTCH PEERS (TWENTE II) randomized trial. Am Heart J. 2016;176:28–35.
21. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2015;9351:234–41.
22. Silva JL, Nobre Menezes M, Rodrigues T, et al. Encoder-decoder architectures for clinically relevant coronary artery segmentation; 2021 Jun 21, arXiv:210611447 [eessIV].
23. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. In: 36th Int Conf Mach Learn ICML 2019. 2019 June. p. 10691–700.
24. Sudre CH, Li W, Vercauteren T, et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Deep Learn Med image Anal multimodal Learn Clin Decis Support Third Int Work DLMIA 2017, 7th Int Work ML-CDS 2017, held conjunction with MICCAI 2017 Quebec City, QC, vol. 2017. Europe PMC Funders; 2017. p. 240.

The International Journal of Cardiovascular Imaging (2023) 39:1385–1396
https://doi.org/10.1007/s10554-023-02839-5

ORIGINAL PAPER

Coronary X-ray angiography segmentation using Artificial Intelligence: a multicentric validation study of a deep learning model

Miguel Nobre Menezes[1,2] · João Lourenço Silva[3] · Beatriz Silva[1,2] · Tiago Rodrigues[1,2] · Cláudio Guerreiro[4] · João Pedro Guedes[5] · Manuel Oliveira Santos[6,7] · Arlindo L. Oliveira[3] · Fausto J. Pinto[1,2]

Received: 20 November 2022 / Accepted: 18 March 2023 / Published online: 7 April 2023

Abstract

Introduction We previously developed an artificial intelligence (AI) model for automatic coronary angiography (CAG) segmentation, using deep learning. To validate this approach, the model was applied to a new dataset and results are reported.

Methods Retrospective selection of patients undergoing CAG and percutaneous coronary intervention or invasive physiology assessment over a one month period from four centers. A single frame was selected from images containing a lesion with a 50–99% stenosis (visual estimation). Automatic Quantitative Coronary Analysis (QCA) was performed with a validated software. Images were then segmented by the AI model. Lesion diameters, area overlap [based on true positive (TP) and true negative (TN) pixels] and a global segmentation score (GSS – 0 -100 points) - previously developed and published - were measured.

Results 123 regions of interest from 117 images across 90 patients were included. There were no significant differences between lesion diameter, percentage diameter stenosis and distal border diameter between the original/segmented images. There was a statistically significant albeit minor difference [0,19 mm (0,09–0,28)] regarding proximal border diameter. Overlap accuracy ((TP + TN)/(TP + TN + FP + FN)), sensitivity (TP / (TP + FN)) and Dice Score (2TP / (2TP + FN + FP)) between original/segmented images was 99,9%, 95,1% and 94,8%, respectively. The GSS was 92 (87–96), similar to the previously obtained value in the training dataset.

Conclusion the AI model was capable of accurate CAG segmentation across multiple performance metrics, when applied to a multicentric validation dataset. This paves the way for future research on its clinical uses.

Keywords Deep learning · Artificial Intelligence · Machine learning · Coronary angiography · Coronary artery disease · Percutaneous coronary intervention.

Abbreviations

GSS	Global Segmentation Score
AI	Artificial Intelligence
CAG	Coronary Angiography
STEMI	ST-elevation myocardial infarction
CTO	Chronic Total Occlusion

✉ Miguel Nobre Menezes
mnmenezes.gm@gmail.com

[1] Structural and Coronary Heart Disease Unit, Faculdade de Medicina, Cardiovascular Center of the University of Lisbon, Universidade de Lisboa (CCUL@RISE), Av Prof. Egas Moniz, Lisboa 1649-028, Portugal

[2] Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Av Prof. Egas Moniz, Lisboa 1649-028, Portugal

[3] INESC-ID / Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

[4] Centro Hospitalar de Vila Nova de Gaia, Porto, Portugal

[5] Unidade de Hemodinâmica e Cardiologia de Intervenção, Serviço de Cardiologia, Centro Hospitalar Universitário do Algarve, Hospital de Faro, Faro, Portugal

[6] Unidade de Intervenção Cardiovascular, Serviço de Cardiologia do Centro Hospitalar e Universitário de Coimbra, Praceta Professor Mota Pinto, Coimbra 3004-561, Portugal

[7] Faculdade de Medicina da Universidade de Coimbra, R. Larga 2, Coimbra 3000-370, Portugal

Appendix

Introduction

The application of artificial intelligence (AI) to coronary angiography (CAG) has only been ascertained in very few medical/biology publications [1–4]. While the possibilities of such an approach are vast, the first step is arguably to produce accurate segmentation of CAGs, i.e., clearly identifying the coronary tree while excluding other structures.

We have previously published the first results of deep learning models capable of good quality CAG segmentation [5]. In this paper, we aim to validate the results, by applying the model to a new, previously unseen, dataset of coronary angiographies from multiple centers. A well-known validated software was used as reference for segments with non-occlusive lesions, where detailed measurements were undertaken, while also applying the previously described Global Segmentation Score for broad assessment of segmentation quality [5].

Methods

Participating centers and equipment

Four centers from across Portugal participated in this study. Images were acquired in Siemens Axiom Artis and Philips Azureon equipment.

Inclusion criteria

Retrospective selection of consecutive patients who had undergone CAG and percutaneous coronary intervention (PCI) and/or underwent invasive physiology assessment (Fractional Flow Reserve and/or other indexes), within a 1-month period of 2022, regardless of clinical context (i.e. both acute and chronic coronary syndrome). This ensures the model was tested in a real-world context where revascularization was either being considered or performed, thereby excluding a population with normal or near-normal coronary arteries.

Exclusion criteria

We excluded cases where any of the following applied:

1) Patients with previous cardiac surgery, cardiac devices or other sources of potential artifact.
2) Absence of coronary lesions 50–99% stenosis by visual estimation (i.e. single-vessel ST-elevation myocardial infarction – STEMI - or chronic total occlusions – CTO alone).
3) Poor image quality.
4) Unclear individualization of lesion outline with no overlapping vessels.
5) Unsuccessful automatic measurements with validated software (details below).
6) Unsuccessful software extraction and superimposition of lesion markers on segmented image (details below).

Image selection

For each selected lesion, a single end-diastolic frame with clear outline definition of the vessel and target lesion was selected. More than one segment per patient and/or image could be used. With an original training dataset of 416 images as previously published [5], we aimed to have a validation dataset of at least 100 images.

Brief description of previous work and AI model

In our previous work [5] we trained AI models for CAG segmentation using 416 images from patients undergoing physiology or PCI in a single center. The images were manually annotated by a small group (two Cardiology Fellows and an Interventional Cardiologist, who both annotated and supervised the process) and continuously reviewed and corrected, in order to minimize heterogeneity and errors.

We then performed segmentation using an encoder-decoder fully convolutional neural networks based on the U-Net [6], commonly used in medical image segmentation. These are composed of an encoder for extracting image features and a decoder to process those features and produce segmentation masks. To derive the best approach for this task, we conducted a comparative study of encoder and decoder architectures, which resulted in the proposal of the EfficientUNet++, a computationally efficient and high-performing decoder architecture [7], which obtained the best results when combined with an EfficientNet-B5 encoder [8].

To ensure fair evaluation and minimize any bias induced by the input data, each model was tested on data it had not seen during training. The dataset was thus split at the patient level, into 13 subsets of approximately 32 angiograms each. Each subset's segmentation was performed using a neural network trained exclusively on the remaining data. This enabled the assessment of the segmentation results for the entire cohort, as the usual splitting into a training and testing dataset would have yielded a much smaller group of images for result assessment. The training hyperparameters, namely the number of training epochs and the learning rate decay schedule, were set on the first train-test split, using 1 of the 12 training data subsets for validation. The selected values were then used on every other train-test split, and to train the model on the whole training set of the first split.

We also considered cross-validation, but it would be very compute-heavy.

This process resulted in an early AI model, which was then further improved by a second round of manual annotation, where the annotators corrected the resulting imperfections, thereby producing a final training dataset. An "enhanced" model was then trained once again using the same process with the new improved annotated dataset, yielding superior results to the early model, with a final Generalized Dice Score of 93,48/ +/- 2,84%. While we continue to work on improving our model, because the aim of this study is to validate the aforementioned "enhanced" model as previously published [5], no additional training was performed.

Original images analysis and segmentation

A well-established and validated software (CAAS Workstation 8.5.1) capable of semi-automatic segmentation and Quantitative Coronary Angiography (QCA) was used to generate a reference dataset for comparison. Because it is especially important for a model to correctly segment diseased segments, QCA analysis was performed in selected segments with a stenosis severity of 50–99% by visual estimation. For QCA measurements, calibration was performed either automatically (based on the DICOM information) or by measuring the catheter (5 or 6 Fr), provided it was clearly visible and measurable. The region of interest was then selected and automatic QCA measurements were undertaken.

For each region of interest where successful automatic QCA measurements were undertaken, the lesion diameter, reference diameter, diameter at proximal obstruction border and diameter at distal obstruction border were recorded. The diameter stenosis percentage was calculated as follows: ((reference diameter − lesion diameter) / reference diameter) x 100 [9]. No manual adjustments were accepted, in order to exclude human bias or human-induced imperfection. If the automated outline and measurements were not clearly accurate by visual inspection, the case was excluded (supplementary Fig. 1).

The original images (i.e., without the measurement annotations generated by the CAAS software) were then segmented using our best AI model to date [5], which segments the coronary tree in white and the catheter in red. This process is fully automatic and the only required human input is the image itself. These images were used for testing only, not training.

Performance assessment

Diameters and percentage diameter stenosis

A dedicated python script was written to extract the CAAS markers and superimpose them on the segmentation obtained by the model. The lesion diameter, diameter at proximal obstruction border and diameter at distal obstruction border were then measured using a dedicated python script as well, by verifying the superimposition of the markers with the coronary tree. Because the reference diameter does not exist in the segmented image (which only contains the coronary artery tree and catheter), the CAAS-generated value was used. Percentage stenosis was then calculated using the same equation. Finally, we also compared the measured catheter diameter on the original image versus the segmented image with another adaption of the same script, by measuring the distance between the two parallel lines generated in the original image from the CAAS software. The resulting measurements obtained in the original and the segmented images were then compared.

Overlap between original and segmented images

A dedicated python script was also used for assessing the overlap between the original and the segmented images in the region of interest, using the CAAS output as reference. Pixels were then classified as follows:

- True positive (TP): a pixel marked as coronary in both the segmented and original image.
- False positive (FP): a pixel marked as coronary only in the segmented image.
- True negative (TN): a pixel marked as non-coronary in both the segmented and original image.
- False negative (FN): a pixel marked as non-coronary only in the original image.

Using this classification, the following parameters were calculated:

- Accuracy: ([TP+TN]/[TP+TN+FP+FN])
- Sensitivity: TP / (TP+FN).
- Specificity: TN/(TN+FP),
- Positive Predictive Value: TP / (TP+FP).
- Negative predictive value TN/(TN+FN).
- Intersection over Union (IoU): TP / (TP+FN+FP).
- Dice Score: 2TP / (2TP+FN+FP).

Appendix

Global segmentation score

While the above-mentioned criteria offer a detailed account of the model's accuracy, they do not provide a broad overview of the quality of segmentation as assessed by experts in CAG interpretation (i.e. Cardiologists). As a result, we have previously developed the Global Segmentation Score (GSS), which we have previously applied on the original CAG dataset used to train the AI model (details on its application on supplementary data file) [5]. The GSS was scored by consensus by four Interventional Cardiologists (one from each contributing center).

Figure 1 summarizes the above-mentioned steps for assessing coronary segmentation.

Statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. Quantitative variables are shown in average ± standard deviation (if normally distributed) or median (interquartile range) if non-normally distributed. If distribution was normal, we used the paired samples T-test to assess for differences in related samples quantitative variables. If distribution was not normal, we used the Mann-Whitney test (two independent groups) or the Kruskal Wallis test (multiple independent groups) to assess for differences in quantitative variables. A p-value < 0,05 was used for statistical significance. SPSS 27 was used for analysis.

Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local Ethics' Institutional Review Board.

Results

Baseline characteristics

We included 123 measurements from 117 images, from a total of 90 patients (flowchart in Fig. 2; clinical data on Table 1). The left anterior descending artery (LAD) was the most common target vessel (three measurements were taken on diagonals, two emerging proximally and one emerging in the middle segment of the LAD; all were taken on the proximal segment of the collateral), with measurements taking place more frequently in the middle and proximal segments. As measured by QCA, most lesions had a 50–69% diameter stenosis, with a minority of ≥ 70% lesions (Table 2).

Performance

Diameters and percentage diameter stenosis

Detailed metrics of images (Fig. 3) are depicted in Tables 3 and 4. There were no significant differences for all parameters except for diameter at proximal obstruction border, where the median difference between groups was 0,19 mm. All difference parameters (Table 3) had a non-normal distribution, with the interquartile range demonstrating that there is a clear predominant difference towards the lower-end values, as the 25th quartile is either 0 or very close to 0.

There were no significant differences across stenosis severity (supplementary Table 1) or target vessel (supplementary Table 2). There were also no significant differences considering across centers (supplementary Tables 3 and 4).

With regards to the catheter diameters (Fig. 4), results are shown on supplementary Table 5. A significant number of cases (26/117 – 22%) had to be excluded, either because of collimation (rendering the catheter not visible – 8 cases) or segmentation gaps leading to inaccurate border definition (18 cases). The latter occur because the model focuses especially on segmenting the distal part of the catheter for correctly identifying the transition between catheter and coronary, whereas in the original images calibration occurred predominantly in less distal portions. Because the presence of two groups (5 and 6 Fr) of catheters renders the overall distribution of the sample non-normal, the two groups were analysed separately. There were no significant differences between the original and segmented images. Again, the difference parameter had a non-normal distribution, with the interquartile range demonstrating that there is a clear predominant difference towards the lower-end values.

Overlap between original and segmented images

Results are detailed on Table 5. The model scored ≥ 90% in all metrics (Fig. 5). There were some significant differences between target vessel (supplementary Table 6) and stenosis severity (supplementary Table 7) which, in absolute terms, were between 1 and 3%. There were no differences between centers (supplementary Table 8).

Global segmentation score

Results are shown on supplementary Table 9. The model scored well above or close to 90% in most criteria. Catheter gaps were common, usually due to contrast backflow impeding proper visualization of such portions. Catheter artifacts were common and mild gaps in distal parts of small collaterals were quite common as well.

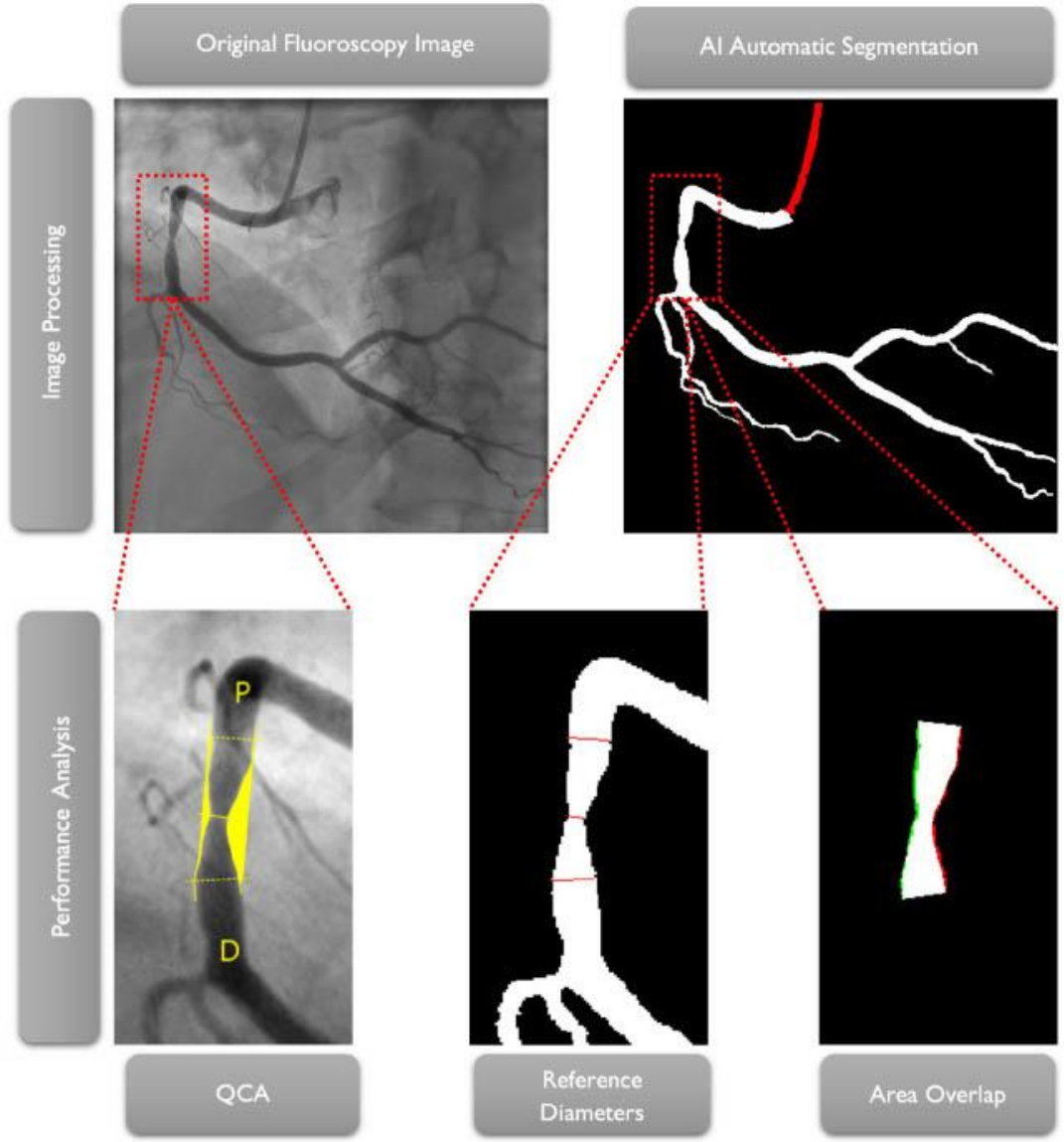


Fig. 1 Graphical Abstract: Overview of the segmentation and analysis process. Top left: Baseline CAG of a right coronary artery. Top right: AI automated segmented image. Bottom left: automatic QCA analysis image output in detail. Bottom middle: transposition of the lesion markers on the segmented image in detail. Bottom right: area overlap between the region of interest in the auto-QCA and the segmented image; white pixels are true positives; green pixels are false negatives; red pixels are false positives

N is lower than overall measurements due to assessment of more than one lesion per image and 8 cases of collimation where the catheter could not be scored, thereby excluding those cases from assessment.

Discussion

Main findings

A deep learning AI segmentation model was capable of fully

Appendix

Fig. 2 Flowchart of patient and image selection

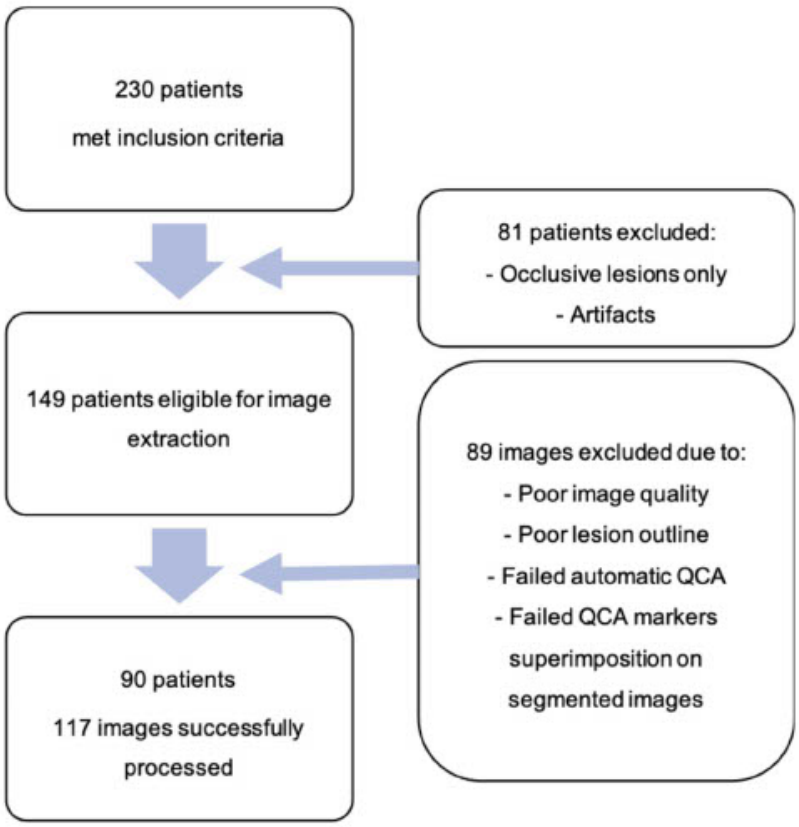


Table 1 Clinical characteristics of included patients

Factor	N +/- SD or N(%)
Age	65 +/- 12
Sex (male)	73 (81%)
Hypertension	62 (68,9%)
Diabetes mellitus	28 (31,1%)
Dyslipidemia	55 (61,1%)
Smoker (past or present)	50 (55,6%)
Chronic coronary syndrome	37 (41,1%)
Acute coronary syndrome	53 (58,9%)
Revascularization during/after CAG	76 (84,4%)
Invasive Physiology during procedure	19 (21,1%)

automatic accurate CAG segmentation, as checked by a reference segmentation obtained with validated software and also when assessed by a broad assessment score we previously developed [5].

Diameters at both healthy segments (proximal and distal lesion borders) and diseased segments (diameter at maximum obstruction zone) were similar between the two groups, with statistically significant differences only at the

Table 2 Distribution of target vessel and lesion severity. LAD: Left Anterior Descending Artery; RCA: Right Coronary Artery; CX: Left Circumflex Artery

Parameter			N (%)
Target Vessel	LAD	Proximal	19
		Middle	23
		Distal	8
		Total	50 (41)
	RCA	Proximal	9
		Middle	24
		Distal	8
		Total	41 (33)
	CX	Proximal	9
		Middle	21
		Distal	2
		Total	32 (26)
Lesion severity		≥ 70%	22 (18)
		50–69%	58 (47)
		< 50%	43 (35)

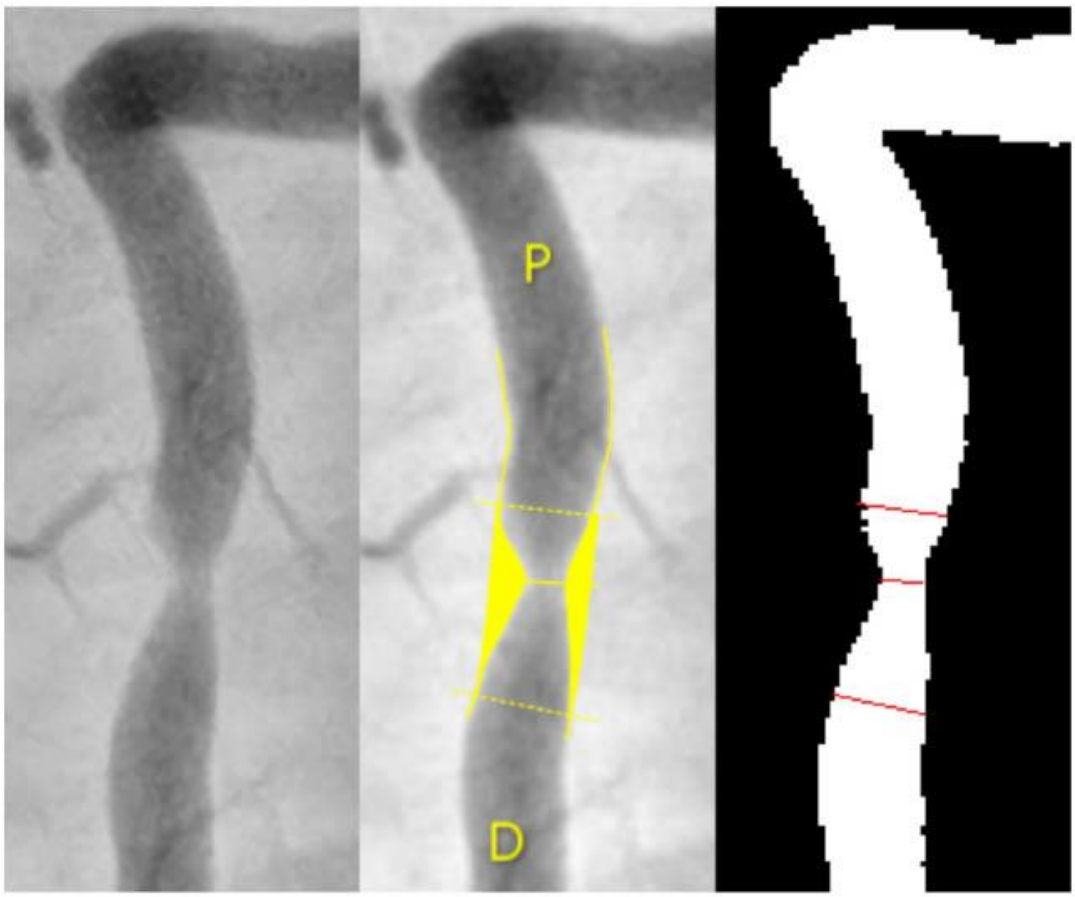


Fig. 3 Comparative view of a right coronary artery (56% stenosis by QCA). Left-to-right: original image, auto-QCA, transposition of lines (proximal border diameter, lesion diameter and distal border diameter) to segmented image

Table 3 Detailed measurements between the original and the segmented images. Values shown as mean ± standard deviation. AI – artificial intelligence. *Paired samples T-test;

Parameter	Original Image	AI Generated Segmented Image	P-value*
Diameter Stenosis (%)	56 ± 13	55 ± 13	0,071
Diameter at lesion (mm)	1,06 ± 0,39	1,08 ± 0,37	0,146
Diameter at proximal obstruction border (mm)	2,27 ± 0,54	2,09 ± 0,53	**< 0,01**
Diameter at distal obstruction border (mm)	2,19 ± 0,56	2,15 ± 0,58	0,133

proximal obstruction border. However, in absolute terms, the difference was very small (0,19 mm, a < 10% difference considering the proximal diameter in either group) and we therefore believe it is unlikely to be of clinical significance.

