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, MSc² | Beatriz Silva MD, MSc² | João Lourenço Silva MSc³ | Tiago Rodrigues MD, MSc¹,² | João Silva Marques MD, MSc¹,² | Cláudio Guerreiro MD, MSc⁴ | João Pedro Guedes MD, MSc⁵ | Manuel Oliveira-Santos MD, MSc⁶,⁷ | Arlindo L. Oliveira PhD³ | Fausto J. Pinto PhD¹,²

¹Structural and Coronary Heart Disease Unit, Cardiovascular Center of the University of Lisbon, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
²Departamento de Coração e Vasos, Serviço de Cardiologia, CHULN Hospital de Santa Maria, Lisboa, Portugal
³INESC-ID/Instituto Superior Técnico, Lisbon, Portugal
⁴Department of Cardiology, Centro Hospitalar de Vila Nova de Gaia, Vila Nova de Gaia, Portugal
⁵Unidade de Hemodinâmica e Cardiologia de Intervenção, Serviço de Cardiologia, Centro Hospitalar Universitário do Algarve, Hospital de Faro, Faro, Portugal
⁶Unidade de Intervenção Cardiovascular, Serviço de Cardiologia do Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
⁷Pólo das Ciências da Saúde, Unidade Central, Azinhaga de Santa Comba, Celas, Faculty of Medicine, University of Coimbra, Coimbra, Portugal

Abstract
Background: Visual assessment of the percentage diameter stenosis (%DSVE) 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’ %DSVE in angiography versus AI-segmented images.

Methods: Quantitative coronary analysis (QCA) %DS (%DSQCA) was previously performed in our published validation dataset. Operators were asked to estimate %DSVE of lesions in angiography versus AI-segmented images in separate sessions and differences were assessed using angiography %DSQCA as reference.

Results: A total of 123 lesions were included. %DSVE 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 %DSQCA 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 %DSQCA < 50% lesions, but not %DSQCA > 70% lesions. Agreement between %DSQCA/%DSVE across %DSQCA strata (<50%, 50%–70%, >70%) was approximately twice in the segmentation group (60.4% vs. 30.1%; p < 0.001). %DSVE inter-operator differences were smaller with segmentation.

Conclusion: %DSVE was much less discrepant with segmentation versus angiography. Overestimation of %DSQCA < 70% lesions with angiography was especially common. Segmentation may reduce %DSVE 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; %DSQCA, percentage diameter stenosis by QCA; %DSVE, percentage diameter stenosis by visual estimation.
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 (%DSVE) or by means of direct semi-automatic measurement with quantitative coronary angiography (CAG) (QCA—%DSQCA). 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. Disagreements of lesion severity by visual inspection versus QCA may be clinically relevant, as they have been associated to the likelihood of clinical events.

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