Table 4 Median differences between the original and segmented images. Values shown as median (IQ 25th – 75th)

Parameter	Difference
Diameter Stenosis (%)	4,5 (0–7,7)
Diameter at lesion (mm)	0,10 (0–0,17)
Diameter at proximal obstruction border (mm)	0,19 (0,09–0,28)
Diameter at distal obstruction border (mm)	0,10 (0–0,19)

The stenosis severity as assessed by percentage stenosis only differed by < 5% in absolute terms, a difference not meaningful either statistically or clinically. The latter is perhaps the single most important finding, as percentage diameter stenosis is the fundamental criteria assessed in clinical practice for proceeding with either revascularization or functional testing, as recommended in current guidelines [10]. Importantly, there were no significant differences

Appendix

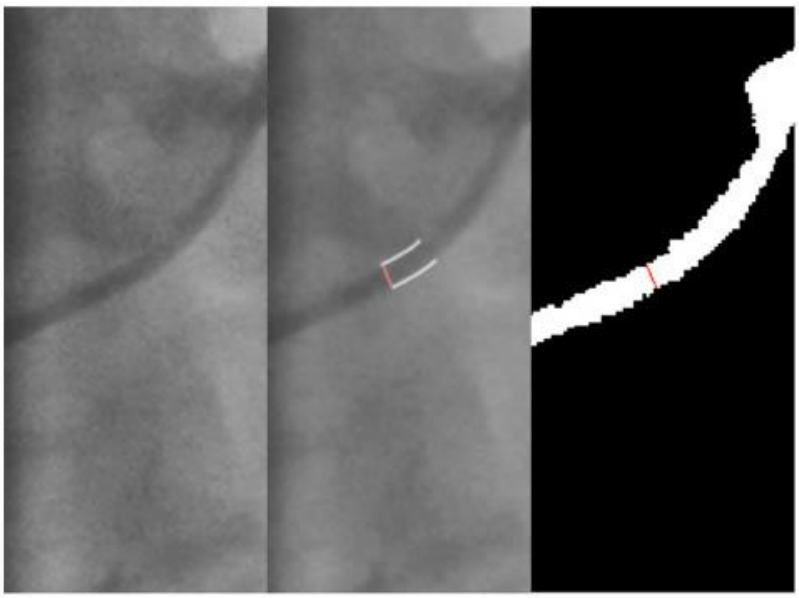


Fig. 4 Catheter segmentation assessment. Left-to-right: original image, auto-border detection by reference software, transposition of lines in proximal border to segmented image

Table 5 Overlap metrics. Values shown as median (IQ 25th – 75th)

Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Intersection over Union (%)	Dice Score (%)
99,9 (99,9–99,9)	95,1 (92,8–96,4)	99,9 (99,9–99,9)	94,9 (93,1–96,5)	99,9 (99,9–99,9)	90,1 (87,6–91,7)	94,8 (93,4–95,7)

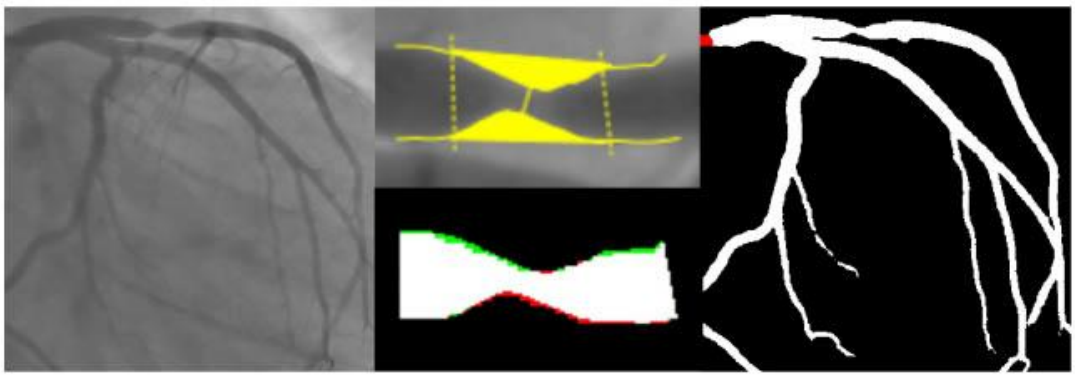


Fig. 5 Area overlap in a and left anterior descending 64% stenosis (as measured by QCA).

in performance regarding target vessel, stenosis severity or centers.

When considering the overlap between the segmented image and the original image, accuracy, specificity and negative predictive value scored close to 100%. This was expected, because most of the image is composed by background rather than artery. As a result, we believe metrics that do not take into account true negatives provide a more faithful indication of the actual model performance. In that regard, sensitivity and positive predictive value still scored quite high, at approximately 95%. The metric that more directly assesses the true overlap between the original and segmented images in the region of interest (correctly identifying all of the vessel while avoiding non-artery pixels) is the intersection over union criteria, which fell just short of 90%. Lastly, the Dice Score puts greater emphasis on the

fundamental task of segmentation – correctly identifying the target structure i.e. true positives – in this case, the coronary tree. With an average score of approximately 95%, while also considering all the remaining metrics, we believe our model can be described as accurate. Importantly, the Dice Score in our previous study was 93%, thus very similar to what we now found [5]. There were statistically significant differences in the IoU and Dice Scores between target vessel stenosis severity. Notwithstanding, the absolute differences were very minor (around 1–2%) and therefore of little or no clinical relevance.

With regards to the GSS, our model achieved a high score with a median of 92/100 points, exceedingly similar to what we had previously described in the dataset used to train and develop the model. The model scored very high in almost all tasks, while maintaining minor imperfections with regards to mild gaps in collateral branches, which were very frequent. Catheter segmentation was not as good as coronary segmentation, as usually small catheter artifacts or gaps in the vicinity of the coronary tree origin were common. This was due not only to contrast backflow, but also because of how AI models are trained and function. Indeed, performance is very dependent on class frequency. Because the catheter is a less frequent class (i.e. corresponds to much fewer pixels), the models receive less penalty for errors regarding its segmentation when compared to the coronary tree. This is partly mitigated by the use of an appropriate loss function, but the imbalance nevertheless persists to some extent. Once again, this was very similar to what we saw in the training dataset [5]. With regards to precise catheter measurements, the differences between original and segmented images (for both 5 and 6 Fr catheters) were not statistically significant, suggesting the catheter's segmentation, from a calliper precision point of view, is accurate. However, due to the above-mentioned limitations and to a small number of images where only a small portion (or none at all) of the catheter was discernible, our sample was somewhat reduced, thereby limiting this assessment.

Other studies in the field

There are very few studies published in medical/biology journals to date where a comparison with our results can be made. With regards to the GSS in particular, no similar application has ever been undertaken, to our knowledge.

The largest published study [3] included a dataset of 1050 images distributed across all incidences and vessels for performance evaluation. An average 98% accuracy was obtained. While specificity and negative predictive values scored very highly, sensitivity and positive predictive value came closer to 80%. The performance was slightly inferior in more distal vessels. Intersection over union or Dice Score were not reported. Importantly, however, that study's evaluation used the baseline human annotation as reference, rather than an external validated software, thereby not enabling the identification of bias or imperfections which might have become embedded in their AI model. In our previous study, we demonstrated that even with a small group of annotators and continuous review of the annotations, there is always some degree of imperfection in human annotation [5], hence the relevance of comparing against an automated and validated external software. Additionally, the reported accuracy focuses on the overlap across the entire coronary tree rather than the percentage stenosis of diseased segments. This is advantageous in the sense that a globally accurate performance can be tested. Notwithstanding, we believe testing only for diseased segments actually renders the comparison more demanding. This is because the segmentation of stenotic segments is harder from a technical point of view and also due to the fact that the number of true positive pixels is necessarily smaller in such segments – leading to a lower likelihood of true positives. Whichever interpretation is made, it is clear that an exact comparison with Du et al. [3] is not possible. However, broadly speaking, the accuracy of both models seems quite high and our model seems at least as accurate, if not more.

Su Yang et al. [4] also produced AI models for CAG segmentation. Their validation dataset was somewhat larger (181 images), but their performance seems slightly lower, with all overlap metrics generally scoring just short of 90% and a Dice Score of 89%. Importantly, they also only segmented diseased segments, with a minimum lesion of 30% and used the same reference software as we did. Thus, their results are more directly comparable to ours and our model seems to have superior performance. Two other works [1, 2], from the same baseline dataset, also went on to develop AI-based CAG segmentation, this time with a validation dataset of 550 images. While the model performed well, with an accuracy of 98% and a sensitivity of 87%, they also based their validation dataset on human annotation of the coronary tree without using external software. Thus the above-mentioned considerations for Du et al. [3] also apply.

Recently, Gao et al. [11] published the results of a CAG segmentation model trained on only 130 images. Their methodology, however, is somewhat different, since they combined features from deep learning segmentation models' features and non-AI image filters to perform pixel-wise classification using gradient-boosting decision trees [12] and deep forests [13]. Their results also show good performance, with a Dice Score of 87,4%, sensitivity of 90,2% and specificity of 99,2%. This highlights that merging deep learning with traditional computer vision methods can yield good results, when working with relatively small datasets. However, no external validation software was used and the

Appendix

whole coronary tree was evaluated. As a result, once more, the previous considerations for Du et al. [3] apply.

Other works in the application of AI to coronary segmentation are primarily technical and featured in engineering publications. A detailed review of these falls outside the scope of this paper and can be consulted in our previous technical publication [7]. However, some considerations regarding these provide further contextualization of our findings.

Xian et al. [14] used a very large dataset of 3200 manually annotated images and experimented with the U-Net architecture as well, with a sensitivity of 90,1%, positive predictive value of 89,8% and Dice Score 90%. However, the annotations were undertaken with a specific software for the purpose of coarsely signaling the vessel route, and focused only on the main vessels. Since we achieved higher performance metrics, it seems a smaller but higher quality dataset, with very precise and cumbersome manual annotations, may be a better approach.

Yang et al [15] have obtained a sensitivity, positive predictive value and Dice Score of 91,3%, 92,5% and 91,9%, respectively, by using popular image classification backbones pre-trained on ImageNet instead of the U-Net's encoder, while also using a modified generalized dice loss function. Their findings were influential in our training method, as we used a combination of their proposed loss function and the focal loss [16]. Other authors have explored the use of dense connections, improving on the performance of the standard U-Net [17]. This approach is also present in the U-Net ++ [18], which we used in our approach.

In all of the above studies, metrics regarding vessel diameters were not performed. Thus, a direct comparison with this study regarding those is not possible. M'hiri et al. addressed the issue of CAG diameter measurements, when dealing with the issue of diameter variation during the cardiac cycle due to vessel distensibility. They focused mainly in measuring specific segments of the coronary tree, as we did. However, they used a graph-based segmentation method, then tracked the changes across the cardiac cycle using a spatio-temporal segmentation method. They obtained a Dice Score of 98%, with a very small diameter mean error (0,18 mm) [19]. However, they did not focus on diseased regions. While this study is not focused on AI methods, it highlights that other methods may be of use for accurate CAG segmentation, potentially in combination with AI tools [11].

In light of all these studies, the performance of our model seems at least as good, if not better, than previously proposed AI models. We believe this is related to its neural network architecture, which was carefully chosen over a series of experiments [7], taking into consideration the invaluable contributions of previously mentioned studies. In addition to that, we also believe that our manual annotations methodology was essential, as it allowed us to obtain a highly reliable training dataset: a small number of annotators (to reduce heterogeneity) well trained in the interpretation of coronary angiograms; very careful review of annotations with recurrent iterations of quality checks and improvements; and further manual improvement of the already accurate segmentation images produced by an earlier AI model, thus combining the best of AI and human annotations into a final training dataset, as mentioned in the methods section and previous publication [5].

Limitations

Our study is not without limitations. Despite the multicentric approach, our dataset is relatively small when compared to previously published studies. We also only tested the model performance against validated software in diseased locations, rather than on the whole coronary tree. Therefore, we cannot affirm that the performance would be identical in the remaining areas. However, as previously explained, segmenting zones with lesions is actually more challenging for the model than segmenting broad, mostly healthy segments. In addition to that, we did not find differences regarding target vessel or lesion severity. Plus, considering the results of the GSS, the overall performance regarding CAG segmentation was quite appropriate. Thus, we believe that it is unlikely that performance would be significantly different had we tested for the whole coronary tree. Importantly, if we had chosen to segment whole vessels, it would be very likely that some manual corrections had to be undertaken, which might induce bias or imperfections in the reference images. Hence, the decision to proceed as described was deliberate. The assessment of catheter segmentation was also more limited than that of the coronary tree, as described above.

The exclusion of potential sources of artifacts from devices or previous cardiac surgery means our model is not yet applicable to such patients. Notwithstanding, we didn't exclude cases with previous implantation of stents, but we did not perform detailed measurements on such segments.

The total number of patients/images who fully met exclusion criteria was somewhat high, thereby limiting the final amount of available images for analysis, which may raise questions as to whether this sample is representative of everyday CAGs and an therefore constitutes an adequate validation dataset. This was the result of somewhat stringent criteria, which we felt were nonetheless necessary due to basic feasibility (such as excluding single-vessel complete occlusion cases where QCA is not applicable, or excluding imaging artifacts for which the models are not yet trained), reduction of bias (such as not allowing for manual QCA correction), or excluding patients with normal/near-normal

arteries (where testing would be much less challenging or useful in future clinical application). Notwithstanding, we included patients consecutively rather than selectively and the clinical characteristics of included patients are in agreement with everyday clinical practice. We therefore believe our sample to be reasonably representative of real-world practice. Furthermore, we exceeded the minimum validation target of 100 images, yielding relative rates of training vs. validation cases in agreement with other AI studies [2–4].

The imbalance in sample size limits the comparison between centers.

It has long been established that operators significantly differ in their interpretation of lesion severity and have a tendency to overestimate the importance of a stenosis [20–24], as we also saw in this study. Indeed, while visually all lesions were interpreted as > 50% stenosis, a significant amount of the sample actually had a < 50% stenosis, which further reflects the real-world nature of the dataset.

Lastly, the distance between the 2D centerline and the distance to the closest edge would have also been a good metric for assessing model performance in this setting. We did not perform such testing.

In light of all of the above, concerns may be raised regarding generalization from this dataset. However, we believe that the absence of statistically significant differences across all subgroups at least partially attenuates this concern.

Future directions

We are currently working in automatic anatomical interpretation, lesion severity based on auto-QCA and integration with physiology. We believe without effective segmentation models, none of these will be possible. Much like for human interpretation of CAG, separating the coronary arteries from everything else in the image is an essential first step. Our ultimate goal is to produce an intelligence augmentation tool that helps physicians perform a more objective and streamlined interpretation of CAG, hopefully contributing for better patient outcomes. As we continuously improve its performance, while also adding new capabilities, clinical application will potentially be possible in the near future, opening a new perspective and potentially more accurate method to assess coronary artery disease.

We are also continuously working to expand and improve the model, as segmentation alone is not a final goal in itself, but rather a fundamental step. We hope to release a public version in the near future, which other researchers may use for whichever application they may deem useful. Importantly, comparing or even merging with future models from other groups may also be very relevant. Since it uses an inherently data-hungry deep learning model, our coronary artery segmentation system would surely benefit from training on a larger volume of data. Manual annotation of coronary angiography images, however, is very cumbersome and time-consuming, and therefore it is difficult to obtain much larger labeled datasets. Hence, significant improvements to the model could probably be achieved, for example, by using self-supervised learning on existing very large volumes of unlabeled data. These possibilities are described in detail in our previous technical publication [7].

Data Availability

Detailed full-scale study data cannot currently be made publicly available due to limitations imposed by national data protection regulations, as this is a retrospective study and no informed consent was obtainable regarding this particular analysis. Both our research team and others in the national scientific community are working to develop a framework where such would be possible. However, independent replication of our analysis is possible, given that the detailed description of our experimentations and relevant code is publicly available [7].

Conclusions

Our AI model was capable of accurate CAG segmentation when applied to a multicentric validation dataset, with no differences between target vessels or stenosis severity. This paves the way for future research and implementation for its clinical uses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10554-023-02839-5.

Author Contribution MNM was responsible for conceptualization, data/image gathering, processing and analysis and paper drafting. JLS was responsible for technical and AI tasks, data and image processing, model implementation and training. CR, JPG and MSO were responsible for data gathering and image analysis. TR and BS were responsible for data gathering, processing and analysis in the development of the original dataset. ALO was responsible for supervising the work of JLS, having directly taken part in the same tasks. FJP was responsible for supervising the work of MNM, having directly taken part in the same tasks.
All authors revised the paper critically for important intellectual content, gave final approval for its publication and agree to be accountable for all respects of its accuracy and integrity.

Funding Open access funding provided by FCT|FCCN (b-on). Cardiovascular Center of the University of Lisbon, INESC-ID / Instituto Superior Técnico, University of Lisbon.

Declarations

Appendix

Conflict of Interest Not applicable.

References

1. Wang L, Liang D, Yin X et al (2020) Coronary artery segmentation in angiographic videos utilizing spatial-temporal information. BMC Med Imaging 2020 201 20:1–10. https://doi.org/10.1186/S12880-020-00509-9
2. Liang D, Qiu J, Wang L et al (2020) Coronary angiography video segmentation method for assisting cardiovascular disease interventional treatment. BMC Med Imaging 2020 201 20:1–8. https://doi.org/10.1186/S12880-020-00460-9
3. Du T, Xie L, Zhang H et al (2021) Training and validation of a deep learning architecture for the automatic analysis of coronary angiography. EuroIntervention 17:32–40. https://doi.org/10.4244/EIJ-D-20-00570
4. Yang S, Kweon J, Roh J-H et al (2019) Deep learning segmentation of major vessels in X-ray coronary angiography. Sci Rep 2019 91 9:1–11. https://doi.org/10.1038/s41598-019-53254-7
5. Nobre Menezes M, Lourenço-Silva J, Silva B et al (2022) Development of deep learning segmentation models for coronary X-ray angiography: Quality assessment by a new global segmentation score and comparison with human performance. Rev Port Cardiol. https://doi.org/10.1016/J.REPC.2022.04.001
6. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 9351:234–241
7. Silva JL, Menezes MN, Rodrigues T et al (2021) Encoder-decoder architectures for clinically relevant Coronary artery segmentation. arXiv:2106.11447 [eess.IV]
8. Tan M, Le QV (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 36th Int Conf Mach Learn ICML 2019 2019-June:10691–10700
9. Suzuki N, Asano T, Nakazawa G et al (2020) Clinical expert consensus document on quantitative coronary angiography from the Japanese Association of Cardiovascular intervention and therapeutics. Cardiovasc Interv Ther 35:105. https://doi.org/10.1007/S12928-020-00653-7
10. Neumann FJ, Sousa-Uva M, Ahlsson A et al (2019) 2018 ESC/EACTS guidelines on myocardial revascularization. Eur Heart J 40:87–165. https://doi.org/10.1093/EURHEARTJ/EHY394
11. Gao Z, Wang L, Soroushmehr R et al (2022) Vessel segmentation for X-ray coronary angiography using ensemble methods with deep learning and filter-based features. BMC Med Imaging 22:1–17. https://doi.org/10.1186/S12880-022-00734-4/TABLES/5
12. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29:1189–1232. https://doi.org/10.1214/AOS/1013203451
13. Zhou ZH, Feng J (2017) Deep Forest. Natl Sci Rev 6:74–86. https://doi.org/10.48550/arxiv.1702.08835
14. Xian Z, Wang X, Yan S et al (2020) Main coronary vessel segmentation using deep learning in Smart Medical. Math Probl Eng. https://doi.org/10.1155/2020/8858344
15. Yang S, Kweon J, Kim Y-H (2022) Major vessel segmentation on X-ray coronary angiography using deep networks with a Novel Penalty loss function.Proc Mach Learn Res Rev1–5
16. Lin TY, Goyal P, Girshick R et al (2017) Focal loss for dense object detection. Proc IEEE Int Conf Comput Vis 2017-October 2999–3007. https://doi.org/10.1109/ICCV.2017.324
17. Jun TJ, Kweon J, Kim YH, Kim D (2020) T-Net: nested encoder–decoder architecture for the main vessel segmentation in coronary angiography. Neural Netw 128:216–233. https://doi.org/10.1016/J.NEUNET.2020.05.002
18. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J (2018) Unet++: A nested u-net architecture for medical image segmentation. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 11045 LNCS:3–11. doi: https://doi.org/10.1007/978-3-030-00889-5_1/COVER
19. M'hiri F, Duong L, Desrosiers C et al (2017) Automatic evaluation of vessel diameter variation from 2D X-ray angiography. Int J Comput Assist Radiol Surg 12:1867–1876. https://doi.org/10.1007/S11548-017-1639-9/FIGURES/9
20. Fischer JJ, Samady H, McPherson JA et al (2002) Comparison between visual assessment and quantitative angiography versus fractional flow reserve for native coronary narrowings of moderate severity. Am J Cardiol 90:210–215. https://doi.org/10.1016/S0002-9149(02)02456-6
21. Adjedj J, Xaplanteris P, Toth G et al (2017) Visual and quantitative Assessment of Coronary Stenoses at Angiography Versus Fractional Flow Reserve: the impact of risk factors. Circ Cardiovasc Imaging. https://doi.org/10.1161/CIRCIMAGING.117.006243
22. Nallamothu BK, Spertus JA, Lansky AJ et al (2013) Comparison of clinical interpretation with visual assessment and quantitative coronary angiography in patients undergoing percutaneous coronary intervention in contemporary practice: the assessing angiography (A2) project. Circulation 127:1793–1800. https://doi.org/10.1161/CIRCULATIONAHA.113.001952
23. Zhang H, Mu L, Hu S et al (2018) Comparison of Physician Visual Assessment with quantitative coronary angiography in Assessment of Stenosis Severity in China. JAMA Intern Med 178:239–247. https://doi.org/10.1001/JAMAINTERNMED.2017.7821
24. Shah R, Yow E, Jones WS et al (2017) Comparison of visual assessment of coronary stenosis with independent quantitative coronary angiography: findings from the PROMISE trial. Am Heart J 184:1. https://doi.org/10.1016/J.AHJ.2016.10.014

Received: 7 June 2023 | Accepted: 1 August 2023

DOI: 10.1002/ccd.30805

ORIGINAL ARTICLE - CLINICAL SCIENCE

WILEY

Segmentation of X-ray coronary angiography with an artificial intelligence deep learning model: Impact in operator visual assessment of coronary stenosis severity

Miguel Nobre Menezes MD, MSc[1,2] | Beatriz Silva MD, MSc[1,2] | João Lourenço Silva MSc[3] | Tiago Rodrigues MD, MSc[1,2] | João Silva Marques MD, MSc[1,2] | Cláudio Guerreiro MD, MSc[4] | João Pedro Guedes MD, MSc[5] | Manuel Oliveira-Santos MD, MSc[6,7] | Arlindo L. Oliveira PhD[3] | Fausto J. Pinto PhD[1,2]

[1]Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

[2]Departamento de Coração e Vasos, Serviço de Cardiologia, CHULN Hospital de Santa Maria, Lisboa, Portugal

[3]INESC-ID/Instituto Superior Técnico, Lisbon, Portugal

[4]Department of Cardiology, Centro Hospitalar de Vila Nova de Gaia, Vila Nova de Gaia, Portugal

[5]Unidade de Hemodinâmica e Cardiologia de Intervenção, Serviço de Cardiologia, Centro Hospitalar Universitário do Algarve, Hospital de Faro, Faro, Portugal

[6]Unidade de Intervenção Cardiovascular, Serviço de Cardiologia do Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal

[7]Pólo das Ciências da Saúde, Unidade Central, Azinhaga de Santa Comba, Celas, Faculty of Medicine, University of Coimbra, Coimbra, Portugal

Correspondence
Miguel Nobre Menezes, MD MSc, Serviço de Cardiologia, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal.
Email: mnmenezes.gm@gmail.com

Funding information
INESC-ID; Cardiovascular Center of the University of Lisbon, INESC-ID/Instituto Superior Técnico, University of Lisbon

Abstract

Background: Visual assessment of the percentage diameter stenosis ($\%DS_{VE}$) of lesions is essential in coronary angiography (CAG) interpretation. We have previously developed an artificial intelligence (AI) model capable of accurate CAG segmentation. We aim to compare operators' $\%DS_{VE}$ in angiography versus AI-segmented images.

Methods: Quantitative coronary analysis (QCA) %DS ($\%DS_{QCA}$) was previously performed in our published validation dataset. Operators were asked to estimate $\%DS_{VE}$ of lesions in angiography versus AI-segmented images in separate sessions and differences were assessed using angiography $\%DS_{QCA}$ as reference.

Results: A total of 123 lesions were included. $\%DS_{VE}$ was significantly higher in both the angiography (77% ± 20% vs. 56% ± 13%, $p < 0.001$) and segmentation groups (59% ± 20% vs. 56% ± 13%, $p < 0.001$), with a much smaller absolute %DS difference in the latter. For lesions with $\%DS_{QCA}$ of 50%–70% (60% ± 5%), an even higher discrepancy was found (angiography: 83% ± 13% vs. 60% ± 5%, $p < 0.001$; segmentation: 63% ± 15% vs. 60% ± 5%, $p < 0.001$). Similar, less pronounced, findings were observed for $\%DS_{QCA} < 50\%$ lesions, but not $\%DS_{QCA} > 70\%$ lesions. Agreement between $\%DS_{QCA}/\%DS_{VE}$ across $\%DS_{QCA}$ strata (<50%, 50%–70%, >70%) was approximately twice in the segmentation group (60.4% vs. 30.1%; $p < 0.001$). $\%DS_{VE}$ inter-operator differences were smaller with segmentation.

Conclusion: $\%DS_{VE}$ was much less discrepant with segmentation versus angiography. Overestimation of $\%DS_{QCA} < 70\%$ lesions with angiography was especially common. Segmentation may reduce $\%DS_{VE}$ overestimation and thus unwarranted revascularization.

KEYWORDS
artificial intelligence, coronary angiography, coronary artery disease, deep learning, machine learning, percutaneous coronary intervention

Abbreviations: AI, artificial intelligence; CAG, coronary angiography; %DS, percentage diameter stenosis; $\%DS_{QCA}$, percentage diameter stenosis by QCA; $\%DS_{VE}$, percentage diameter stenosis by visual estimation.

Appendix

1 | INTRODUCTION

The assessment of the severity of coronary stenosis is essential for revascularization decisions. In clinical practice, operators often begin by assessing the percentage diameter stenosis (%DS) of lesions, which can either be estimated visually ($\%DS_{VE}$) or by means of direct semi-automatic measurement with quantitative coronary angiography (CAG) (QCA—$\%DS_{QCA}$).[1] However, multiple studies have shown that visual inspection tends to result in average higher percent diameter stenosis than QCA, with heterogeneity across operators and/or hospitals.[2–6] Disagreements of lesion severity by visual inspection versus QCA may be clinically relevant, as they have been associated to the likelihood of clinical events.[7]

Furthermore, the assessment of the functional significance of stenosis by means of fractional flow reserve (FFR) has been proven to be superior to that of angiography alone regarding clinical outcomes, despite the fact that physiology-guided revascularization results in lower rates of percutaneous coronary intervention, once again highlighting that operators tend to overestimate the severity of lesion severity by visual estimation.[8] As a result, current guidelines strongly emphasize the role of physiology or ischemia testing in the assessment of coronary lesion severity, rather than angiography alone.[9,10] Despite this, the adoption of physiology remains low.[11,12]

As a result, non-invasive and automatic tools that reduce the heterogeneity of CAG interpretation are desirable.[13] Artificial intelligence (AI) may be of use for such a task, but few studies are available in medical/biology publications regarding its application for CAG.[14–17] We have recently developed AI models capable of accurate CAG segmentation.[18,19] In this study, we sought to evaluate how CAD lesion severity is perceived by operators when CAGs are viewed in AI-segmented versus fluoroscopy images, using QCA as reference.

2 | METHODS

2.1 | Previous work and study population

We have previously trained AI models for CAG segmentation based on manual CAG annotation of patients undergoing invasive physiology assessment (FFR and/or other indexes) or PCI, with an original sample of 416 images.[18,20] Recently, we published the results of our validation study with an additional dataset of 117 images.[19] Briefly, consecutive patients who had undergone (PCI) and/or invasive physiology assessment in four centers from across Portugal were selected. The images were then automatically segmented with our AI model. Lesions were measured by QCA with a validated software (CAAS Workstation 8.5.1) in the original images, which were then compared to the segmented images. We have shown that the AI-generated segmentation was highly accurate, with no significant differences between percentage diameter stenosis in fluoroscopy versus segmented images, across all degrees of lesion severity, target lesion or fluoroscopy equipment.[19]

In this study, we chose to use the validation cohort for assessing the impact of segmentation in the perception of lesion severity, given that significant differences between segmented and original images are thus a priori excluded, as they have all been previously measured. As a result, potential differences can be attributed to visual perception rather than actual dataset discrepancies.

2.2 | Stenosis severity assessment

One operator from each participating center was shown all images consecutively in random order. First, the fluoroscopy images were shown, followed by the AI segmented images. Two sessions, with at least a 1-week interval, were scheduled for each of the two datasets and the order of randomization was different for each of the datasets, to avoid a carry-over effect. Operators had not seen any of the images prior the sessions and were blinded regarding clinical data, equipment or originating center. $\%DS_{VE}$ was visually estimated for each target lesion where QCA had been previously measured and operators were asked to provide a specific %DS value of their choosing, rather than a range interval.

Differences were then assessed both in terms of absolute %DS values overall, as well as according to three $\%DS_{QCA}$ Strata, which were defined as <50%, 50%–70%, and >70%.

2.3 | Statistical analysis

Descriptive variables are shown in absolute and relative (percentage) numbers. Quantitative variables are shown in average ± standard deviation or median (interquartile range). To assess for differences in related samples quantitative variables we used the Wilcoxon test (paired samples) or the Friedman test (multiple related samples). To assess differences in qualitative variables we used the Chi-square test.

For further illustrating discrepancies between %DS across the angiography and segmentation groups, we calculated the $\%DS_{VE}/\%DS_{QCA}$ plotted in a scatter graphic in the Y-axis, against the $\%DS_{QCA}$ in the X-axis.

A $p < 0.05$ was used for statistical significance. SPSS 27 was used for analysis.

2.4 | Ethical issues

This study complies with the Declaration of Helsinki and was approved by the local Ethics' Committee.

3 | RESULTS

3.1 | Baseline characteristic

A total of 123 measurements (117 images), from a total of 90 patients, were included. Each operator performed %DS by visual estimation in both the angiography and segmented images,

TABLE 1 Clinical characteristics of included patients.

Factor	*N* ± *SD* or *N* (%)
Age	65 ± 12
Sex (male)	73 (81)
Hypertension	62 (68.9)
Diabetes mellitus	28 (31.1)
Dyslipidemia	55 (61.1)
Smoker (past or present)	50 (55.6)
Chronic coronary syndrome	37 (41.1)
Acute coronary syndrome	53 (58.9)
Revascularization during/after CAG	76 (84.4)
Invasive physiology during procedure	19 (21.1)

TABLE 2 Distribution of target vessel and lesion severity.

Parameter	*N* (%)
Target vessel	
LAD	
Proximal	19
Middle	23
Distal	8
Total	50 (41)
RCA	
Proximal	9
Middle	24
Distal	8
Total	41 (33)
CX	
Proximal	9
Middle	21
Distal	2
Total	32 (26)
Lesion severity	
>70%	22 (18)
50%–70%	58 (47)
<50%	43 (35)

Abbreviations: CX, left circumflex artery; LAD, left anterior descending artery; RCA, right coronary artery.

generating a total of 984 $\%DS_{VE}$ estimates. Most lesions had a $\%DS_{QCA}$ of 50%–70% (Tables 1 and 2).

3.2 | Lesion severity assessment

3.2.1 | Overall results

Lesion severity by $\%DS_{VE}$ was estimated to be higher with angiography than with segmentation. While there were significant differences between $\%DS_{VE}$ and $\%DS_{QCA}$ in both groups, the overall absolute difference in %DS was lower in the segmentation group. Additionally, no statistical difference was found between the segmentation $\%DS_{VE}$ and the $\%DS_{QCA}$ for two of the four operators in the segmentation group, whereas the difference was significant for all operators in the angiography group (Tables 3, Figure 1).

When lesions were grouped in three strata of severity by QCA (<50%, 50%–70%, and >70%), agreement with QCA was generally low, albeit significantly higher (approximately double) with segmentation than angiography, both considering the overall sample and individual operators (Table 4).

3.3 | Results per QCA severity strata

For lesions with $\%DS_{QCA} > 70\%$, there was a statistically significant higher $\%DS_{VE}$ estimation in both the angiography and segmentation groups. In the segmentation group, $\%DS_{VE}$ values were lower and closer to $\%DS_{QCA}$ with no statistically significant differences for one operator. Detailed results are outlined in Tables 3 and 4, Figures 2 and 3, Figures S1–S4.

There was a clear strata agreement between visual estimation and QCA with either angiography (100% agreement) or segmentation (88.6%).

For lesions with $\%DS_{QCA}$ of 50%–70%, there was a very large and significant difference between $\%DS_{VE}$ with angiography, with median estimates of 80% or 90% across all operators and very low rates of strata agreement (0%–19%).

When $\%DS_{VE}$ was undertaken with segmentation, differences with $\%DS_{QCA}$ were small and significantly different only for two of the four operators. The rates of strata agreement between visual estimation with segmentation and QCA were significantly higher when compared with angiography in the overall sample, with individual rates across operators between 39.7% and 51.7%.

For lesions with $\%DS_{QCA} < 50\%$, there were also significantly higher estimates $\%DS_{VE}$ with angiography for all operators and low rates of strata agreement (11.6%–32.6%), but to a lesser degree than in the $\%DS_{QCA}$ 50%–70% strata.

For $\%DS_{VE}$ with segmentation, there was no statistically significant difference with $\%DS_{QCA}$ overall and for two of the four operators, with higher rates of agreement (34.9%–69.8%) than in the angiography group.

3.4 | Operator heterogeneity

There were no significant differences across operators in the $DS_{QCA} > 70\%$ strata for either the angiography or segmentation

Appendix

TABLE 3 diameter stenosis assessed by visual estimation and QCA across operators, lesion severity strata and overall.

$\%DS_{QCA}$ Stratum	Group/ Parameter	$\%DS_{VE}$ - Angiography					$\%DS_{VE}$ - Segmentation					$\%DS_{QCA}$
		Operator 1	Operator 2	Operator 3	Operator 4	All	Operator 1	Operator 2	Operator 3	Operator 4	All	
All cases	%DS	84 +/- 15 90 (80 - 95)	78 +/- 19 80 (50 – 90)	71 +/- 21 70 (50 - 90)	76 +/- 21 80 (70 - 90)	77 +/- 20 80 (70 – 90)	64 +/- 17 60 (50 – 80)	54 +/- 23 60 (40 – 70)	55 +/- 19 50 (40 – 70)	60 +/- 21 60 (40 – 75)	59 +/- 20 60 (40 – 75)	56 +/- 13 56 (46 – 67)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	0,327	0,998	**< 0,001**	**< 0,001**	-
	Inter-Operator Difference**	**< 0,001**				-	**< 0,001**				-	-
> 70%	%DS	95 +/- 4 95 (90 - 99)	96 +/- 4 99 (90 – 99)	94 +/- 5 90 (90 - 99)	96 +/- 4 99 (90 - 99)	95 +/- 4 97 (90 – 99)	82 +/- 12 80 (78 – 90)	82 +/- 12 83 (70 – 90)	79 +/- 14 80 (70 – 90)	82 +/- 11 83 (74 – 90)	81 +/- 12 80 (71 – 90)	74 +/- 3 74 (71 – 76)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**0,008**	**0,016**	0,088	**0,004**	**< 0,001**	-
	Inter-Operator Difference**	0,115*					0,334*					-
50-70%	%DS	88 +/- 8 90 (80 - 95)	83 +/- 12 80 (75 – 90)	76 +/- 15 80 (70 - 90)	84 +/- 12 90 (75 - 90)	83 +/- 13 85 (75 – 90)	66 +/- 14 60 (58 – 80)	60 +/- 17 60 (50 – 70)	59 +/- 13 60 (50 – 70)	67 +/- 13 70 (60 – 75)	63 +/- 15 60 (50 – 75)	60 +/- 5 59 (55 – 64)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**0,002**	0,707	0,837	**< 0,001**	**< 0,001**	-
	Inter-Operator Difference**	**< 0,001**					**< 0,001**					-
< 50%	%DS	71 +/- 17 80 (60 - 80)	62 +/- 20 65 (40 – 80)	53 +/- 17 50 (40 - 70)	57 +/- 20 50 (40 - 70)	61 +/- 20 60 (43 – 80)	52 +/- 13 50 (40 – 60)	33 +/- 12 30 (20 – 40)	40 +/- 10 40 (30 – 50)	40 +/- 17 40 (30 – 50)	41 +/- 15 40 (30-50)	41 +/- 7 43 (38 – 47)
	$\%DS_{VE}$ vs $\%DS_{QCA}$ P Value*	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	0,116	0,538	0,380	-
	Inter-Operator Difference**	**< 0,001**					**< 0,001**					-

Note: *Wilcoxon test paired samples; **Friedman test; significant p values in bold.

Abbreviations: %DSQCA, diameter stenosis by QCA; %DSVE, diameter stenosis by visual estimation; QCA: quantitative coronary angiography.

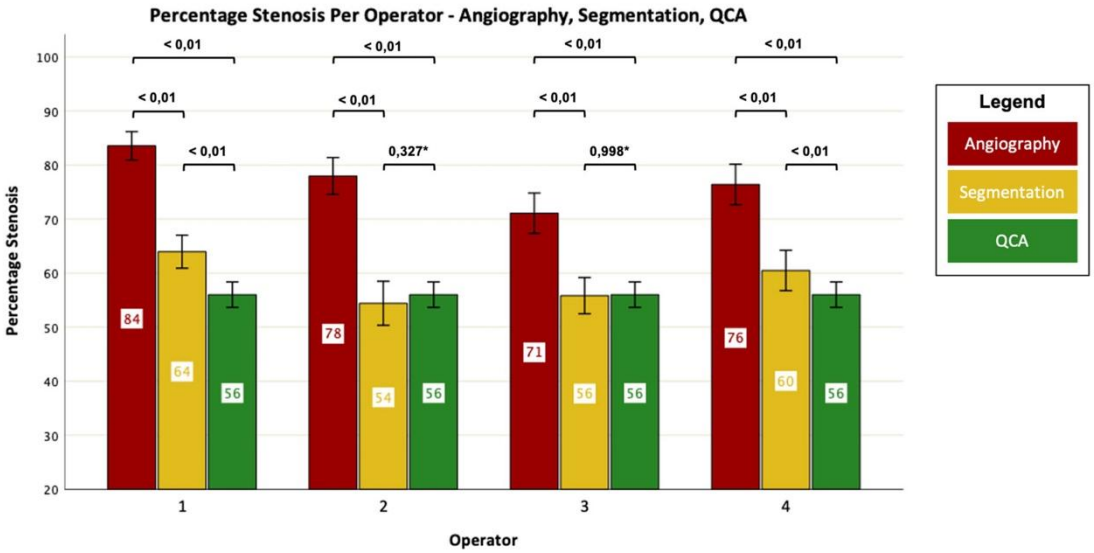


FIGURE 1 $\%DS_{VE}$ according to operator for both angiography and segmentation, compared with $\%DS_{QCA}$. Significant differences are always found between angiography and both other groups, but not between segmentation $\%DS_{VE}$ and $\%DS_{QCA}$ for operators 2 and 3. [Color figure can be viewed at wileyonlinelibrary.com]

groups. Considering the overall sample and the remaining strata, there were significant differences between operators in absolute $\%DS_{VE}$ values both in the angiography and segmentation group. However, in absolute terms, the differences were smaller in the segmentation group. Detailed results are outlined in Tables 3 and 4, Table S1.

When considering agreement by QCA strata rather than absolute %DS values, there were significant differences in the angiography group overall and in the $\%DS_{QCA}$ 50%–70% strata. For the segmentation group, significant differences were only found in the $\%DS_{QCA}$ < 50% strata.

4 | DISCUSSION

4.1 | Main findings

When considering QCA as reference, operators generally tended to overestimate lesion severity in angiography images, but much less so in segmented images. Indeed, the overall rate of agreement in severity strata was approximately double in the segmentation group.

For lesions with a $\%DS_{QCA}$ > 70%, visual estimation was usually in agreement with QCA in both the fluoroscopy and segmented datasets. These were lesions operators deemed as very severe, as is evident by mean $\%DS_{VE}$ > 80%–90% in both groups. However, even in this stratum, absolute $\%DS_{VE}$ values were less discrepant with $\%DS_{QCA}$ in the segmentation group.

When considering lesions with $\%DS_{QCA}$ of 51%–69%, overestimation was very frequent and pronounced in the angiography group, but not in the segmentation group. Indeed, the mean $\%DS_{VE}$ difference in the angiography group exceeded that of QCA by approximately 15–30 percentage points versus 1–7 percentage points in the segmentation group.

For the least severe lesions, the findings are somewhat similar to those in the intermediate group, albeit to a lesser degree in all respects: the discrepancy between %DS is not as large and the rates of agreement not as low, in the angiography group. In the segmentation group, differences in %DS are small and overall not significant, with much higher rates of agreement with the $\%DS_{QCA}$ lesion severity strata.

The scatterplots (Figure 2, Figures S1 and S2) clearly illustrate these findings, as the discrepancy outside $\%DS_{QCA}$ > 70% stratum is much more evident in the angiography group than the segmentation group. Additionally, when values are plotted based on %DS ratios, it is clear that the trendline for angiography only closely matches that of QCA in the $\%DS_{QCA}$ > 70% stratum, whereas in the segmentation group it is constantly much closer to 1 (Figure 3, Figures S3 and S4).

Significant differences across operators were found for both the angiography and segmentation groups. However, the absolute differences in %DS were much lower in the latter, with overall rates of agreement across $\%DS_{QCA}$ severity strata not significantly different in the segmentation group.

The results of our study therefore suggest that visualization of segmented images seems to render visual estimation of stenosis severity more objective, significantly reducing the tendency to overestimate, while possibly reducing operator heterogeneity as well (Figure 4).

Appendix

TABLE 4 agreement between %DSQCA strata and %DSVE strata between angiography and segmentation, across operators and overall.

$\%DS_{QCA}$ Strata/Group		Angiography All	Segmentation All	Angiography Operator 1	Segmentation Operator 1	Angiography Operator 2	Segmentation Operator 2	Angiography Operator 3	Segmentation Operator 3	Angiography Operator 4	Segmentation Operator 4
All cases	$\%DS_{VE}/\%DS_{QCA}$ Strata Agreement (n/%)	148 (30,1)	297 (60,4)	27 (22,0)	65 (52,8)	35 (28,5%)	79 (64,2%)	47 (38,2)	80 (65,0)	39 (31,7)	73 (59,3)
	Angiography vs Segmentation* P Value	**< 0,001**		**< 0,001**		**0,007**		**< 0,001**		**< 0,001**	
> 70%	$\%DS_{VE}/\%DS_{QCA}$ Strata Agreement (n/%)	88 (100)	78 (88,6%)	22 (100)	29 (90,9)	22 (100)	19 (86,4)	22 (100)	18 (81,8)	22 (100)	21 (95,5)
	Angiography vs Segmentation* P Value	NA		NA		NA		NA		NA	
50-70%	$\%DS_{VE}/\%DS_{QCA}$ Strata Agreement (n/%)	17 (7,3)	107 (46,1)	0 (0)	30 (51,7)	2 (3,4%)	23 (39,7%)	11 (19)	32 (55,2)	4 (6,9)	22 (37,9)
	Angiography vs Segmentation* P Value	**0,009**		NA		0,243		**0,008**		0,113	
< 50%	$\%DS_{VE}/\%DS_{QCA}$ Strata Agreement (n/%)	43 (25,0)	112 (65,1)	5 (11,6)	15 (34,9)	11 (25,6)	37 (86,0)	14 (32,6)	30 (69,8)	13 (30,2)	30 (69,8)
	Angiography vs Segmentation* P Value	**< 0,001**		**0,001**		0,590		**0,003**		**0,004**	

Note: *Chi-square test; significant *p* values in bold.

Abbreviations: %DSQCA, diameter stenosis by QCA; %DSVE, diameter stenosis by visual estimation; NA, not applicable; QCA: quantitative coronary angiography.

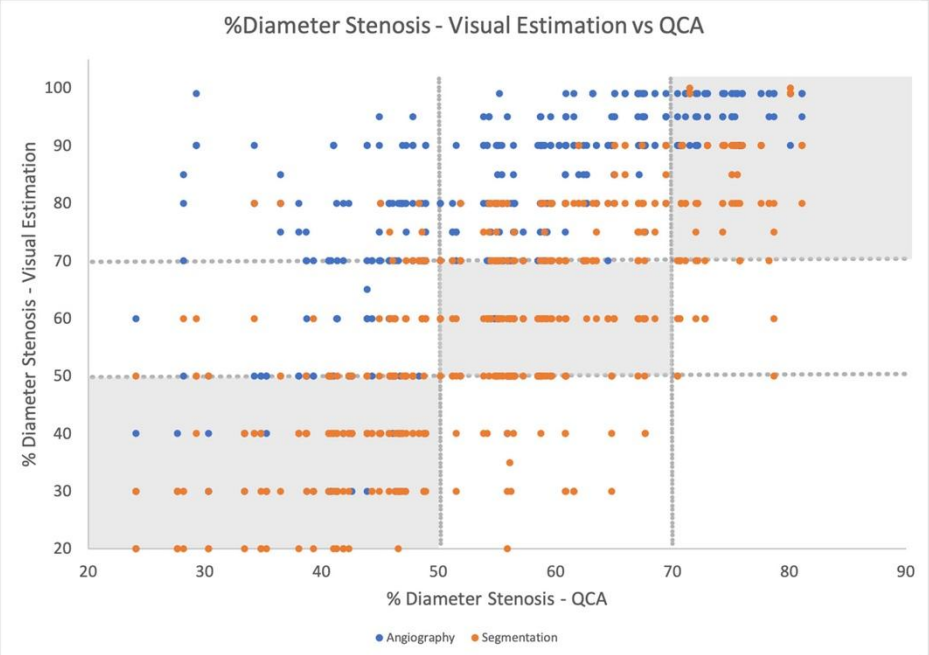


FIGURE 2 Combined scatterplot of angiography and segmentation $\%DS_{VE}$ vs. $\%DS_{QCA}$. The difference between the two $\%DS_{VE}$ is visually clear. [Color figure can be viewed at wileyonlinelibrary.com]

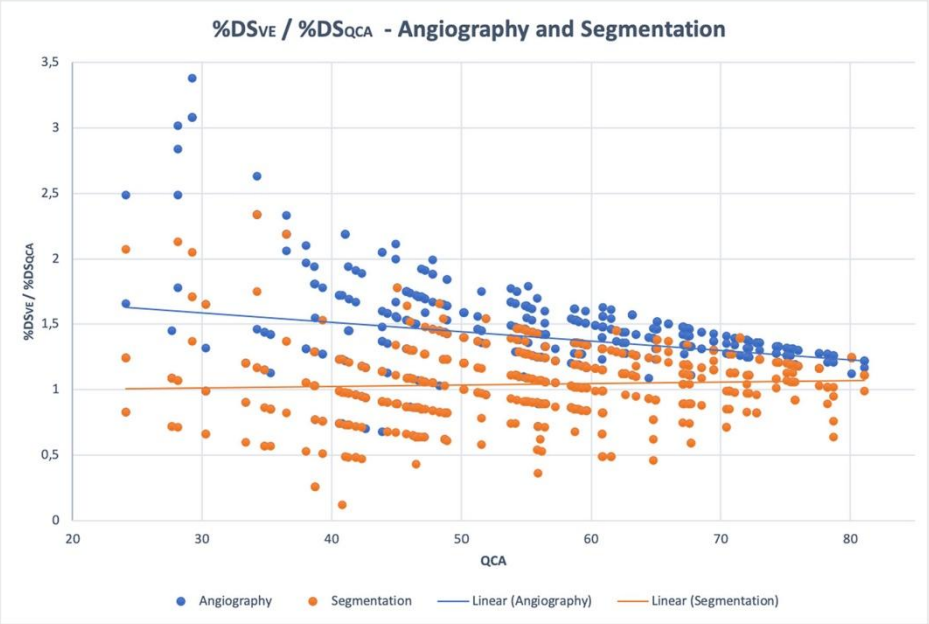


FIGURE 3 Combined scatterplots of angiography and segmentation $\%DS_{VE}/\%DS_{QCA}$ plotted in the Y-axis vs. $\%DS_{QCA}$ plotted in the X-axis. The difference between the two $\%DS_{VE}$ is visually clear. [Color figure can be viewed at wileyonlinelibrary.com]

Appendix

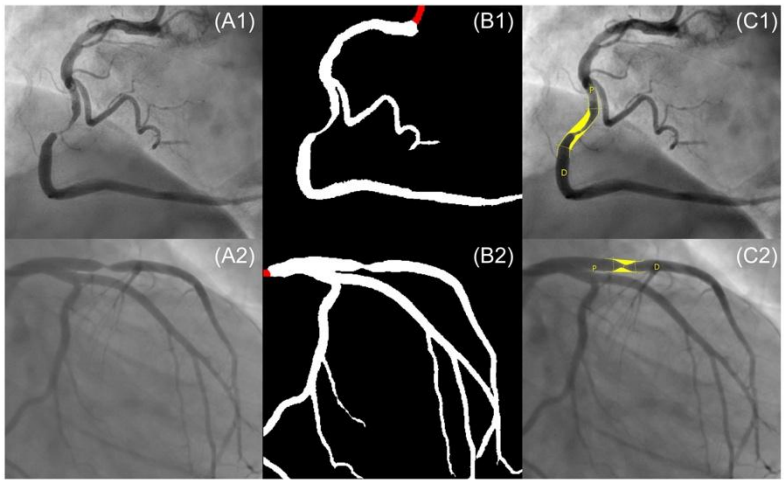


FIGURE 4 Image examples. The upper three images represent a lesion with $\%DS_{QCA}$ of 78% (C1), with average operator $\%DS_{VE}$ of 90% based on angiography (A1) vs. 80% based on segmentation (B1). The lower three images represent a lesion with $\%DS_{QCA}$ of 63% (C2), with average operator $\%DS_{VE}$ of 88% based on angiography (A2) vs. 69% based on segmentation (B2). [Color figure can be viewed at wileyonlinelibrary.com]

The reason for these findings is not entirely clear. The most likely explanation is that segmented images display a stenosed segment in homogeneous fashion, with the transition between artery (white) and background (black) very clearly visible. In contrast, angiography images display the artery in shades of gray, with stenosed regions with poorer contrast filling and less clear demarcation of artery and background. As a result, the human eye seems more prone to underestimating the actual size of the artery lumen in these segments, thereby estimating the stenosis as more severe.

4.2 | Other studies in the literature

The above-mentioned findings for stenosis severity assessment in fluoroscopy images have long been described in the literature for more than 30 years[2] and continue to be found today across the world, even in very large cohorts.[4–6] The idea that visual estimation is inaccurate was one of the triggers that spawned the development of QCA analysis and software. Indeed, core-lab QCA in PCI trials continues to be advocated for in international research consortiums[21] or scientific societies.[1] However, in clinical practice, the estimation of stenosis severity continues to be undertaken mostly by visual estimation.[6] The fact that correlation between QCA and FFR has not been demonstrated to be superior to that of visual estimation,[5] or that clinical trials often don't require QCA measurements for including patients,[22,23] further reinforces this approach.

However, discrepancies between QCA and visual estimation are likely to be of clinical relevance. In a subanalysis of the PROMISE trial of patients who underwent CAG, patients without obstructive disease by QCA, but classified as such by visual estimation, had lower event rates than those with obstructive disease by both criteria.[7] As operators perceive these patients as high risk, revascularization—sometimes potentially unwarranted—may become more likely. And, indeed, the decision to proceed with physiology, which has been shown to be superior to angiography-guided revascularization[8] and is recommended in current guidelines,[9,10] implies that the lesion be classified as intermediate rather than severe. Thus, the persistence of the visual estimation approach and the resulting overestimation tendency may arguably be contributing to the seemingly low usage rates (6%–13%) of physiology in cath labs.[10,11,24] Last, the growing relevance of imaging for decision-making in the cath lab further adds to this issue, as recently highlighted in a major clinical trial,[25] further points to the importance of avoiding the so-called "oculostenotic reflex."

4.3 | Practical clinical implications

In keeping with all of the above, the findings of our study may have an important clinical application: if visual estimation of segmented images makes operators less prone to overestimating lesions, unwarranted revascularization may become less likely and the use of physiology/ischemia/imaging testing may increase. The advantage of AI-based segmentation is that it is fast, fully automatic and requires no human input, other than the image itself. The operator can be simply and immediately be exposed to the segmented image almost effortlessly. Conversely, QCA software is semi-automatic, as it requires the manual annotation of the region of interest and may require manual adjusting of vessel contours. The simple fact that it is seldom employed in clinical practice further emphasizes the point.

If these exploratory results are confirmed in subsequent studies, AI-based segmentation of the coronary tree may become a relevant tool capable of improving CAG interpretation, hopefully contributing to improved outcomes, without complicating or lengthening the procedure.

4.4 | Other studies in AI applied to CAG

Few medical papers regarding the application of AI to CAG have been published in medical literature so far. Two studies focused on segmentation alone, with accurate results.[14,17] We too have previously published our results regarding the development of AI models capable of highly accurate segmentation, encompassing more than 500 images as a whole.[18,19] The largest published study to date focused on developing models capable of segmentation with high accuracy as well, with the added feature of lesion type recognition (i.e., calcium, thrombus, among others), thus potentially enabling future clinical application.[16] Last, a small exploratory study also tested the hypothesis of estimating FFR from CAG images using AI, but the study population was small and the group has published no further data in medical journals.[26]

Thus, our study is one of the few and first in the field of AI applied to CAG with potential clinical implications.

4.5 | Limitations

Our study has important limitations. The dataset size is small and composed predominantly of 51%–70% lesions, with a small amount of lesions >70%, as assessed by QCA. The fact that lesion assessment was based on the evaluation by only four operators is also a limitation. We did not perform sub analysis regarding clinical context, risk factors or lesion characteristics (such as presence of calcium), due to the small sample size. The concept of the study is thus exploratory in nature, requiring external validation in further works.

4.6 | Future directions

We are currently working not only in improving the segmentation capabilities of our model, but also exploring other applications of AI to CAG. Automatic anatomical interpretation, lesion severity based on auto-QCA and integration with physiology are areas of active research. We aim to develop an integrated suite of AI models capable of enhancing CAG interpretation, improving decision-making and hopefully patient outcomes.

5 | CONCLUSIONS

When considering QCA as reference, the visual estimation of stenosis severity was much less discrepant with automatic AI-segmented images than angiography images, with reduced inter-operator discrepancies as well. Operators were essentially prone to over-estimation of lesion severity, especially of intermediate lesions (% DS_{QCA} of 50%–70%). The visual assessment of coronary lesions with segmented images may, therefore, lead to a lower likelihood of unwarranted revascularization, while potentially increasing the use of functional assessment, as recommended by current guidelines.

ACKNOWLEDGMENTS

This work is supported by Cardiovascular Center of the University of Lisbon, INESC-ID/Instituto Superior Técnico, University of Lisbon.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Detailed full-scale study data cannot currently be made publicly available due to limitations imposed by national data protection regulations, as this is a retrospective study and no informed consent was obtainable regarding this particular analysis. Both our research team and others in the national scientific community are working to develop a framework where such would be possible. However, independent replication of our analysis is possible, given that the detailed description of our experimentations and relevant code is publicly available.

ORCID

Miguel Nobre Menezes http://orcid.org/0000-0001-8363-0363
Cláudio Guerreiro http://orcid.org/0000-0002-4825-0979
João Pedro Guedes http://orcid.org/0000-0002-7145-2843
Manuel Oliveira-Santos http://orcid.org/0000-0002-5980-6880

REFERENCES

1. Suzuki N, Asano T, Nakazawa G, et al. Clinical expert consensus document on quantitative coronary angiography from the Japanese Association of Cardiovascular Intervention and Therapeutics. *Cardiovasc Interv Ther*. 2020;35:105-116. doi:10.1007/S12928-020-00653-7
2. Beauman GJ, Vogel RA. Accuracy of individual and panel visual interpretations of coronary arteriograms; implications for clinical decisions. *JACC*. 1990;16:108-113. doi:10.1016/0735-1097(90)90465-2
3. Fischer JJ, Samady H, McPherson JA, et al. Comparison between visual assessment and quantitative angiography versus fractional flow reserve for native coronary narrowings of moderate severity. *Am J Cardiol*. 2002;90:210-215. doi:10.1016/S0002-9149(02)02456-6
4. Nallamothu BK, Spertus JA, Lansky AJ, et al. Comparison of clinical interpretation with visual assessment and quantitative coronary angiography in patients undergoing percutaneous coronary intervention in contemporary practice: the assessing angiography (A2) project. *Circulation*. 2013;127:1793-1800. doi:10.1161/CIRCULATIONAHA.113.001952
5. Adjedj J, Xaplanteris P, Toth G, et al. Visual and quantitative assessment of coronary stenoses at angiography versus fractional flow reserve: the impact of risk factors. *Circ Cardiovasc Imaging*. 2017;10(7):e006243. doi:10.1161/CIRCIMAGING.117.006243
6. Zhang H, Mu L, Hu S, et al. Comparison of physician visual assessment with quantitative coronary angiography in assessment of stenosis severity in China. *JAMA Intern Med*. 2018;178:239-247. doi:10.1001/JAMAINTERNMED.2017.7821

Appendix

7. Shah R, Yow E, Jones WS, et al. Comparison of visual assessment of coronary stenosis with independent quantitative coronary angiography: findings from the prospective multicenter imaging study for evaluation of chest pain (PROMISE) trial. *Am Heart J*. 2017;184:1-9. doi:10.1016/J.AHJ.2016.10.014
8. Tonino PAL, De Bruyne B, Pijls NHJ, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N Engl J Med*. 2009;360:213-224. doi:10.1056/NEJMoa0807611
9. Lawton JS, Tamis-Holland JE, Bangalore S, et al. 2021 ACC/AHA/SCAI guideline for coronary artery revascularization: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2022;145:18. doi:10.1161/CIR.0000000000001038/FORMAT/EPUB
10. Neumann FJ, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS guidelines on myocardial revascularization. *Eur Heart J*. 2019;40:87-165. doi:10.1093/EURHEARTJ/EHY394
11. Petraco R, Park JJ, Sen S, et al. Hybrid iFR-FFR decision-making strategy: implications for enhancing universal adoption of physiology-guided coronary revascularisation. *EuroIntervention*. 2013;8:1157-1165. doi:10.4244/EIJV8I10A179
12. Tebaldi M, Biscaglia S, Fineschi M, et al. Evolving routine standards in invasive hemodynamic assessment of coronary stenosis. *JACC Cardiovasc Interv*. 2018;11:1482-1491. doi:10.1016/j.jcin.2018.04.037
13. Pyxaras SA, Harmel EK. Artificial intelligence for the assessment of coronary artery disease—the Dawn of a new era? *Int J Cardiol*. 2021;343:3-4. doi:10.1016/J.IJCARD.2021.09.002
14. Wang L, Liang D, Yin X, et al. Coronary artery segmentation in angiographic videos utilizing spatial-temporal information. *BMC Med Imaging*. 2020;20:110. doi:10.1186/S12880-020-00509-9
15. Liang D, Qiu J, Wang L, et al. Coronary angiography video segmentation method for assisting cardiovascular disease interventional treatment. *BMC Med Imaging*. 2020;20:65. doi:10.1186/S12880-020-00460-9
16. Du T, Xie L, Zhang H, et al. Training and validation of a deep learning architecture for the automatic analysis of coronary angiography. *EuroIntervention*. 2021;17:32-40. doi:10.4244/EIJ-D-20-00570
17. Yang S, Kweon J, Roh J-H, et al. Deep learning segmentation of major vessels in X-ray coronary angiography. *Sci Rep*. 2019;9:16897. doi:10.1038/s41598-019-53254-7
18. Nobre Menezes M, Lourenço-Silva J, Silva B, et al. Development of deep learning segmentation models for coronary X-ray angiography: quality assessment by a new global segmentation score and comparison with human performance. *Rev Port Cardiol*. 2022;41(12):1011-1021. doi:10.1016/J.REPC.2022.04.001
19. Nobre Menezes M, Silva JL, Silva B, et al. Coronary X-ray angiography segmentation using artificial intelligence: a multicentric validation study of a deep learning model. *Int J Cardiovasc Imaging*. 2023;2023:1-12. doi:10.1007/S10554-023-02839-5
20. Silva JL, Menezes MN, Rodrigues T, Silva B, Pinto FJ, Oliveira AL. Encoder-decoder architectures for clinically relevant coronary artery segmentation. Bansal, M.S., et al., eds. *Computational Advances in Bio and Medical Sciences*. ICCABS 2021. Lecture Notes in Computer Science, vol 13254. Springer. 2022. doi:10.1007/978-3-031-17531-2_6
21. Garcia-Garcia HM, McFadden EP, Farb A, et al. Standardized end point definitions for coronary intervention trials: the academic research Consortium-2 consensus document. *Circulation*. 2018;137:2635-2650. doi:10.1161/CIRCULATIONAHA.117.029289/-/DC1
22. Perera D, Clayton T, O'Kane PD, et al. Percutaneous revascularization for ischemic left ventricular dysfunction. *N Engl J Med*. 2022;387:1351-1360. doi:10.1056/NEJMOA2206606/SUPPL_FILE/NEJMOA2206606_DATA-SHARING.PDF
23. De Bruyne B, Pijls NHJ, Kalesan B, et al. Fractional flow reserve–guided PCI versus medical therapy in stable coronary disease. *N Engl J Med*. 2012;367:991-1001. doi:10.1056/NEJMoa1205361
24. Raposo L, Gonçalves M, Roque D, et al. Adoption and patterns of use of invasive physiological assessment of coronary artery disease in a large cohort of 40821 real-world procedures over a 12-year period. *Rev Port Cardiol*. 2021;40:771-781. doi:10.1016/J.REPC.2021.01.010
25. Koo B-K, Hu X, Kang J, et al. Fractional flow reserve or intravascular ultrasonography to guide PCI. *N Engl J Med*. 2022;387:779-789. doi:10.1056/NEJMOA2201546
26. Roguin A, Abu Dogosh A, Feld Y, Konigstein M, Lerman A, Koifman E. Early feasibility of automated artificial intelligence angiography based fractional flow reserve estimation. *Am J Cardiol*. 2021;139:8-14. doi:10.1016/j.amjcard.2020.10.022

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Nobre Menezes M, Silva B, Silva JL, et al. Segmentation of X-ray coronary angiography with an artificial intelligence deep learning model: Impact in operator visual assessment of coronary stenosis severity. *Catheter Cardiovasc Interv*. 2023;102:631-640. doi:10.1002/ccd.30805

JOURNAL OF Invasive Cardiology

ORIGINAL CONTRIBUTION

PEER REVIEWED

Coronary Physiology Instantaneous Wave-Free Ratio (iFR) Derived From X-Ray Angiography Using Artificial Intelligence Deep Learning Models: A Pilot Study

Miguel Nobre Menezes, MD, MSc[1,2]; João Lourenço Silva, MSc[3,4]; Beatriz Silva, MD[1,2]; Rita Marante de Oliveira, BSc[5]; Tiago Rodrigues, MD[1,2]; Arlindo L. Oliveira, PhD[3,4]; Fausto J. Pinto, PhD[1,2]

Keywords
Deep Learning
Artificial Intelligence
Machine Learning
Coronary Angiography
Coronary Artery Disease
Percutaneous Coronary Intervention
Coronary Physiology
Fractional Flow Reserve
Instantaneous Wave-Free Ratio (IFR)

March 2024
ISSN 1557-2501

Index J INVASIVE CARDIOL 2024;36(3). doi:10.25270/jic/23.00285. Epub February 22, 2024.

Abstract

Objectives. Coronary angiography (CAG)-derived physiology methods have been developed in an attempt to simplify and increase the usage of coronary physiology, based mostly on dynamic fluid computational algorithms. We aimed to develop a different approach based on artificial intelligence methods, which has seldom been explored.

Methods. Consecutive patients undergoing invasive instantaneous free-wave ratio (iFR) measurements were included. We developed artificial intelligence (AI) models capable of classifying target lesions as positive (iFR ≤ 0.89) or negative (iFR > 0.89). The predictions were then compared to the true measurements.

Results. Two hundred-fifty measurements were included, and 3 models were developed. Model 3 had the best overall performance: accuracy, negative predictive value (NPV), positive predictive value (PPV), sensitivity, and specificity were 69%, 88%, 44%, 74%, and 67%, respectively. Performance differed per target vessel. For the left anterior descending artery (LAD), model 3 had the highest accuracy (66%), while model 2 the highest NPV (86%) and sensitivity (91%). PPV was always low/modest. Model 1 had the highest specificity (68%). For the right coronary artery, model 1 had an accuracy of 86%, NPV was 97%, and specificity was 87%, but all models had low PPV (maximum 25%) and low/modest sensitivity (maximum 60%). For the circumflex, model 1 performed best: accuracy, NPV, PPV, sensitivity, and specificity were 69%, 96%, 24%, 80%, and 68%, respectively.

Conclusions. We developed 3 AI models capable of binary iFR estimation from CAG images. Despite modest accuracy, the consistently high NPV is of potential clinical significance, as it would enable avoidance of further invasive maneuvers after CAG. This pivotal study offers proof of concept for further development.

Introduction

The use of invasive coronary physiology has been extensively studied and is clearly recommended in clinical guidelines today.[1,2] The most widely studied index is the Fractional Flow Reserve (FFR). Three major trials established its use in selecting lesions where revascularization had an additional benefit to medical therapy,[3-5] alongside large observational data.[6] More recently, another index gained importance: instantaneous free-wave ratio (iFR). It was initially studied using FFR as the gold standard, with high accuracy.[7,8] Our own experience showed similar results.[9] Compared to FFR, iFR achieved non-inferiority in 2 major outcome trials,[10,11] including over a 5-year follow-up period.[12]

Appendix

Coronary physiology is vastly underused, ranging from 7% to 13% of procedures.[13,14] The risk of complications, time consumption, and cost are likely reasons for this. Thus, calculating a physiological index (either FFR and/or iFR) digitally from coronary angiography (CAG) images is desirable, as it would bypass these limitations and potentially broaden physiology adoption. While this has already been achieved with several different software approaches relying on 3-dimensional (3D) vessel reconstruction and complex fluid dynamics computational algorithms,[15-17] some limitations remain.

In medicine, artificial intelligence (AI) has shown great potential, especially in imaging, as several publications have demonstrated excellent results with regards to electrocardiogram (ECG),[18] echocardiography,[19,20] and magnetic resonance imaging (MRI).[21,22] However, the use of AI regarding physiology estimation derived from CAG has seldom been explored.[23,24] The potential advantages of using AI for this task could be either fully automating the process with minimal user input and/or improving the reliability of current systems, either functioning as a standalone approach or an added layer to existing software.

In this pilot study, we aimed to develop fully automated AI models capable of binary iFR lesion classification from CAG images alone, using measured invasive iFR as a reference.

Methods

Inclusion criteria. We conducted a single-center retrospective selection of consecutive patients over a 3-year period (2017-2019), who had undergone both CAG and invasive physiology assessment with iFR (Philips Volcano System), regardless of clinical context (ie, both acute and chronic coronary syndrome).

Exclusion criteria. We excluded cases where any of the following applied:

1. Imaging criteria:
    1. Patients with cardiac devices or other sources of potential imaging artifacts overlapping with the coronary tree image
    2. Poor image quality
    3. Unsuccessful segmentation with AI models
    4. Unclear individualization of lesion outline with overlapping vessels
2. Clinical criteria
    1. History of coronary artery bypass grafting (CABG) or valvular intervention (surgical or percutaneous)
    2. Significant left heart valvular disease (severe aortic stenosis or regurgitation, severe mitral regurgitation, moderate or severe mitral stenosis, moderate valvular disfunction in both aortic and mitral valves)
    3. Target culprit vessel culprit of acute coronary syndrome (ACS)
    4. Target vessel non-culprit of ACS in patient with ST-segment elevation myocardial infarction (STEMI) within 48 hours of presentation
    5. Previous transmural myocardial infarction in target-vessel
    6. Chronic total occlusion not previously treated with percutaneous coronary intervention (PCI) in any vessel
    7. Left ventricular systolic disfunction, defined as an ejection fraction < 50%
    8. Cardiogenic shock
    9. Hemodynamic instability
    10. Left main lesions

All the exclusion criteria were selected because of their potential impact on physiology index measurements, which could generate confoundment in the training process of an AI algorithm.

AI models development. We have previously trained AI models capable of fully automatic AI segmentation of CAG images.[25-28] We used these models to segment all the CAGs of included patients. Images were then annotated with the target vessel and the location of the pressor sensor wires used to measure iFR using a single telediastolic frame where the target vessel was outlined. All available projections were annotated for each case.

Preliminary testing for AI models showed that performance was better when a single image that best outlined the target vessel was used, rather than using a combination of different projections. As a result, we proceeded to train our models as such. The model was given only the CAG image, both original (greyscale) and manually annotated after automatic segmentation **(Figure 1)**. No further details, clinical or otherwise (such as the target vessel), were provided. The models were trained to binarily classify whether a given target had an iFR less than or equal to 0.89 or greater than 0.89, henceforth defined as a positive or negative iFR result, respectively.

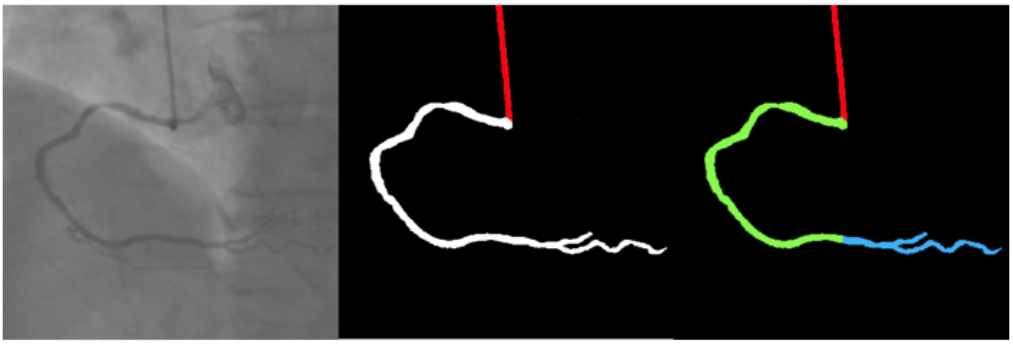


Figure 1. Original diagnostic CAG image **(left)**, automatic AI-based segmentation **(center)** and annotated image with coronary vessel proximal to wire sensor position in green and distal in blue **(right)**. *AI = artificial intelligence; CAG = coronary angiography.*

A total of 3 models were trained. The first model used as input the sequence of diameters along the main branch of the analyzed vessel, automatically computed by a preprocessing algorithm. This sequence was then processed by a transformer encoder,29 with a classification head on top, which inherently took into account the sequential nature of the data and allowed prediction of the iFR value at any given point within the artery. During training, the loss wa only computed at the point for which the ground-truth iFR was available, using a Cross-Entropy Loss, weighted by the inverse of each class's frequency. The Transformer encoder has 6 layers, 8 heads, a hidden dimension of 768, Gaussian Error Linear Unit (GELU) activation, 0.3 dropout, a maximum of 1024 linear positional embeddings, and is followed by a linear classification layer with 0.3 dropout.

Models 2 and 3 are Convolutional Neural Networks (CNNs), which took as input the concatenation of the single-channel angiography image and its segmentation, with 1 channel per class. Model 2 used a simple Cross-Entropy Loss, and model 3 used a Cross-Entropy Loss weighted by the inverse of each iFR class's frequency, aimed to mitigate the negative effects of class imbalance. The chosen CNN was an EfficientNet-B5,30 with which we had already had success in previous work,25–28 followed by a linear classification layer.

Theory suggests model 1 should be much more suitable for this task. Unlike model 1, the CNNs in models 2 and 3 cannot inherently consider the 1-dimensional characteristic of the artery and must learn to do so during training, possibly requiring more data to achieve the same level of performance. Additionally, the transformer in model 1 operates on a much lower-dimension input space than CNNs, making the former's task theoretically much easier. Finally, since the location of the iFR predicted by models 2 and 3 is directly tied to the input segmentation, they require an additional inference per iFR prediction.

Performance assessment and statistical analysis. Descriptive variables are shown in absolute and relative (percentage) numbers. Quantitative variables are shown in average ± standard deviation (if normally distributed) or median (interquartile range) if non-normally distributed. The chi-square test was used for statistically comparing the binary classification of measured iFR vs that of the models. A *P*-value of .05 was used for statistical significance.

The results of the models' classification of target lesions as either iFR positive or negative were compared to those of the true (ie, real) invasive iFR measurements, as follows:

- True positive (TP): both estimated and real iFR were positive
- False positive (FP): positive estimated iFR and negative real iFR
- True negative (TN): both estimated and real iFR were negative
- False negative (FN): negative estimated iFR and positive real iFR

Using this classification, the following parameters were calculated:

- Accuracy: ([TP + TN] / [TP + TN + FP + FN])
- Sensitivity: TP / (TP + FN)
- Specificity: TN / (TN + FP),
- Positive predictive value (PPV): TP / (TP + FP)
- Negative predictive value (NPV): TN / (TN + FN)

To ensure proper evaluation of the results, the models could not be tested on data already seen during training. Hence, we used a cross-validation split at the patient level into 10 subsets, retaining the relative distribution of target vessel and iFR classification per split. The models' iFR classification in each subset was then undertaken using neural networks trained exclusively on the remaining data. This enabled the assessment of the models' performance for the whole cohort, whereas

Appendix

the usual splitting in a fixed train/test datasets would have resulted in a much smaller testing group, limiting the ability to test the models' performance. We have successfully used this approach in the past, when developing our segmentation models.25-27 SPSS 27 (IBM) was used for analysis.

Ethical issues. This study complies with the Declaration of Helsinki and was approved by the local Ethics' Institutional Review Board.

Results

Baseline characteristics. A total of 334 patients were screened. After applying the exclusion criteria, a total of 250 measurements, from a total of 223 patients, were included **(Figure 2, Table 1)**. Most lesions had an iFR greater than 0.89. There was a large imbalance between positive and negative iFR lesions in the right and circumflex coronary arteries subgroups. The difference was much less pronounced in the LAD measurements **(Table 2)**.

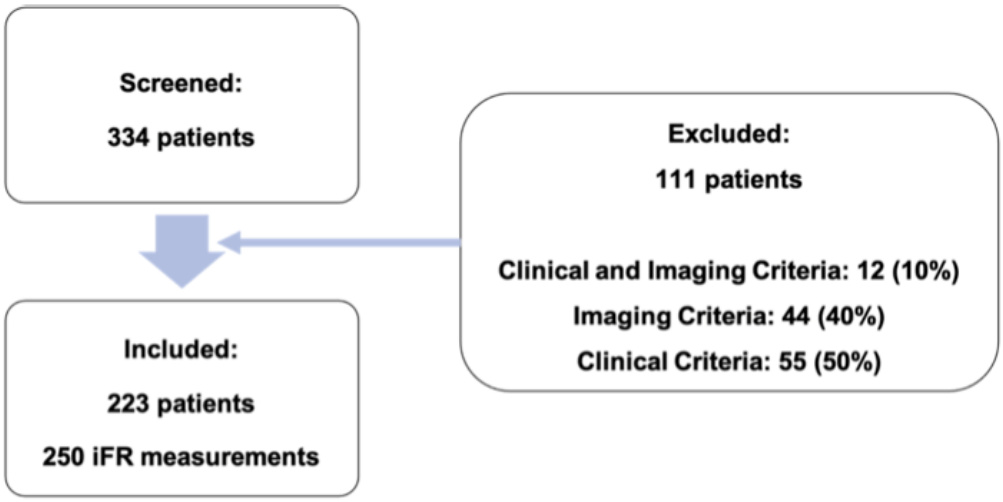


Figure 2. Inclusion flowchart. *iFR = instantaneous free-wave ratio.*

TABLE 1. CLINICAL CHARACTERISTICS OF INCLUDED PATIENTS

Parameter	N +/- SD or N (%)
Age	68 ± 11
Sex (male)	148 (66.37%)
Hypertension	180 (80.72%)
Diabetes mellitus	96 (43.05%)
Dyslipidemia	132 (59.19%)
Smoker (past or present)	89 (39.91%)
Chronic coronary syndrome	148 (66.37%)
Acute coronary syndrome	74 (33.18%)

TABLE 2. IFR RESULTS OVERALL AND STRATIFIED PER ARTERY

	N (%)	iFR (mean ± SD)	iFR ≤ 0.89 (N / %)	iFR > 0.89 (N / %)
Total	250 (100%)	0.91 ± 0.006	65 (26.0%)	185 (74.0%)
LAD	129 (51.6%)	0.88 ± 0.009	55 (42.5%)	74 (57.4%)
RCA	76 (30.4%)	0.95 ± 0.009	5 (6.6%)	71 (93.4%)
Cx	45 (18.0%)	0.96 ± 0.009	5 (33.33%)	40 (88.9%)

Cx = circumflex artery; iFR = instantaneous free-wave ratio; LAD = left anterior descending artery; RCA = right coronary artery.

Physiology Prediction of AI Models

Overall results. The difference between measured iFR classification and that of the AI models was not statistically significant for model 1 ($P = .063$), whereas models 2 and 3 differed significantly ($P < .001$). Models 1 and 3 classified most lesions as negative, contrary to model 2, with only model 1 distributing classifications in similar proportions to those of the ground truth. Agreement was strongest for negative lesions. Details are presented in **Table 3**.

TABLE 3. IFR CLASSIFICATION OF LESIONS AS MEASURED VS AS PER EACH MODEL PREDICTION FOR ALL CASES

Classification (All Cases)		Model 1		Model 2		Model 3		Total
		≤ 0.89	> 0.89	≤ 0.89	> 0.89	≤ 0.89	> 0.89	
Measured iFR	≤ 0.89	24	41	53	12	48	17	65
	> 0.89	46	139	92	93	61	124	185
Total		70	180	145	105	109	141	250

iFR = instantaneous free-wave ratio.

All models performed with modest accuracy, with the worst results for model 2 (58%) and the best for model 3 (nearing 70%), followed closely by model 1 (65%). NPV was high for all models, performing close to 80% or 90%, in contrast with PPV. Models 2 and 3 had the highest sensitivity, whereas model 1 had the highest specificity. Details are presented in **Figure 3**.

Appendix

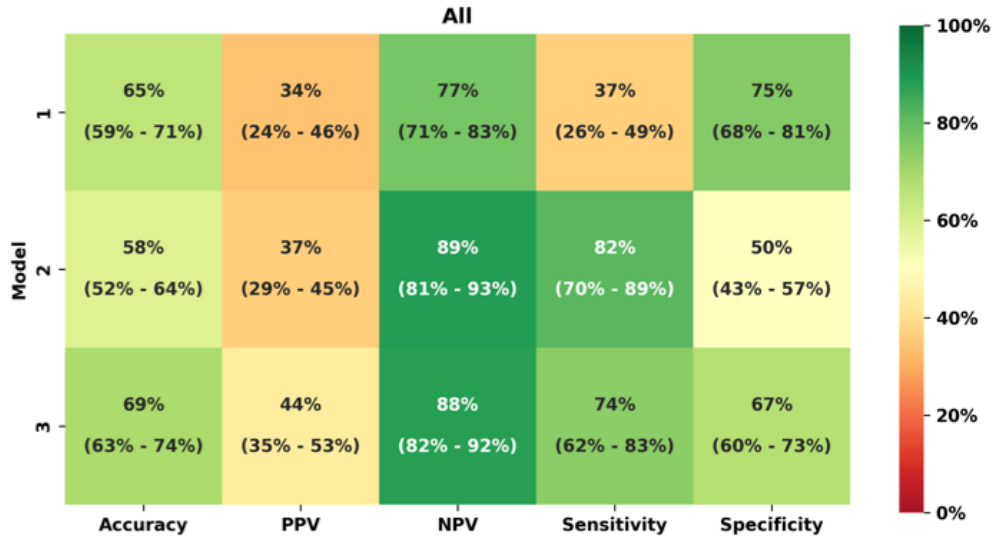


Figure 3. Performance metrics (with 95% CI) of each AI model regarding iFR binary classification using the measured iFR classification as reference, for all iFR measurements across all target vessels. Accuracy, PPV, NPV, sensitivity, and specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values. *AI = artificial intelligence; iFR = instantaneous free-wave ratio; NPV = negative predictive value; PPV = positive predictive value.*

Left anterior descending (LAD) lesions. The models' classification of iFR as compared to the measured iFR was not statistically significant for model 1 ($P = .854$), whereas models 2 and 3 differed significantly ($P < .001$). Models 1 and 2 classified most lesions as negative, where agreement was more common. Model 3 classified most lesions as positive. Details are presented in **Table 4**.

TABLE 4. IFR CLASSIFICATION OF LESIONS AS MEASURED VS AS PER EACH MODEL PREDICTION FOR LEFT ANTERIOR DESCENDING ARTERY (LAD) CASES.

Classification (LAD)		Model 1		Model 2		Model 3		Total
		≤0.89	>0.89	≤0.89	>0.89	≤0.89	>0.89	
Measured iFR	≤0.89	17	38	50	5	43	12	55
	>0.89	24	50	43	31	32	42	74
Total		41	88	93	36	75	54	129

iFR = instantaneous free-wave ratio.

Models 2 and 3 displayed modest accuracy, with better results for the latter, while model 1's performance barely surpassed 50%. The NPV and sensitivity was high or very high for model 2, nearing 90%, while model 3's performance was 78% for both. Model 1 performed reasonably only for specificity (68%). The PPV was above 50% for models 2 and 3. Details are presented in **Figure 4**.

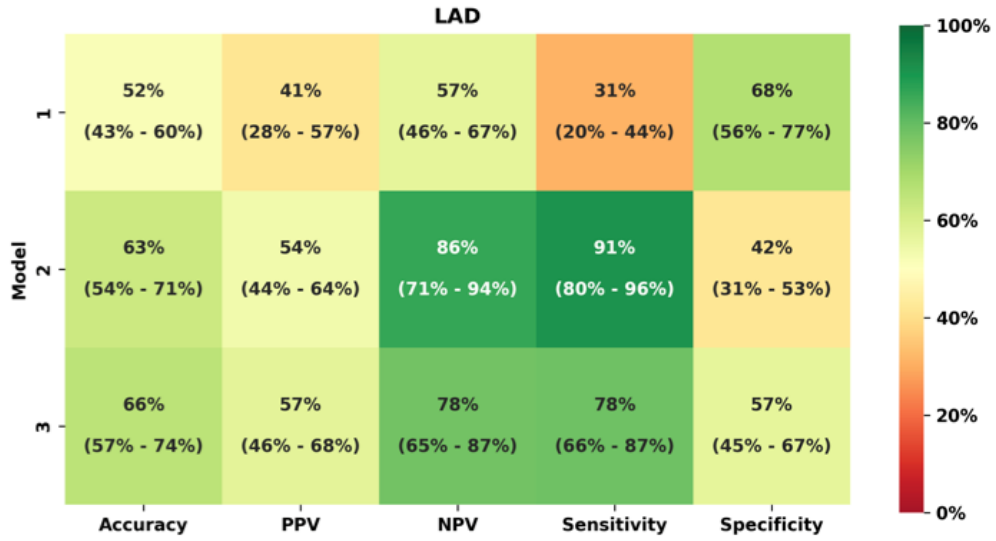


Figure 4. Performance metrics (with 95% CI) of each AI model regarding iFR binary classification prediction using the measured iFR classification as reference, for iFR measurements in the LAD. Accuracy, PPV, NPV, sensitivity, and specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values. *AI = artificial intelligence; iFR = instantaneous free-wave ratio; LAD = left anterior descending artery; NPV = negative predictive value; PPV = positive predictive value.*

Right coronary artery (RCA) lesions. The iFR classification of model 1 differed significantly from the measured iFR (*P* = .005), whereas for models 2 and 3 there were no significant differences (*P* = .282 and .357, respectively). All models classified most lesions as negative (albeit in smaller proportion to the actual measurements distribution), especially model 1. Agreement was highest for negative lesions in all models. Details are presented in **Table 5**.

TABLE 5. IFR CLASSIFICATION OF LESIONS AS MEASURED VS AS PER EACH MODEL PREDICTION FOR RIGHT CORONARY ARTERY (RCA) CASES

Classification (RCA)		Model 1		Model 2		Model 3		Total
		≤0.89	>0.89	≤0.89	>0.89	≤0.89	>0.89	
Measured iFR	≤0.89	3	2	1	4	2	3	5
	>0.89	9	62	29	42	14	57	71
Total		12	64	30	46	16	60	76

iFR = instantaneous free-wave ratio.

Models 1 and 3 displayed the highest accuracy, especially the former (86%), which was in contrast with model 2. The NPV was always very high, with a maximum of 97% for model 1. On the opposite spectrum of performance, the PPV was very low for all models. Sensitivity was high for both models 1 and 3. Lastly, specificity was low for models 2 and 3, and modest for model 1. Details are presented in **Figure 5**.

Appendix

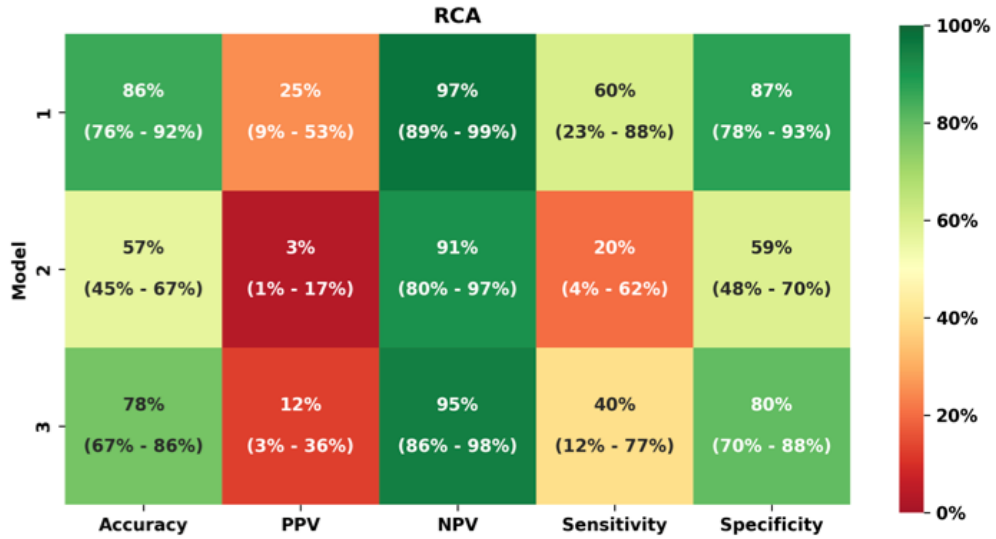


Figure 5. Performance metrics (with 95% CI) of each AI model regarding iFR binary classification prediction using the measured iFR classification as reference, for iFR measurements in the RCA. Accuracy, PPV, NPV, sensitivity, and specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values. *AI = artificial intelligence; iFR = instantaneous free-wave ratio; LAD = left anterior descending artery; NPV = negative predictive value; PPV = positive predictive value; RCA = right coronary artery.*

Circumflex artery (Cx) lesions. The neasured iFR classification differed significantly from the iFR classification of model 1 (P = .039), but not models 2 and 3 (P < .673 and .333, respectively). All models classified most lesions as negative, but always in quite shorter proportion to the actual measurements' distribution. Agreement was highest for negative lesions in all models. Details are presented in **Table 6**.

TABLE 6. IFR CLASSIFICATION OF LESIONS AS MEASURED VS AS PER EACH MODEL PREDICTION FOR CIRCUMFLEX ARTERY (Cx) CASES

Classification (Cx)		Model 1		Model 2		Model 3		Total
		≤0.89	>0.89	≤0.89	>0.89	≤0.89	>0.89	
Measured iFR	≤0.89	4	1	2	3	3	2	5
	>0.89	13	27	20	20	15	25	40
Total		17	28	22	23	18	27	45

iFR = instantaneous free-wave ratio.

Model 1 scored higher in all metrics, with an accuracy of 69% for a very high NPV (96%) and high specificity (80%). Sensitivity was 68% and the PPV was low (24%). Model 3 followed with similar, albeit inferior performance. Model 2 only scored high in NPV (87%). Details are presented in **Figure 6**.

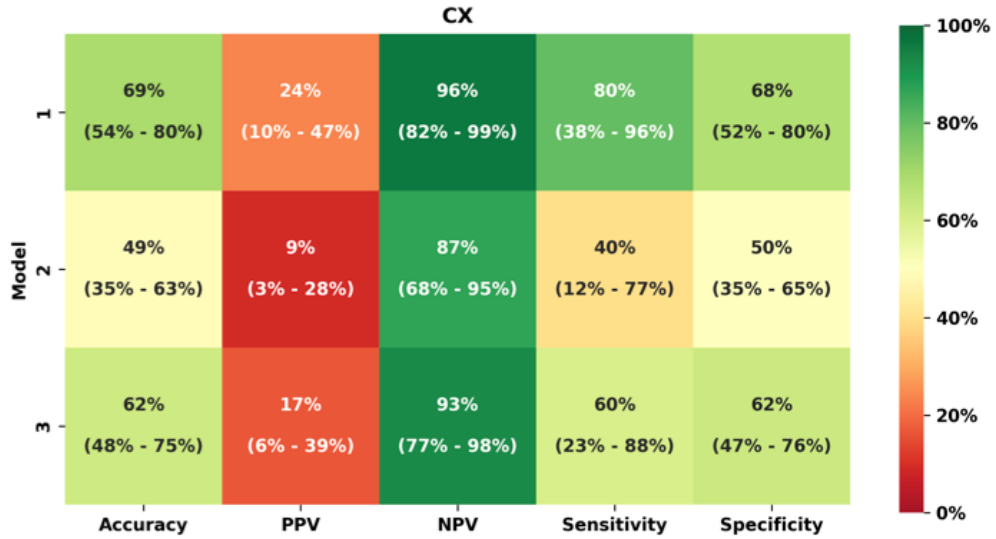


Figure 6. Performance metrics (with 95% CI) of each AI model regarding iFR binary classification prediction using the measured iFR classification as reference, for iFR measurements in the Cx. Accuracy, PPV, NPV, sensitivity, and specificity are displayed. Background in green for values close to 100%, red for values approaching 0% and mixed (yellow) for mid-range values. *AI = artificial intelligence; Cx = circumflex artery; iFR = instantaneous free-wave ratio; LAD = left anterior descending artery; NPV = negative predictive value; PPV = positive predictive value; RCA = right coronary artery.*

Discussion

Main findings: a proof of concept. In this study, we were able to develop AI models capable of binary iFR lesion classification. The accuracy of all models was modest, close to 70% for model 3. For both the LAD and Cx, the best that any model could achieve was also just below 70%. However, for the RCA, an accuracy of 78% and 86% was achieved for models 3 and 1, respectively, which, in the latter case, is quite high.

The models displayed very different reliability in correctly classifying lesions as either positive (ie, iFR ≤ 0.89) or negative (ie, iFR > 0.89). Indeed, the PPV was low or very low for all models, likely because much fewer positive cases were available for training. Only for the LAD, where there was a larger number of positive lesions, were the models able to achieve modest performance, surpassing the 50% PPV mark. The NPV, however, was generally high or very high, ranging from 77% to 89% overall. For both the RCA and the Cx, at least 1 model neared 100%.

At first glance, one might interpret these findings to be a result of the predominance of negative iFR cases, overwhelmingly so for the RCA and the Cx. Therefore, the models could simply be producing a result based on the statistically higher likelihood of a negative result, rather than extrapolating from CAG images.

We believe several factors suggests otherwise. First, the models did not necessarily reproduce the distribution of positive/negative lesions of the true measurements, neither globally nor regarding specific vessels. For example, model 3, which achieved the highest accuracy overall, only classified lesions as negative in 56% of cases (vs 74% for real measurements). The difference was even greater for the Cx, where the models' negatively classified lesions ranged from 51% to 62% (vs a real result of 89%).

Furthermore, the models' NPV was also greater than the proportion of negative cases, reaching 86% for the LAD (vs 54% real negative cases) and very close to 100% for both the RCA and Cx (where the real proportion of negative cases was close to 90%). Thus, the number of iFR-negative cases provided enough training data for the models to correctly learn to classify a lesion as negative with high reliability.

Despite this, a significant number of true negative lesions were not identified, as evidenced by modest specificity overall. Models 1 and 3 were exceptions in the case of RCA lesions, where higher performances were obtained.

No single model proved to be ideal. Model 1, followed by model 3, seemed to be the best option for the RCA and the Cx, whereas both models 2 and 3 were a better fit for the LAD. This suggests that fine-tuning a model for the target vessel may improve results.

Appendix

Considering all of the above, we believe our models offer proof of concept that the derivation of coronary physiology data by deep learning AI models based on X-ray angiography alone is feasible.

Potential practical clinical implications. When faced with either a positive or negative result by any such software, the operator's main question is often whether the result is likely correct (ie, how high is the PPV and/or NPV). Arguably, in the context of invasive coronary physiology, the NPV is of particular importance: a negative result enables the operator to conclude the procedure without engaging in further invasive maneuvers (ie, deploying a guide catheter, wiring the target vessel, administering further drugs), as opposed to a positive result. The fact that the majority of measurements in invasive physiology are negative further strengthens this point.[9–11,15,17,23,31] For example, in this case series, 185 (74%) cases had an iFR greater than 0.89. Since the models classified lesions as negative in 42% to 72% of cases and the NPV approaches 89%, in the best-case scenario, arguably around one- to two-thirds of all measurements could have been avoided. For the RCA and Cx, where the NPVs neared 100% and the proportion of cases classified as negative was even higher, the impact would have been even more significant.

Thus, while the models are by no means reliable enough for current clinical deployment, this somewhat simplified analysis illustrates the potential practical implications of this technology.

Analyzing lesions on coronary angiography using AI: other studies. Two studies of FFR estimation from CAG using primarily AI methods were published. However, to the best of our knowledge, ours is the first published study to report fully automatic derivation of iFR from CAG using such methods.

Roguin et al[23] conducted a pilot feasibility study in a single-center population consisting of 31 patients with predominantly LAD lesions (80%). They reported an accuracy of 90%, an NPV of 87%, a PPV of 94%, sensitivity of 88%, and specificity of 93% when conducting a binary analysis of FFR ≤ 0.80, similar to our approach. Their single model is able to derive an estimated FFR value, with an area under the curve of 0.91 and an r correlation coefficient of 0.71 ($P < .001$). The study does not report the exact AI methods. However, the binary approach to classify lesions, fully based on AI, using routine angiography projections (rather than predefined angulations) with a fully automatic method, is conceptually very similar to ours. While these results are impressive, with a seemingly superior performance compared to ours, the small sample size and single center nature of the study are limiting factors.

Cho et al[24] used a different approach. With a very large sample of 1501 lesions from a single center (predominantly LAD [67%]), the authors plotted the target vessel diameters together with clinical characteristics (age, sex, body surface area, and target segment) to binarily classify FFR measurements with a threshold of less than or equal to 0.80. An overall accuracy of 82%, an NPV of 84%, a PPV of 81%, sensitivity of 84%, and specificity of 80% were reported in the test set, similar to the external validation dataset of 79 patients. The main limitation of this study, from an AI research perspective, is that it does not provide a fully automated AI physiology estimation, as it requires manual segmentation and diameter calculations of the vessel with an external software, thereby rendering the process semi-automatic and somewhat time-consuming.

Another group is currently launching an initiative for fully automated AI-based PCI guidance interpretation, including FFR[32], but the project is underway and results are not yet available.

Lastly, one of the largest groups in iFR research has applied AI for interpreting iFR pullback curves and found non-inferiority to human performance.[33] While this is a very different task from the one we explored, it highlights how AI is also being applied to this important index which is so commonly used in clinical practice.

There has also been some exploration of the estimation of Coronary Flow Reserve (CFR) from CAG using AI methods, mainly to perform microcirculatory studies.[34] Other authors have tested the application of deep learning methods to classifying stenosis using Quantitative Coronary Analysis (QCA)[35] or automatically detecting significant coronary stenosis using bounding boxes,[36,37] which could be of use given the operators' heterogeneity and tendency for lesion overestimation when interpreting CAG, as we have recently demonstrated ourselves.[28]

Physiology derivation from CAG images without a primary AI approach: other studies. The estimation of physiology from CAG images has been explored in recent years, with commercial software made recently available.

Most studies focused on FFR, using a threshold of less than or equal to 0.80. The FAST-FFR pivotal trial was a multicenter international study of 301 subjects with a predominance of LAD lesions (54.2%). The correlation between estimated (FFRangio) and measured FFR was r = 0.80 ($P < .001$), with an accuracy of 92.2%, an NPV of 94.8%, a PPV of 89%, sensitivity of 93.5%, and specificity of 91.2%.[17] A pooled analysis of 5 cohort studies yielded similar results.[38]

The FAST I[31] and II[16] studies tested a similar approach. In the larger FAST II study[16], 334 patients from 6 centers were enrolled (with 66% LAD lesions). The correlation with invasive FFR was r = 0.74 (*P* < .001). Two multicenter trials are ongoing to test the clinical outcomes of virtual (vFFR) vs invasive FFR approach.[39,40]

The Quantitative Flow Ratio (QFR) is perhaps the most extensively "virtual" FFR index studied to date. After the encouraging results of the multicenter FAVOR pilot study (73 patients, 84 vessels, mostly LAD [54.8%]),[41] 2 further studies enrolled over 600 patients from Europe and Asia.[15,42] In the one with a large European cohort, values of 86.8%, 93%, 76.3%, 86.5%, and 86.9% for accuracy, NPV, PPV, sensitivity, and specificity were obtained, respectively, with a correlation of r = 0.83 (*P* < .001).[15] A Chinese multicentric trial comparing a PCI QFR-guided strategy vs an invasive FFR-guided strategy yielded better outcomes for the QFR group.[43]

The derivation of iFR from CAG has also been explored on the REVEAL iFR trial.[44] Published results are expected soon.

All of the above studies employ primarily non-AI methods to derive FFR from CAG, using a combination of 3D-image reconstruction and computational fluid dynamics. They demonstrate that physiology can successfully be derived from CAG images alone and have a meaningful impact on clinical outcomes. Further ongoing research is likely to strengthen this approach. While all of the above post-pilot studies yielded better performance results than our model, one thing seems ubiquitous: the NPV is always high, which is of particular significance.

However, these approaches are not without disadvantages. A reasonable amount of manual input is necessary – marking the target vessel, proximal and distal points, defining "healthy" regions, or correcting for imperfections in manual segmentation – as we have experienced ourselves when testing these platforms at our own catheterization laboratory. They are therefore semi-automatic, somewhat time consuming, and may potentially not be as reliable in less experienced hands. The FAST II trial clearly illustrated this, as the performance at specific sites was lower than at the core lab; the reported overall accuracy, NPV, PPV, specificity, and sensitivity were 83% vs 90%, 85% vs 90%, 79% vs 90%, 71% vs 81%, and 89% vs 95%, respectively.[16]

Lastly, except for one,[44] the above-mentioned methods require more than 1 projection (sometimes prespecified), which is probably the result of multiple factors. Indeed, the 3D nature of the coronary anatomy and the existence of energy losses in very distal segments (with resulting lower pressures), along with the limitations of CAG resolution and motion artifacts, all render a 3D approach more reliable, and are likely playing a role in hampering our models' accuracy, since it was based on a single 2-dimensional (2D) frame. Thus, while our simplified approach may initially be perceived as advantageous, it was, in all likelihood, a limitation.

Limitations and future directions. As AI training is highly volume dependent, the dataset size was the most important limitation. This was especially relevant for cases with iFR less than or equal to 0.89 and cases pertaining to the Cx. However, our relative distribution regarding target vessel and positive/negative cases is in agreement with previously published datasets.[10,11,15,17,23,31] As a result, obtaining a dataset with enough iFR-positive cases for successful training, especially concerning the RCA and Cx, will require a much larger dataset. The dataset size was also limiting with regards to the method employed for model testing. Using a classical train/test split of 80% / 20% would have resulted in a small testing dataset and reduced our ability to test the models' performance, particularly for subanalyzing performance per target vessel, given the naturally unbalanced characteristics of the dataset. We therefore preferred a cross-validation split in 10 subsets, as mentioned in the Methods section. While this is a common approach in the field of machine learning, it may be regarded as a limitation as well.

Some may view the extrapolation of iFR rather than FFR as a limitation, because FFR was directly compared to angiography in clinical outcomes trials,[3-5] whereas iFR was only directly compared to FFR itself[10-12] However, iFR has repeatedly been shown to be non-inferior to FFR and today is (together with other resting indexes) the default tool of epicardial physiology assessment in many labs due to its simplicity. As a result, the current amount of iFR measurements far outpaces those of FFR in our lab, and thus iFR provides a much larger base for future training, improvement, and validation.

Another limitation is the fact that the model provides a binary classification, but not yet the iFR value itself. During preliminary testing, it was clear that determining the exact iFR value would require a much larger dataset, which was beyond the scope of a pivotal study.

The use of a single image (human-optimized and labelled end-diastolic 2D frame) for the models' training, instead of a 3D reconstruction based on multiple 2D projections, was also a limitation, given the truly 3D nature of coronary anatomy and lesions. The fact that the above-mentioned non-AI approaches obtained a superior performance with 3D reconstruction, rather than a 2D approach like ours, further supports this consideration.

Appendix

The single-center, retrospective dataset is another limitation, as external validation will be required in the future. Notwithstanding, physiology results have been shown to be quite reproducible and thus the impact of this particular limitation is not very likely to be of significance.

The concept of this study is thus exploratory and aimed at proof of concept. We aim to greatly enlarge the training dataset through multi-institutional collaboration (as we have done for our segmentation models[27,28]), which will be essential for performance improvements and external validation. We also aim to further improve our models using multiple projections and 3D reconstruction, which may enhance performance. The ultimate aim is deployment in clinical practice; this may occur either as a standalone solution, or by improving existing software.

Conclusions

We developed deep learning AI models capable of binary lesion classification using an iFR threshold of 0.89. While the overall accuracy of the models is not yet high enough for clinical deployment, the high negative predictive capacity of these models is of clinical significance and potential clinical application. This pivotal study therefore offers proof of concept for further development, with larger and multicentric training and validation datasets poised for the future. This approach has the potential to evolve into a standalone software aid in the catheterization laboratory, or for further integration into existing, non-AI based software solutions by streamlining workflows and/or improving their performance. This could prove to be of great value for patient management and to improve catheterization laboratory flows.

Affiliations and Disclosures

From the [1]Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon (CCUL@RISE), Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal; [2]Serviço de Cardiologia, Departamento de Coração e Vasos, CHULN Hospital de Santa Maria, Lisbon, Portugal; [3]INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal; [4]Neuralshift Inc., Lisboa, Portugal; [5]Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal.

Data Availability: Detailed full-scale study data cannot currently be made publicly available due to limitations imposed by national data protection regulations, as this is a retrospective study and no informed consent was obtainable regarding this particular analysis. Both our research team and others in the national scientific community are working to develop a framework where such would be possible. However, independent replication of our segmentation models is possible, given that the detailed description of our experimentations and relevant code is publicly available.[25-28]

Disclosures: The authors report no financial relationships or conflicts of interest regarding the content herein.

Funding: Cardiovascular Center of the University of Lisbon, INESC-ID / Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal.

Address for correspondence: Miguel Nobre Menezes, MD, MSc, Serviço de Cardiologia, Avenida Professor Egas Moniz, Lisbon, 1649-028, Portugal. Email: mnmenezes.gm@gmail.com; X: @nobremenezes

References

1. Neumann FJ, Sousa-Uva M, Ahlsson A, et al; ESC Scientific Document Group. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J*. 2019;40(2):87–165. doi: 10.1093/eurheartj/ehy394

2. Lawton JS, Tamis-Holland JE, Bangalore S, et al. 2021 ACC/AHA/SCAI Guideline for coronary artery revascularization: A report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2022;145(3): e4-e17. doi: 10.1161/CIR.0000000000001039

3. De Bruyne B, Pijls NHJ, Kalesan B, et al; FAME 2 Trial Investigators. Fractional flow reserve-guided PCI versus medical therapy in stable coronary disease. *N Engl J Med*. 2012;367(11):991-1001. doi: 10.1056/NEJMoa1205361

4. Tonino PAL, De Bruyne B, Pijls NHJ, et al; FAME Study Investigators. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N Engl J Med*. 2009;360(3):213-224. doi: 10.1056/NEJMoa0807611

5. Bech GJW, De Bruyne B, Pijls NHJ, et al. Fractional flow reserve to determine the appropriateness of angioplasty in moderate coronary stenosis: a randomized trial. *Circulation*. 2001;103(24):2928-2934. doi: 10.1161/01.cir.103.24.2928.

6. Baptista SB, Raposo L, Santos L, et al; PRIME-FFR Study Group. Impact of routine fractional flow reserve evaluation during coronary angiography on management strategy and clinical outcome. *Circ Cardiovasc Interv*. 2016;9(7):e003288. doi: 10.1161/CIRCINTERVENTIONS.115.003288

7. van de Hoef TP, Meuwissen M, Escaned J, et al. Head-to-head comparison of basal stenosis resistance index, instantaneous wave-free ratio, and fractional flow reserve: diagnostic accuracy for stenosis-specific myocardial ischaemia. *EuroIntervention*. 2015;11(8):914-925. doi: 10.4244/EIJY14M08_17

8. Escaned J, Echavarría-Pinto M, Garcia-Garcia HM, et al; ADVISE II Study Group. Prospective assessment of the diagnostic accuracy of instantaneous wave-free ratio to assess coronary stenosis relevance: Results of ADVISE II international, multicenter study (ADenosine vasodilator independent stenosis evaluation II). *JACC Cardiovasc Interv*. 2015;8(6):824-833. doi: 10.1016/j.jcin.2015.01.029

9. Nobre Menezes M, Francisco ARG, Carrilho Ferreira P, et al. Comparative analysis of fractional flow reserve and instantaneous wave-free ratio: Results of a five-year registry. *Rev Port Cardiol (Engl Ed)*. 2018;37(6):511-520. doi: 10.1016/j.repc.2017.11.011

10. Davies JE, Sen S, Dehbi H-M, et al. Use of the instantaneous wave-free ratio or fractional flow reserve in PCI. *N Engl J Med*. 2017;376(19):1824–1834. doi: 10.1056/NEJMoa1700445

11. Götberg M, Christiansen EH, Gudmundsdottir IJ, et al. Instantaneous wave-free ratio versus fractional flow reserve to guide PCI. *N Engl J Med*. 2017;376(19):1813–1823. doi: 10.1056/NEJMoa1616540

12. Götberg M, Berntorp K, Rylance R, et al. 5-year outcomes of PCI guided by measurement of instantaneous wave-free ratio versus fractional flow reserve. *J Am Coll Cardiol*. 2022;79(10):965-974. doi: 10.1016/j.jacc.2021.12.030

13. Petraco R, Park JJ, Sen S, et al. Hybrid iFR-FFR decision-making strategy: Implications for enhancing universal adoption of physiology-guided coronary revascularisation. *EuroIntervention*. 2013;8(10):1157–1165. doi: 10.4244/EIJV8I10A179

14. Tebaldi M, Biscaglia S, Fineschi M, et al. Evolving routine standards in invasive hemodynamic assessment of coronary stenosis: The nationwide Italian SICI-GISE cross-sectional ERIS study. *JACC Cardiovasc Interv.* 2018;11(15):1482–1491. doi: 10.1016/j.jcin.2018.04.037

15. Westra J, Andersen BK, Campo G, et al. Diagnostic performance of in-procedure angiography-derived quantitative flow reserve compared to pressure-derived fractional flow reserve: The FAVOR II Europe-Japan study. *J Am Heart Assoc*. 2018;7(14):e009603. doi: 10.1161/JAHA.118.009603

16. Masdjedi K, Tanaka N, Van Belle E, et al. Vessel fractional flow reserve (vFFR) for the assessment of stenosis severity: The FAST II study. *EuroIntervention*. 2022;17(18):1498-1505. doi: 10.4244/EIJ-D-21-00471

17. Fearon WF, Achenbach S, Engstrom T, et al. Accuracy of fractional flow reserve derived from coronary angiography. *Circulation*. 2019;139(4):477-484. doi: 10.1161/CIRCULATIONAHA.118.037350

18. Valente Silva B, Marques J, Nobre Menezes M, Oliveira AL, Pinto FJ. Artificial intelligence-based diagnosis of acute pulmonary embolism: Development of a machine learning model using 12-lead electrocardiogram. *Rev Port Cardiol*. 2023;42(7):643-651. doi: 10.1016/j.repc.2023.03.016

19. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. *J Am Coll Cardiol.* 2016;68(21):2287-2295. doi: 10.1016/j.jacc.2016.08.062

20. Asch FM, Poilvert N, Abraham T, et al. Automated echocardiographic quantification of left ventricular ejection fraction without volume measurements using a machine learning algorithm mimicking a human expert. *Circ Cardiovasc Imaging*. 2019;12(9):e009303. doi: 10.1161/CIRCIMAGING.119.009303

21. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal*. 2017;35:159-171. doi: 10.1016/j.media.2016.05.009

22. Bai W, Sinclair M, Tarroni G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J Cardiovasc Magn Reson*. 2018;20(1):65. doi: 10.1186/s12968-018-0471-x

Appendix

23. Roguin A, Abu Dogosh A, Feld Y, Konigstein M, Lerman A, Koifman E. Early feasibility of automated artificial intelligence angiography based fractional flow reserve estimation. *Am J Cardiol*. 2021;139:8-14. doi: 10.1016/j.amjcard.2020.10.022

24. Cho H, Lee JG, Kang SJ, et al. Angiography-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. *J Am Heart Assoc*. 2019;8(4):e011685. doi: 10.1161/JAHA.118.011685

25. Nobre Menezes M, Lourenço-Silva J, Silva B, et al. Development of deep learning segmentation models for coronary X-ray angiography: Quality assessment by a new global segmentation score and comparison with human performance. *Rev Port Cardiol*. 2022;41(12):1011-1021. doi: 10.1016/j.repc.2022.04.001

26. Silva JL, Menezes MN, Rodrigues T, Silva B, Pinto FJ, Oliveira AL. Encoder-decoder architectures for clinically relevant coronary artery segmentation. *arXiv*. Preprint posted online June 21, 2021. doi: 10.48550/arXiv.2106.11447

27. Nobre Menezes M, Silva JL, Silva B, et al. Coronary C-ray angiography segmentation using artificial intelligence: A multicentric validation study of a deep learning model. *Int J Cardiovasc Imaging*. 2023;39(7):1385-1396. doi: 10.1007/s10554-023-02839-5

28. Menezes MN, Silva B, Silva JL, et al. Segmentation of x-ray coronary angiography with an artificial intelligence deep learning model: Impact in operator visual assessment of coronary stenosis severity. *Catheter Cardiovasc Interv*. 2023 102(4):631-640. doi: 10.1002/ccd.30805

29. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *arXiv*. Preprint posted online June 12, 2017. doi: 10.48550/arXiv.1706.03762

30. Tan M, Le Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 36th Int Conf Mach Learn ICML 2019. *arXiv*. Preprint posted online May 28, 2019. doi: 10.48550/arXiv.1905.11946

31. Masdjedi K, van Zandvoort LJC, Balbi MM, et al. Validation of a three-dimensional quantitative coronary angiography-based software to calculate fractional flow reserve: the FAST study. *EuroIntervention*. 2020;16(7):591–599. doi: 10.4244/EIJ-D-19-00466

32. Ploscaru V, Popa-Fotea NM, Calmac L, et al. Artificial intelligence and cloud based platform for fully automated PCI guidance from coronary angiography-study protocol. *PLoS One*. 2022;17(9):e0274296. doi: 10.1371/journal.pone.0274296

33. Cook CM, Warisawa T, Howard JP, et al. Algorithmic versus expert human interpretation of instantaneous wave-free ratio coronary pressure-wire pull back data. *JACC Cardiovasc Interv*. 2019;12(14):1315-1324. doi: 10.1016/j.jcin.2019.05.025

34. Zhao Q, Li C, Chu M, Gutiérrez-Chico JL, Tu S. Angiography-based coronary flow reserve: The feasibility of automatic computation by artificial intelligence. *Cardiol J*. 2023;30(3):369-378. doi: 10.5603/CJ.a2021.0087

35. Liu X, Wang X, Chen D, Zhang H. Automatic quantitative coronary analysis based on deep learning. *Appl Sci*. 2023;13(5):2975. doi: 10.3390/app13052975

36. Avram R, Olgin JE, Ahmed Z, et al. CathAI: fully automated coronary angiography interpretation and stenosis estimation. *NPJ Digit Med*. 2023;6(1):142. doi: 10.1038/s41746-023-00880-1

37. Freitas SA, Zeiser FA, Da Costa CA, De O, Ramos G. DeepCADD: A Deep Learning Architecture for Automatic Detection of Coronary Artery Disease. In: *2022 International Joint Conference on Neural Networks (IJCNN)*. 2022:1-8. doi: 10.1109/IJCNN55064.2022.9892501

38. Witberg G, De Bruyne B, Fearon WF, et al. Diagnostic performance of angiogram-derived fractional flow reserve: A pooled analysis of 5 prospective cohort studies. *JACC Cardiovasc Interv*. 2020;13(4):488-497. doi: 10.1016/j.jcin.2019.10.045

39. European Cardiovascular Research Institute. FAST III. Accessed August 19, 2023. https://www.ecri-trials.com/studies/fast-iii/

40. US National Library of Medicine. Comparison of Vessel-FFR Versus FFR in Intermediate Coronary Stenoses (LIPSIASTRATEGY). ClinicalTrials.gov. Accessed August 19, 2023. https://classic.clinicaltrials.gov/ct2/show/NCT03497637

41. Tu S, Westra J, Yang J, et al; FAVOR Pilot Trial Study Group. Diagnostic accuracy of fast computational approaches to derive fractional flow reserve from diagnostic coronary angiography: The international multicenter FAVOR pilot study. *JACC Cardiovasc Interv*. 2016;9(19):2024-2035. doi: 10.1016/j.jcin.2016.07.013

42. Xu B, Tu S, Qiao S, et al. Diagnostic accuracy of angiography-based quantitative flow ratio measurements for online assessment of coronary stenosis. *J Am Coll Cardiol*. 2017;70(25):3077–3087. doi: 10.1016/j.jacc.2017.10.035

43. Xu B, Tu S, Song L, et al; FAVOR III China study group. Angiographic quantitative flow ratio-guided coronary intervention (FAVOR III China): A multicentre, randomised, sham-controlled trial. *Lancet*. 2021;398(10317):2149–2159. doi: 10.1016/S0140-6736(21)02248-0

44. Ono M, Serruys PW, Patel MR, et al. A prospective multicenter validation study for a novel angiography-derived physiological assessment software: Rationale and design of the radiographic imaging validation and evaluation for Angio-iFR (ReVEAL iFR) study. *Am Heart J*. 2021;239:19-26. doi: 10.1016/j.ahj.2021.05.004

Supplementary Data

Supplementary tables

Appendix

Supplementary table 1: overall performance by group and individual criteria (significance at p <0,05 for paired differences and p < 0,01 for multiple comparisons).
BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model.

Criteria	Nr / % of images where criteria was met	Group				p-value						
		BH	EH	BAI	EAI	Between all	BH vs EH	BAI vs EAI	BH vs BAI	EH vs EAI	BH vs EAI	EH vs BAI
Main Vessel Segmentation	N	414	416	408	414	**0,007**	0,157	0,056	0,056	0,157	1	**0,004**
	%	99,5%	100,0%	98,1%	99,5%							
Main Vessel Gaps	N	416	416	402	407	**< 0,001**	-	0,29	**< 0,001**	**0,003**	**0,003**	**<0,001**
	%	100,0%	100,0%	96,6%	97,8%							
Catheter to Vessel Transition	N	366	400	334	367	**< 0,001**	**<0,001**	**0,002**	**0,002**	**<0,001**	0,915	**<0,001**
	%	88,0%	96,2%	80,3%	88,2%							
Branch Segmentation	N	414	416	391	405	**< 0,001**	0,157	**0,017**	**< 0,001**	**0,001**	**0,012**	**<0,001**
	%	99,5%	100,0%	94,0%	97,4%							
Branch Gaps	N	406	415	116	140	**< 0,001**	**0,006**	0,071	**< 0,001**	**<0,001**	**<0,001**	**<0,001**
	%	97,6%	99,8%	27,9%	33,7%							
Coronary Artifacts	N	409	411	305	356	**< 0,001**	0,561	**<0,001**	**< 0,001**	**<0,001**	**<0,001**	**<0,001**
	%	98,3%	98,8%	73,3%	85,6%							
Catheter Segmentation	N	415	416	396	360	**< 0,001**	0,317	**< 0,001**	**< 0,001**	**<0,001**	**< 0,001**	**< 0,001**
	%	99,8%	100,0%	95,2%	86,5%							
Catheter Gaps	N	413	415	403	319	**< 0,001**	0,316	**< 0,001**	**0,012**	**<0,001**	**< 0,001**	**< 0,001**
	%	99,3%	99,8%	96,9%	76,7%							
Catheter Artifacts	N	407	415	245	225	**< 0,001**	**0,011**	0,162	**< 0,001**	**<0,001**	**<0,001**	**<0,001**
	%	97,8%	99,8%	58,9%	54,1%							
Catheter Location	N	412	416	393	383	**< 0,001**	**0,045**	0,166	**< 0,001**	**<0,001**	**< 0,001**	**< 0,001**
	%	99,0%	100,0%	94,5%	92,1%							
Catheter Thickness	N	358	390	265	401	**< 0,001**	**< 0,001**	**< 0,001**	**< 0,001**	0,078	**< 0,001**	**< 0,001**
	%	86,1%	93,8%	63,7%	96,4%							

Supplementary table 2: performance by criteria and model –RCA vs LCA. BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model; RCA: right coronary artery; LCA: left coronary artery.

		Left Coronary		Right Coronary		p-value
	Model	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	
Main Vessel Segmentation	BH	255	99,2	159	100,0	0,265
	EH	257	100,0	159	100,0	0,071
	BAI	251	97,7	157	98,7	0,437
	EAI	257	100,0	157	98,7	0,146
Main Vessel Gaps	BH	257	100,0	159	100,0	NA
	EH	257	100,0	159	100,0	NA
	BAI	243	94,6	159	100,0	**0,003**
	EAI	249	96,9	158	99,4	0,091
Catheter to Vessel Transition	BH	218	84,8	148	93,1	**0,012**
	EH	244	94,9	156	98,1	0,102
	BAI	195	75,9	139	87,4	**0,004**
	EAI	218	84,8	149	93,7	**0,006**
Branch Segmentation	BH	255	99,2	159	100,0	0,265
	EH	257	100,0	159	100,0	NA
	BAI	238	92,6	153	96,2	0,131
	EAI	248	96,5	157	98,7	0,166
Branch Gaps	BH	250	97,3	156	98,1	0,588
	EH	156	60,7	159	100,0	0,431
	BAI	64	24,9	52	32,7	0,085
	EAI	67	26,1	73	45,9	**<0,001**
Coronary Artifacts	BH	252	98,1	157	98,7	0,596
	EH	253	98,4	158	99,4	0,399
	BAI	184	71,6	121	76,1	0,313
	EAI	216	84,0	140	88,1	0,259
Catheter Segmentation	BH	256	99,6	159	100,0	0,431
	EH	257	100,0	159	100,0	NA
	BAI	238	92,6	158	99,4	**0,002**
	EAI	211	82,1	149	93,7	**0,001**
Catheter Gaps	BH	256	99,6	157	98,7	0,309
	EH	257	100,0	158	99,4	0,203
	BAI	249	96,9	154	96,9	0,986
	EAI	196	76,3	123	77,4	0,798
Catheter Artifacts	BH	251	97,7	156	98,1	0,76
	EH	256	99,6	159	100,0	**0,043**
	BAI	149	58,0	96	60,4	0,629
	EAI	149	58,0	76	47,8	0,409
Catheter Location	BH	256	99,6	156	98,1	0,128
	EH	257	100,0	159	100,0	0,329
	BAI	237	92,2	156	98,1	**0,011**
	EAI	234	91,1	149	93,7	0,061
Catheter Thickness	BH	218	84,8	140	88,1	0,356
	EH	241	93,8	149	93,7	0,979
	BAI	156	60,7	109	68,6	0,106
	EAI	245	95,3	156	98,1	0,139
	Total	257	100,0	159	100,0	416

Appendix

Supplementary table 3: GSS by model - RCA vs LCA. BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model; GSS: Global Segmentation Score; RCA: right coronary artery; LCA: left coronary artery.

		Left Coronary		Right Coronary		p-value
	Model	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	
GSS	BH	96,5 +/- 5,9	100 (9)	97,5 +/- 5,3	100 (0)	0,058
	EH	99,1 +/- 2,7	100 (0)	98,6 +/- 3,5	100 (0)	0,204
	BAI	85,1 +/- 10,8	87,5 (10,5)	87,5 +/- 8,8	87,5 (9)	**0,017**
	EAI	89,7 +/ 7,9	92 (10,3)	90,4 +/- 7,1	92 (9)	0,389

Supplementary table 4: performance by model, criteria across multiple angulation incidences, for the left coronary artery. A single case of PA incidence was excluded from the analysis. BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model.

Criteria	Model	LAO Cranial		LAO Caudal (*Spider*)		RAO Cranial		RAO Caudal		PA Cranial		PA Caudal		p value
		N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	
Main Vessel Gaps	EAI	22	84,6	44	93,6	80	100,0	60	100,0	10	100,0	32	97,0	**0,002**
Catheter to Vessel Transition	BAI	17	65,4	31	66,0	59	73,8	53	88,3	5	50,0	29	87,9	**0,008**
	EH	24	92,3	47	100,0	71	88,8	60	100,0	10	100,0	31	93,9	0,023
	EAI	25	96,2	37	78,7	62	77,5	56	93,3	6	60,0	31	93,9	**0,004**
Branch Gaps	EAI	5	19,2	11	23,4	19	23,8	12	20,0	4	40,0	16	48,5	0,038
Coronary Artifacts	BAI	21	80,8	28	59,6	51	63,8	50	83,3	6	60,0	27	81,8	0,021
Catheter Segmentation	EAI	13	50,0	39	83,0	64	80,0	57	95,0	6	60,0	31	93,9	**<0,001**
Catheter Gaps	EAI	20	76,9	30	63,8	56	70,0	54	90,0	6	60,0	30	90,9	**0,004**
Catheter Artifacts	BAI	15	57,7	26	55,3	36	45,0	43	71,7	6	60,0	22	66,7	0,046
Total	256	26	100,0	47	100,0	80	100,0	60	100,0	10	100,0	33	100,0	-

Supplementary table 5: performance by model, criteria across multiple angulation incidences, for the right coronary artery. A single case of Spider view was excluded from the analysis. BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model.

Criteria	Model	LAO Cranial		LAO		RAO		PA Cranial		p value
		N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	N meeting criteria	% meeting criteria	
Coronary Artifacts	BAI	35	71,4	51	87,9	30	68,2	4	57,1	0,048
Catheter Location	BAI	49	100,0	56	96,6	44	100,0	6	85,7	0,04
Total	158	49	100,0	58	100,0	44	100,0	7	100,0	-

Appendix

Supplementary table 6: GSS by model, criteria across multiple angulation incidences, for the left coronary artery. A single case of PA incidence was excluded from the analysis. BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model; GSS: Global Segmentation Score

	Model	LAO Cranial		LAO Caudal (*Spider*)		RAO Cranial		RAO Caudal		PA Cranial		PA Caudal		p value
		Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	
GSS	BH	99,2 +/ 2,9	100 (0)	99,1 +/- 2,7	100 (0)	96,4 +/- 5,2	100 (9)	97,4 +/- 4,1	100 (8,5)	92,9 +/- 6,3	91 (10,5)	90,1 +/- 9,7	91 (11,5)	**<0,001**
	EH	99,6 +/- 2,1	100 (0)	99,6 +/- 1,8	100 (0)	99,3 +/- 2,5	100 (0)	99,1 +/- 2,7	100 (0)	95,1 +/- 4,7	94 (9)	98,9 +/- 3	100 (0)	**< 0,001**
	BAI	88,1 +/- 6	87,5 (7)	88,2 +/- 7	87,5 (12)	85,3 +/- 9,3	87 (9,5)	85,6 +/- 7,9	86 (7,8)	81,5 +/- 13,6	87,5 (17,8)	78,4 +/- 19,5	84,5 (16)	0,116
	EAI	90,2 +/- 7,7	92 (11)	89,6 +/- 6,4	92 (10)	90,9 +/- 6,9	92 (6,5)	88,8 +/- 7,2	88,5 (10,8)	87,1 +/- 14,2	90 (14)	88,6 +/- 10,8	92 (11,5)	0,451

Supplementary table 7: GSS by model, criteria across multiple angulation incidences, for the right coronary artery. A single case of Spider view was excluded from the analysis. BH: baseline human model; EH: enhanced human model; BAI: baseline AI model; EAI; enhanced AI model; GSS: Global Segmentation Score

	Model	LAO Cranial		LAO		RAO		PA Cranial		p value
		Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	Mean +/- SD	Median (IQR)	
GSS	BH	98,2 +/- 4,7	100 (0)	98,8 +/- 3,6	100 (0)	95,9 +/- 6,7	100 (9)	92,3 +/- 8	91 (15)	**< 0,001**
	EH	98,6 +/- 3,8	100 (0)	99,5 +/- 2,2	100 (0)	97,9 +/- 4	100 (0)	96,6 +/- 6,1	100 (9)	0,059
	BAI	89,7 +/- 7,3	89,5 (11,5)	86,7 +/- 7,1	87,5 (9,4)	86,4 +/- 11,7	88,5 (12)	88,7 +/- 6	91 (9,5)	0,221
	EAI	91,7 +/- 4,8	92 (5)	90,2 +/- 7,3	92 (11,5)	89,6 +/- 8,7	92 (9,5)	89,8 +/- 6	92 (6)	0,818

Supplementary table 8: differences in diameter stenosis (%), lesion diameter, proximal obstruction border diameter and distal obstruction border diameter per stenosis severity. Values shown as median (IQ 25th – 75th). *Kruskal-Wallis test

Stenosis Severity - original QCA image as reference (N, %)	Diameter stenosis (%) difference	Diameter at lesion (mm) difference	Diameter at proximal obstruction border (mm) difference	Diameter at distal obstruction border (mm) difference
≥ 70% (22, 18%))	4,9 (2,2 – 8,3)	0,12 (0,06 – 0,18)	0,13 (0,08 – 0,27)	0,15 (0,06 – 0,27)
50-69%(58, 47%)	5,2 (0 – 7,7)	0,12 (0 - 0,18)	0,18 (0,10 – 0,31)	0,10 (0 – 0,20)
< 50% (43, 35%)	4,7 (0 – 6,2)	0,08 (0 – 0,15)	0,16 (0,10 – 0,27)	0,08 (0 – 0,15)
P-value*	0,224	0,274	0,651	0,094

Supplementary table 9: differences in diameter stenosis (%), lesion diameter, proximal obstruction border diameter and distal obstruction border diameter per target vessel. Values shown as median (IQ 25th – 75th). *Kruskal-Wallis test; LAD – Left Anterior Descending; RCA; Right Coronary Artery; CX: Circumflex Artery

Target Vessel	Diameter stenosis (%) difference	Diameter at lesion (mm) difference	Diameter at proximal obstruction border (mm) difference	Diameter at distal obstruction border (mm) difference
LAD	4,7 (0 – 7,8)	0,09 (0 – 0,17)	0,17 (0,09 – 0,32)	0,11 (0 – 0,19)
RCA	4,1 (0 – 7,7)	0,11 (0 - 0,20)	0,17 (0,08 – 0,28)	0,11 (0 – 0,20)
CX	4,2 (0 – 6,8)	0,09 (0 – 0,15)	0,15 (0,10 – 0,29)	0,08 (0 – 0,18)
P-value*	0,801	0,817	0,569	0,679

Appendix

Supplementary table 10: detailed measurements between the original and the segmented images per center. Values shown as mean ± standard deviation. AI – artificial intelligence. *Paired samples T-test.

Center (N/%)	Parameter	Original Image	AI Generated Segmented Image	P-value*
A (39 / 32%)	Diameter Stenosis (%)	60 ± 11	59 ± 11	0,203
	Diameter at lesion (mm)	0,88 ± 0,33	0,91 ± 0,33	0,261
	Diameter at proximal obstruction border (mm)	2,14 ± 0,49	1,99 ± 0,45	**< 0,01**
	Diameter at distal obstruction border (mm)	2,04 ± 0,46	2,04 ± 0,53	0,930
B (36 / 29%)	Diameter Stenosis (%)	54 ± 14	53 ± 12	0,223
	Diameter at lesion (mm)	1,13 ± 0,39	1,16 ± 0,34	0,277
	Diameter at proximal obstruction border (mm)	2,38 ± 0,58	2,17 ± 0,61	**< 0,01**
	Diameter at distal obstruction border (mm)	2,27 ± 0,61	2,22 ± 0,61	0,067
C (26 / 21%)	Diameter Stenosis (%)	54 ± 14	54 ± 15	0,818
	Diameter at lesion (mm)	1,22 ± 0,40	1,24 ± 0,42	0,778
	Diameter at proximal obstruction border (mm)	2,49 ± 0,49	2,28 ± 0,48	**< 0,01**
	Diameter at distal obstruction border (mm)	2,47 ± 0,52	2,39 ± 0,60	0,170
D (22 / 18%)	Diameter Stenosis (%)	53 ± 12	52 ± 13	0,504
	Diameter at lesion (mm)	1,06 ± 0,33	1,07 ± 0,31	0,842
	Diameter at proximal obstruction border (mm)	2,07 ± 0,52	1,88 ± 0,53	**< 0,01**
	Diameter at distal obstruction border (mm)	1,99 ± 0,53	1,97 ± 0,55	0,648

Supplementary table 11: differences in diameter stenosis (%), lesion diameter, proximal obstruction border diameter and distal obstruction border diameter per center. Values shown as median (IQ 25th – 75th). *Kruskal-Wallis test.

Center (N / %)	Diameter stenosis (%) difference	Diameter at lesion (mm) difference	Diameter at proximal obstruction border (mm) difference	Diameter at distal obstruction border (mm) difference
A (39 / 32%)	5,0 (0 – 6,9)	0,11 (0 - 0,17)	0,11 (0,08 – 0,27)	0,09 (0 – 0,17)
B (36 / 29%)	3,4 (0 – 7,8)	0,09 (0 – 0,20)	0,19 (0,09 – 0,32)	0,09 (0 – 0,17)
C (26 / 21%)	4,0 (0 – 7,7)	0,09 (0 – 0,19)	0,24 (0,10 – 0,28)	0,10 (0 – 0,21)
D (22 / 18%)	4,7 (0 – 7,7)	0,11 (0 – 0,16)	0,16 (0,10 – 0,28)	0,13 (0 – 0,23)
P-value*	0,635	0,952	0,354	0,527

Supplementary table 12: differences in catheter dimensions. *Mean ± Standard Deviation ** median (IQ 25th – 75th) *** Paired samples T-test AI – artificial intelligence.

Size	Parameter	Original Image	AI Generated Segmented Image	P-value***
5 Fr (78 cases)	Diameter* (mm)	1,79 ± 0,08	1,77 ± 0,14	0,091
	Difference** (mm)	0,04 (0 – 0,13)		
6 Fr (13 cases)	Diameter*	2,05 ± 0,08	2,1 ± 0,3	0,270
	Difference**	0,16 (0 – 0,76)		

Supplementary table 13: differences in overlap metrics per target vessel. Values shown as median (IQ 25th – 75th). *Kruskal-Wallis test ; LAD – Left Anterior Descending; RCA; Right Coronary Artery; CX: Circumflex Artery

Target Vessel	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Intersection over Union (%)	Dice Score (%)
LAD	99,9 (99,9 – 99,9)	94,2 (91,8 – 96,2)	99,9 (99,9 – 99,9)	94,6 (91,9 – 95,9(	99,9 (99,9 – 99,9)	89,1 (85,5 – 90,1)	94,2 (92,2 – 95,0)
RCA	99,9 (99,9 – 99,9)	95,1 (92,9 – 96,6)	99,9 (99,9 – 99,9)	95,1 (93,6 – 96,6)	99,9 (99,9 – 99,9)	90,8 (88,1 – 92,0)	95,2 (93,7 – 95,9)
CX	99,9 (99,9 – 99,9)	95,5 (94,3 – 96,4)	99,9 (99,9 – 99,9)	95,2 (93,4 – 96,6)	99,9 (99,9 – 99,9)	91,9 (90,1 – 93,5)	96,8 (96,3 – 96,9)
P-value*	0,176	0,386	0,140	0,119	0,412	**0,01**	**0,01**

Supplementary table 14: differences in overlap metrics per stenosis severity. Values shown as median (IQ 25th – 75th). *Kruskal-Wallis test.

Stenosis Severity (original image as reference)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Intersection over Union (%)	Dice Score (%)
≥ 70%	99,9 (99,9 – 99,9)	95,2 (91,8 – 96,6)	99,9 (99,9 – 99,9)	95,2 (91,8 – 97,5)	99,9 (99,9 – 99,9)	89,3 (84,4 – 91,6)	94,3 (91,6 – 95,6)
50-69%	99,9 (99,9 – 99,9)	95,3 (92,3 – 96,5)	99,9 (99,9 – 99,9)	94,6 (91,8 – 96,0)	99,9 (99,9 – 99,9)	89,3 (85,5 – 91,4)	94,3 (92,2 – 95,5)
< 50%	99,9 (99,9 – 99,9)	95,0 (94,2 – 96,2)	99,9 (99,9 – 99,9)	95,6 (94,6 – 96,6)	99,9 (99,9 – 99,9)	90,1 (89,2 – 91,9)	95,0 (94,3 – 95,8)
P-value*	0,482	0,969	0,376	**0,024**	0,581	**0,029**	**0,029**

Appendix

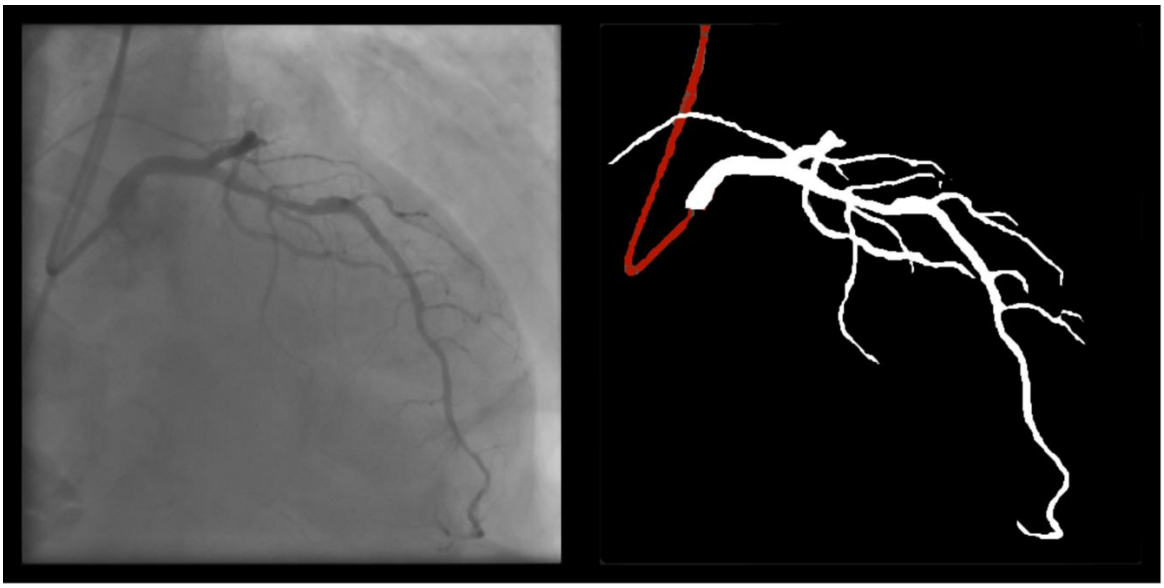
Supplementary table 15: overlap metrics per center. Values shown as median (IQ 25th – 75th).

Center (N / %)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Intersection over Union (%)	Dice Score (%)
A (39 / 32%)	99,9 (99,9 – 99,9)	95,4 (93,3 – 96,6)	99,9 (99,9 – 99,9)	94,7 (92,3 – 96,8)	99,9 (99,9 – 99,9)	89,3 (87,8 – 92,1)	94,4 (93,5 – 95,9)
B (36 / 29%)	99,9 (99,9 – 99,9)	94,9 (93,2 – 96,4)	99,9 (99,9 – 99,9)	94,6 (92,9 – 96,1)	99,9 (99,9 – 99,9)	90,4 (86,7 – 91,9)	94,9 (92,9 – 95,8)
C (26 / 21%)	99,9 (99,9 – 99,9)	95,0 (92,2 – 96,1)	99,9 (99,9 – 99,9)	95,6 (94,9 – 96,8)	99,9 (99,9 – 99,9)	90,6 (89,4 – 92,1)	95,0 (94,4 – 95,9)
D (22 / 18%)	99,9 (99,9 – 99,9)	94,8 (92,6 – 96,2)	99,9 (99,9 – 99,9)	94,6 (91,5 – 95,6)	99,9 (99,9 – 99,9)	89,3 (84,7 – 90,7)	94,4 (91,7 – 95,1)
P-value*	0,756	0,725	0,516	0,059	0,816	0,168	0,168

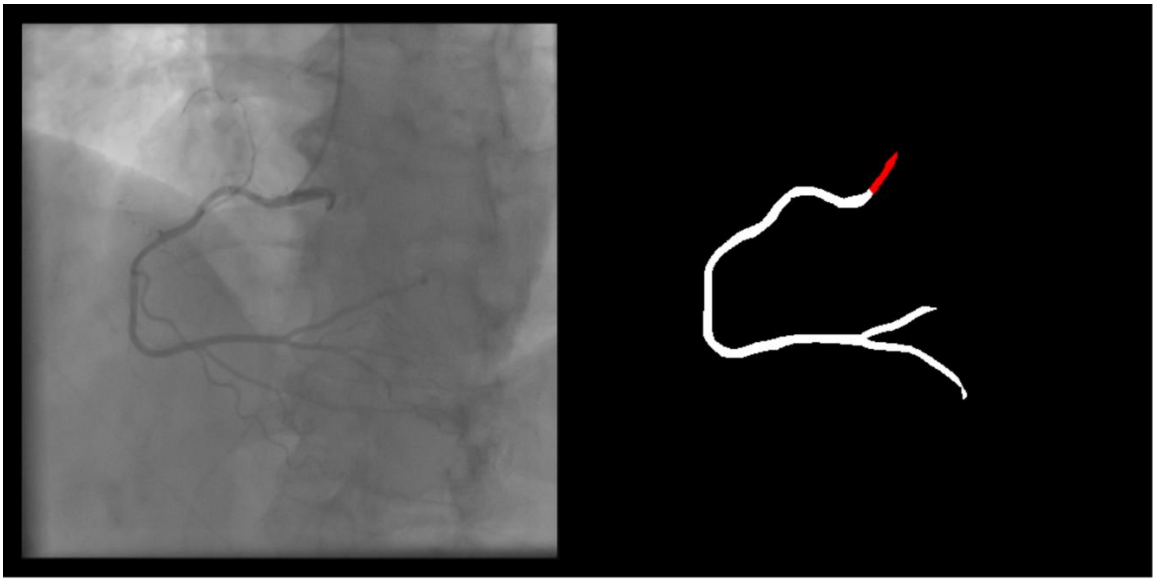
Supplementary table 16: GSS and individual parameter scores. Median (IQR) in bold due to non-normal distribution.

Parameter	Criteria
GSS (Median (IQ 25th – 75th))	**92 (87 – 96)**
GSS (Mean ± Standard Deviation)	90 ± 8
Main Vessel Segmentation (N / %)	108 (99,1)
Main Vessel Gaps (N / %)	104 (95,4)
Catheter to Vessel Transition (N / %)	103 (94,5)
Branch Segmentation (N / %)	101 (92,7)
Branch Gaps (N / %)	28 (25,7)
Coronary Artifacts (N / %)	95 (87,2)
Catheter Segmentation (N / %)	106 (97,2)
Catheter Gaps (N / %)	78 (71,6)
Catheter Artifacts (N / %)	52 (47,7)
Catheter Location (N / %)	105 (96,3)
Catheter Thickness (N / %)	104 (95,4)

Supplementary Figures

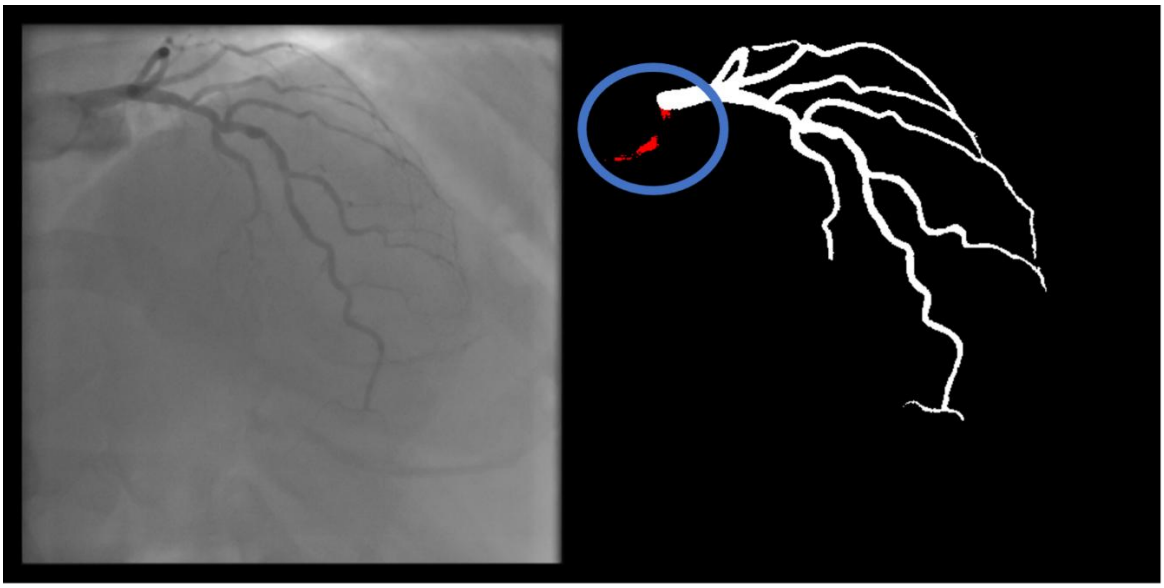


Supplementary figure 1: A test case with too many annotated vessels (not used for training).

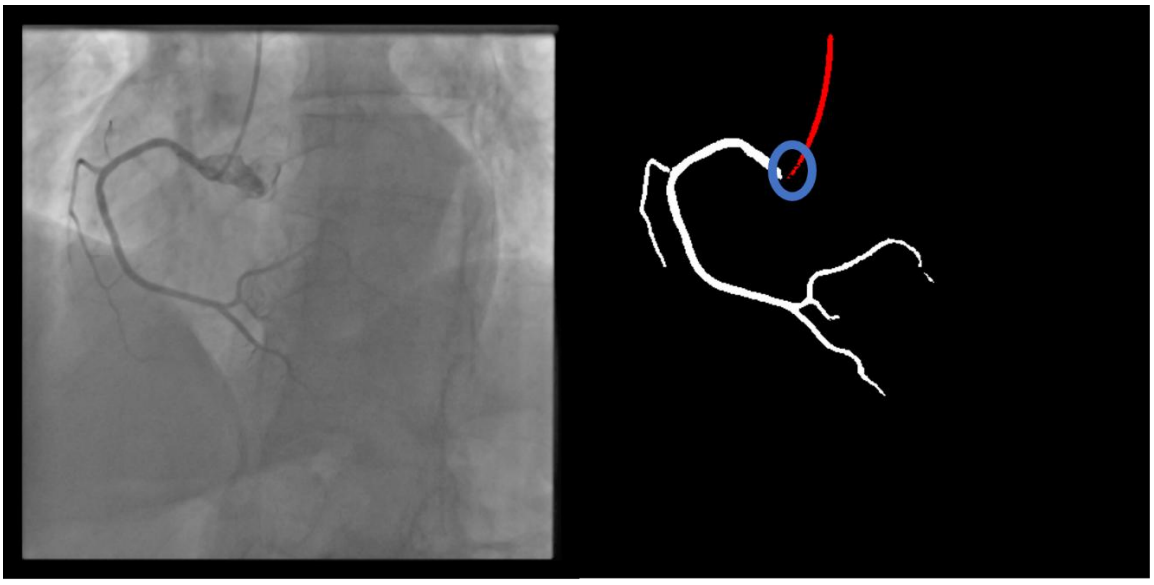


Supplementary figure 2: A test case with too few annotations (not used for training).

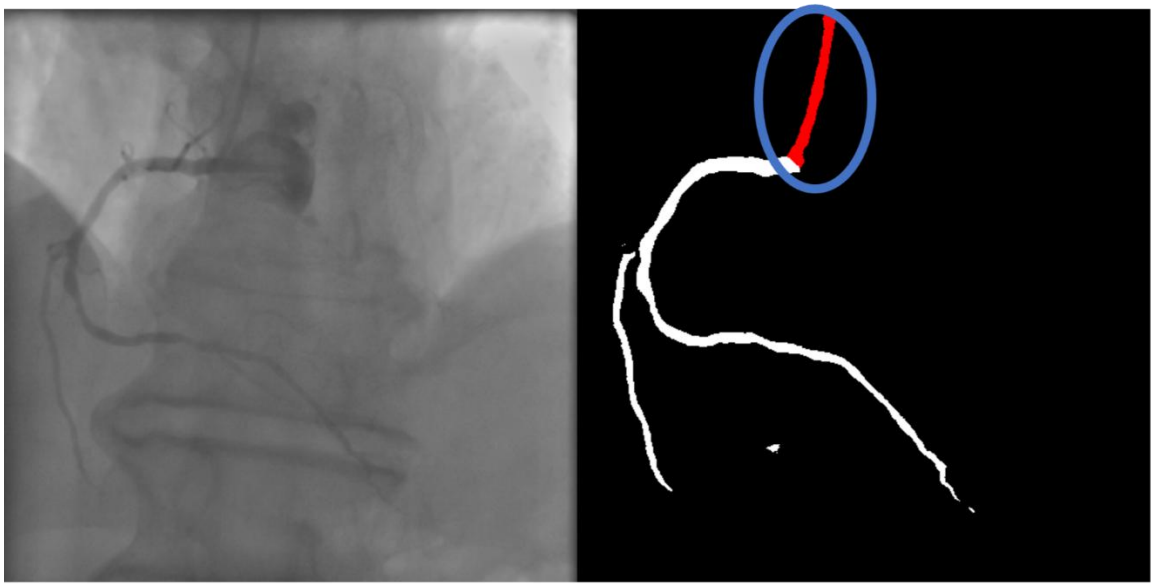
Appendix



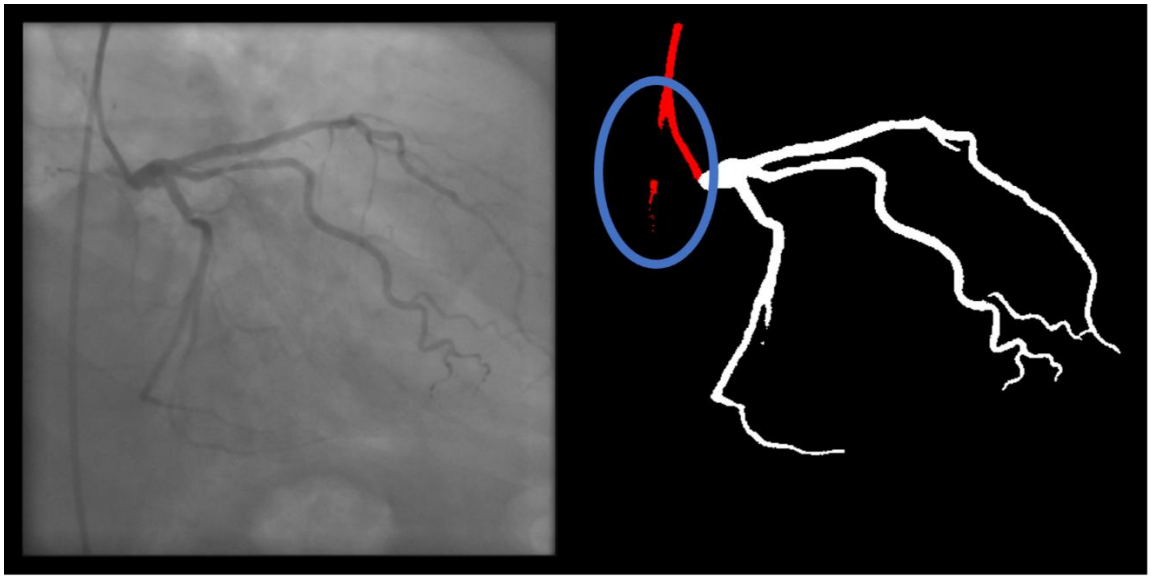
Supplementary figure 3: contrast backflow leads the AI model to disregard most of the catheter.



Supplementary figure 4: minor catheter gaps, possibly facilitated by contrast backflow.

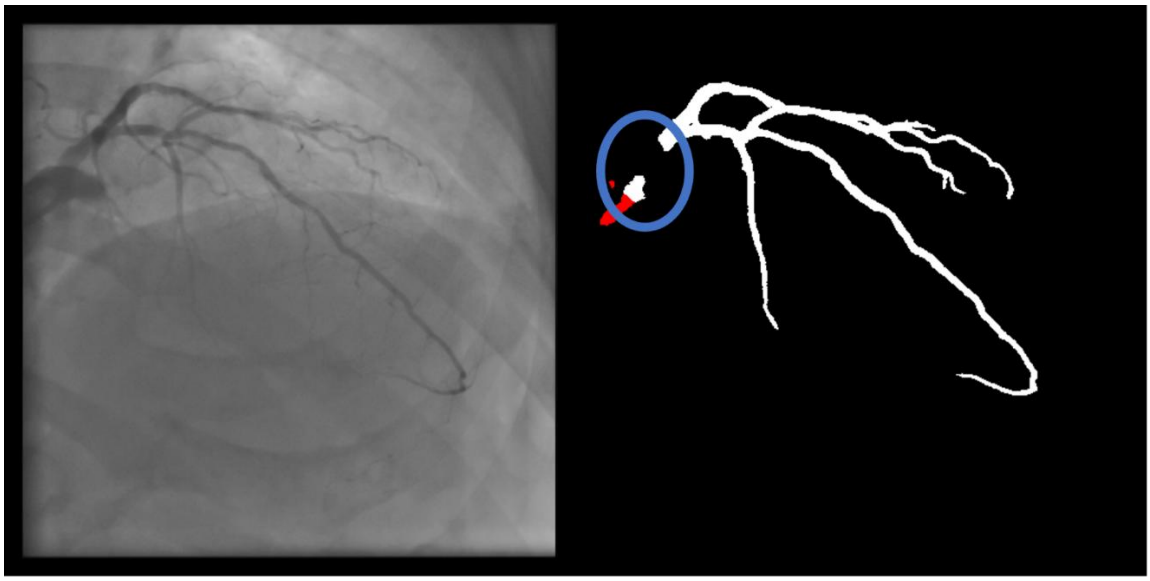


Supplementary figure 5: catheter thickness is overestimated.

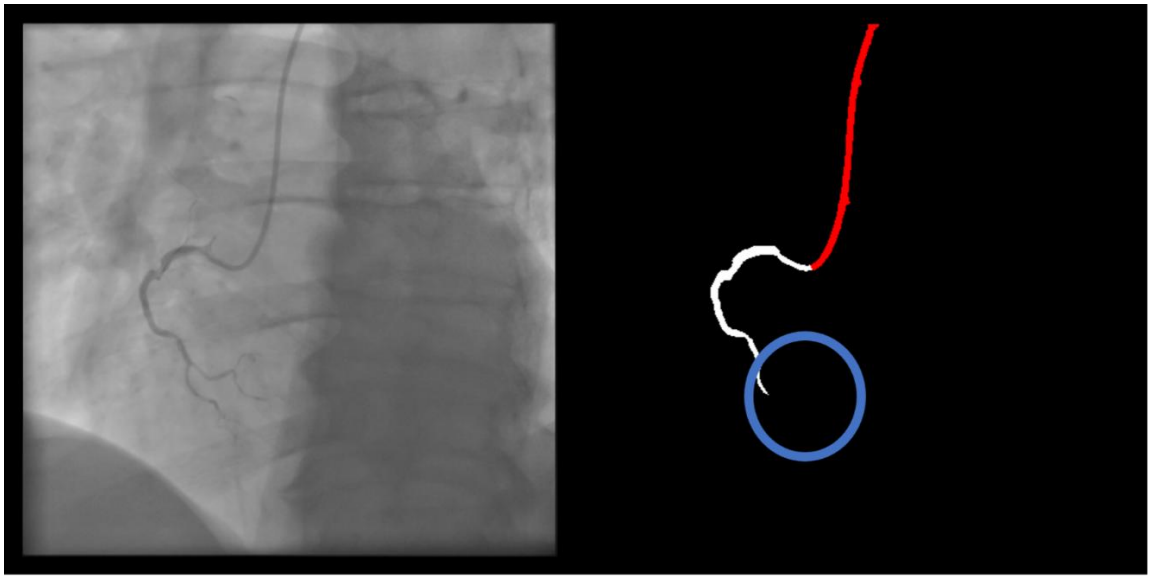


Supplementary figure 6: a small part of the catheter in the descending aorta was segmented in a femoral access case.

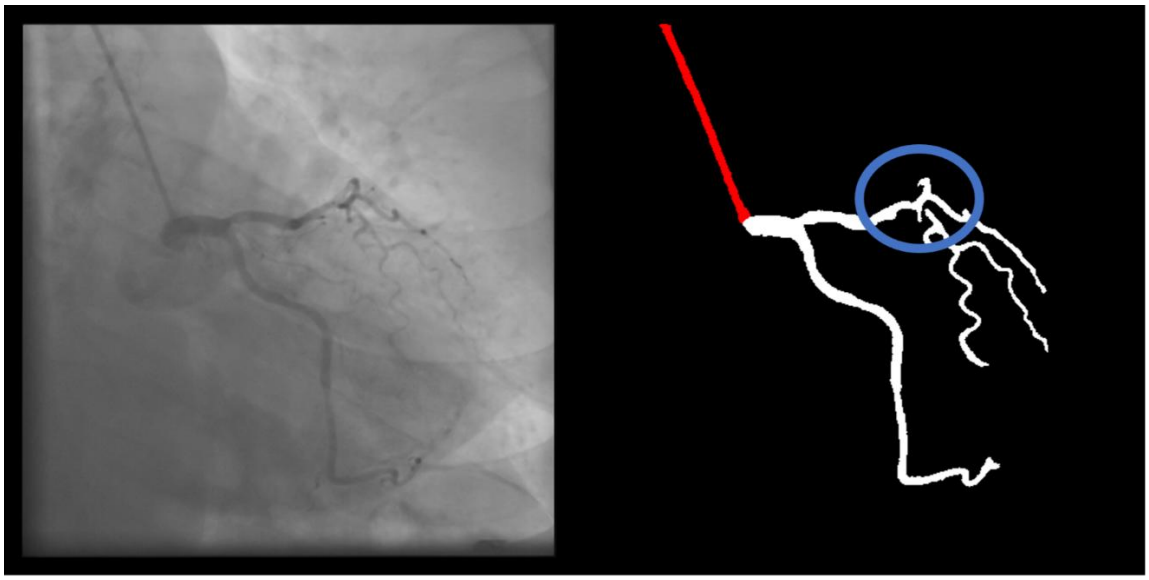
Appendix



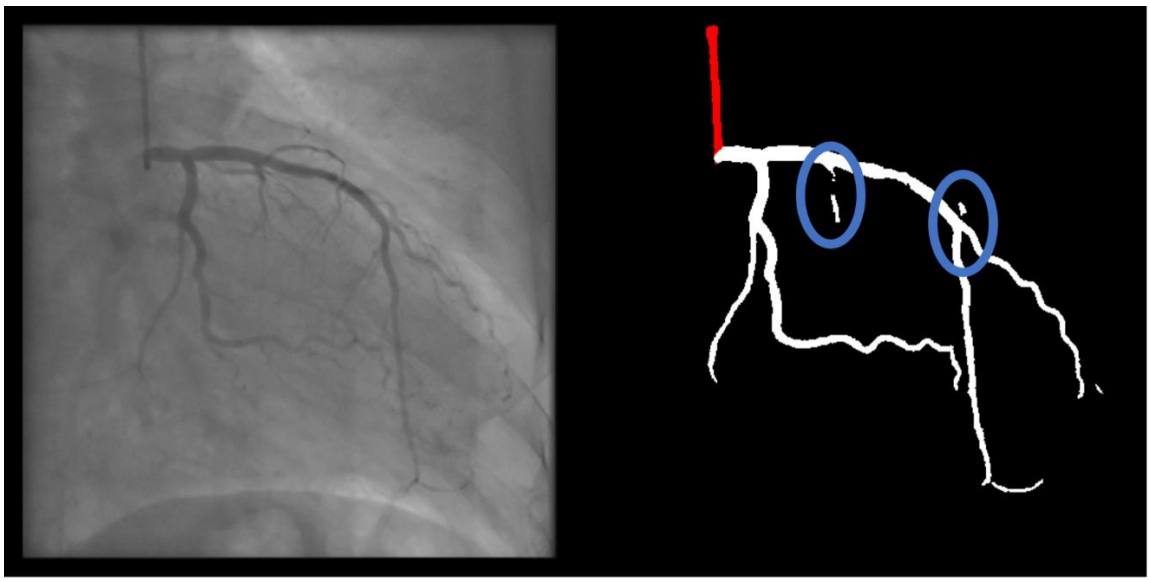
Supplementary figure 7: a major part of the left main was not segmented.



Supplementary figure 8: the postero-lateral branch and much of the posterior descending artery were missed. The enhanced AI model did not miss these vessels.

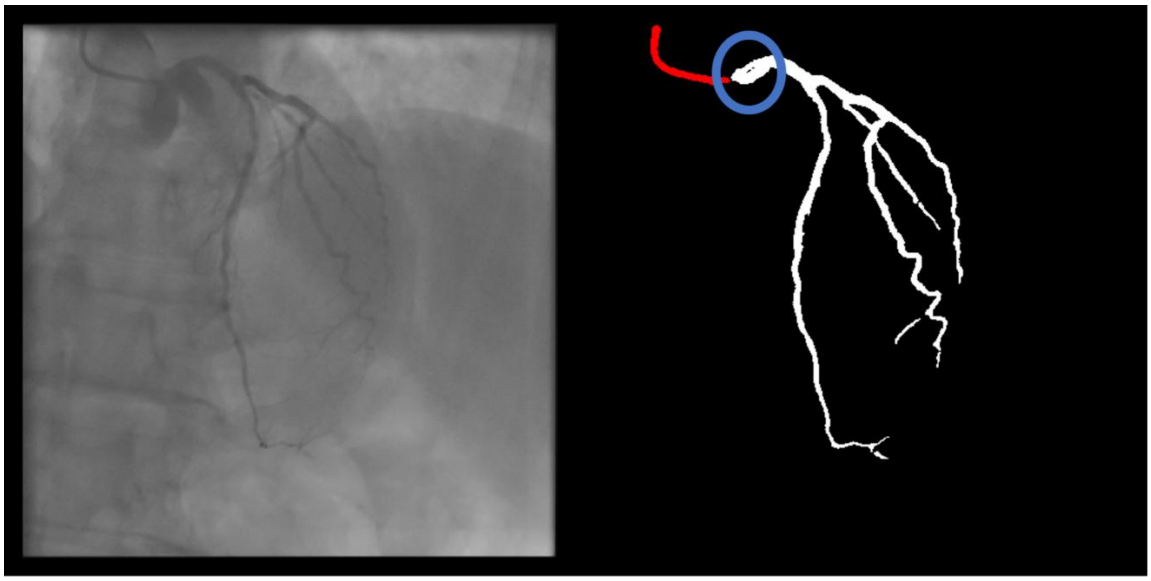


Supplementary figure 9: a small gap is visible in the left anterior descending artery.

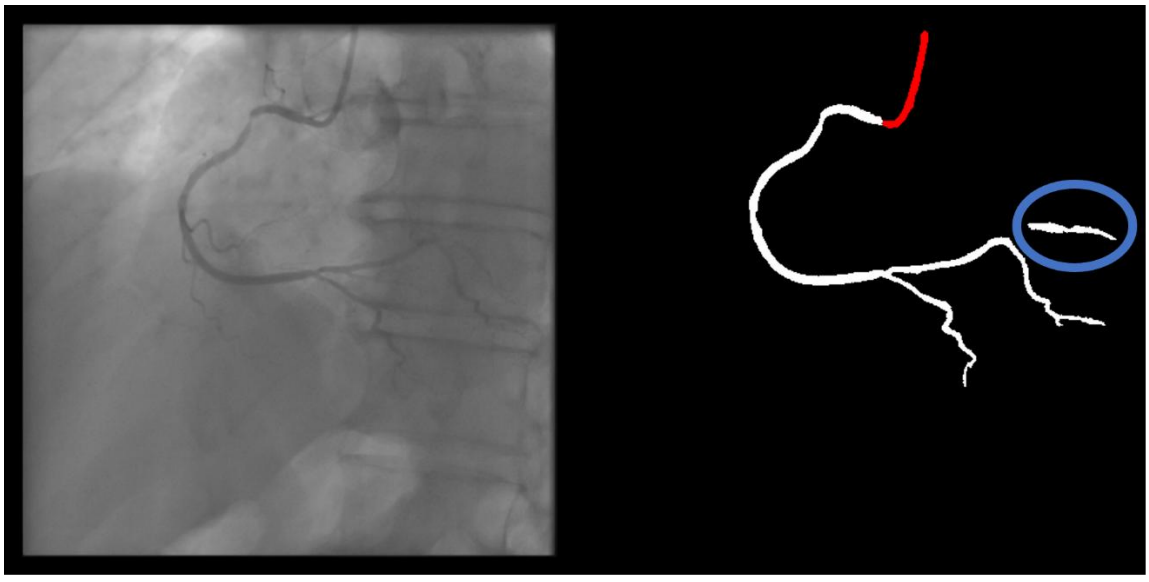


Supplementary figure 10: small gaps are visible in branches.

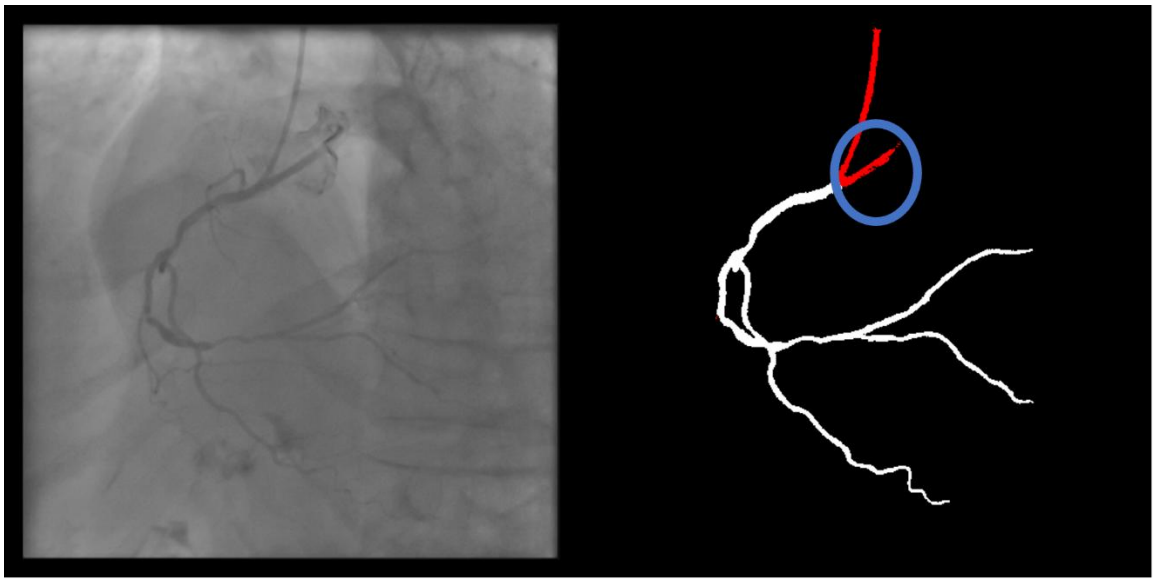
Appendix



Supplementary figure 11: contrast backflow renders the transition less discernible, leading the model to miss the transition zone.

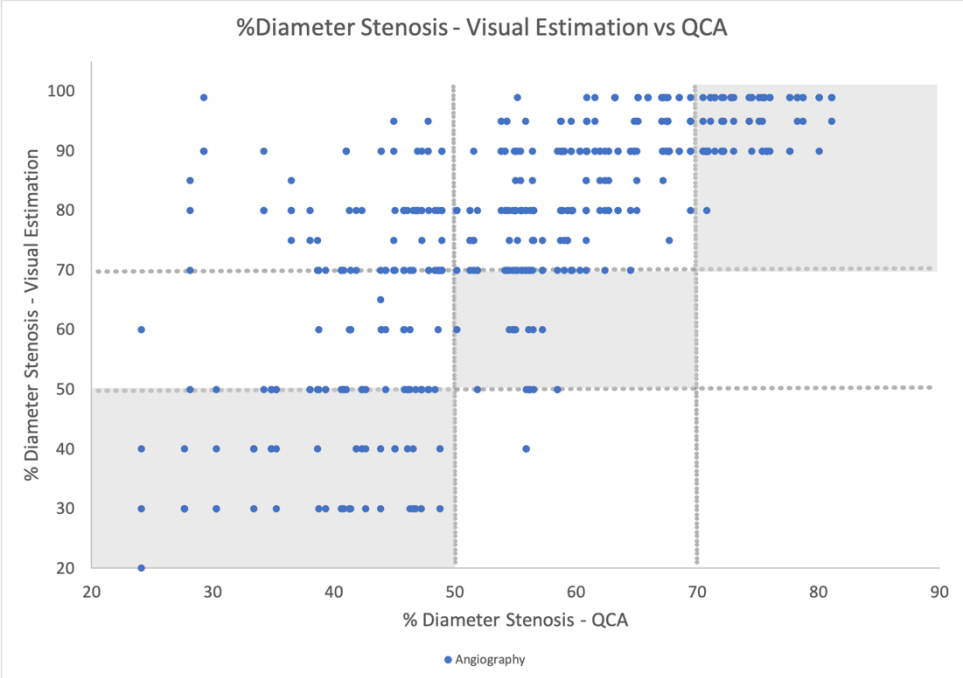


Supplementary figure 12: a part of the intervertebral disk and vertebra are mislabeled as coronary.

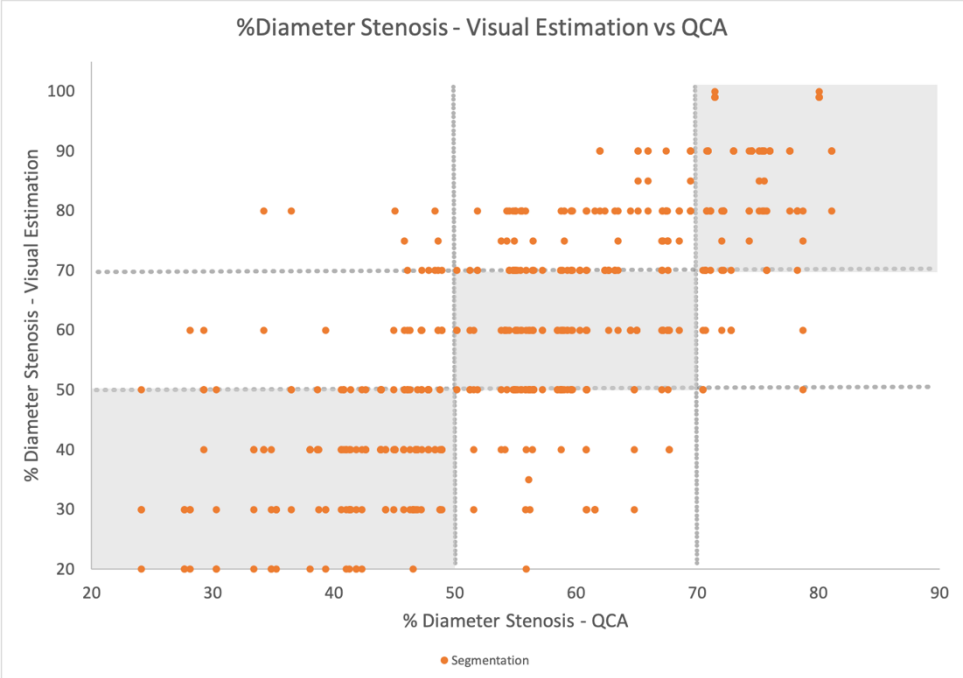


Supplementary figure 13: contrast backflow is mislabeled as catheter.

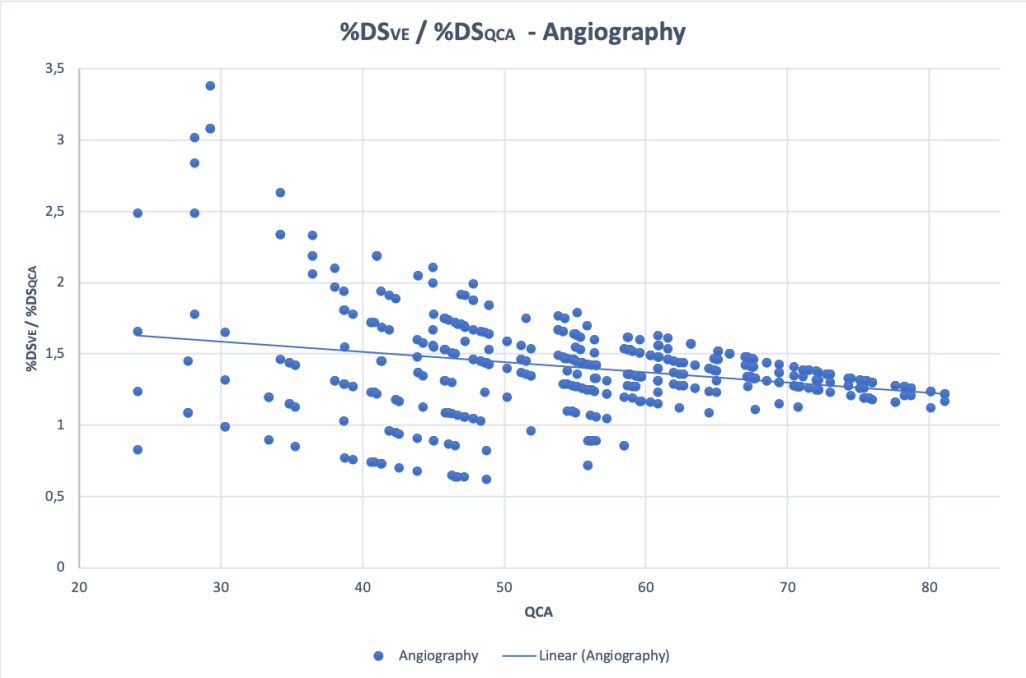
Appendix



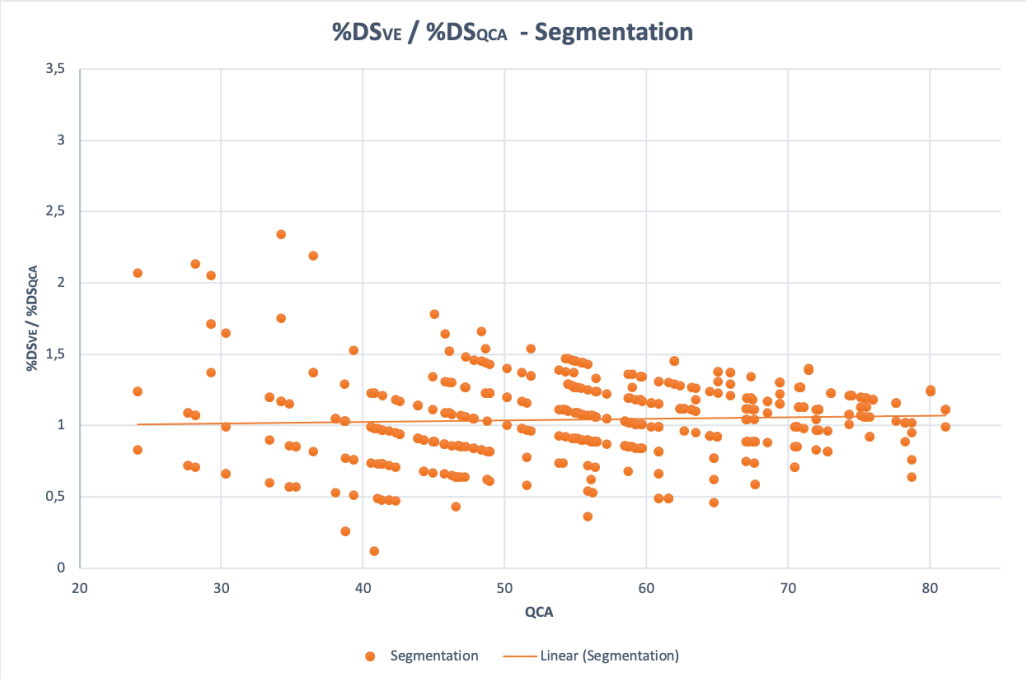
Supplementary figure 14: scatterplot of angiography %DSVE vs %DSQCA. Outside the most severe %DSQCA stratum, the discrepancy and overestimation tendency is clear.



Supplementary figure 15: scatterplot of segmentation %DSVE vs %DSQCA. Measurements are mostly in agreement between the two metrics.



Supplementary figure 16: scatterplot of angiography %DSVE / %DSQCA plotted in the Y axis vs %DSQCA plotted in the X axis. %DSVE only approximates %DSQCA in the most severe lesions, as evident by both the scatter and trendline. Additionally, the ratio is very predominately > 1, illustrating the tendency for overestimation.



Supplementary figure 17: scatterplot of segmentation %DSVE / %DSQCA plotted in the Y axis vs %DSQCA plotted in the X axis. The trendline and measurements are closer to 1 indicated reduced discrepancy and less overestimation tendency.

The role of the author in the research projects

The rules for presentation of PhD dissertations in the Faculty of Medicine, University of Lisbon were approved by the Scientific Council in 24th of June 2014. Details about the role of Miguel Nobre Menezes in the obtention of the results presented in this PhD dissertation are presented ahead as required by the referred document. Following the principles supported by the International Committee of Medical Journal Editors, Miguel Nobre Menezes had significant roles in the obtention of results presented on this PhD dissertation. Miguel Nobre Menezes was the main investigator and first author of all projects. Thus, essential contributions were made to the conception or design of the work; to the acquisition, analysis, and interpretation of data for the work; as well as drafting and revising it critically for important intellectual content. The final versions were approved to be submitted and published, ensuring the accuracy and integrity of all parts of the work.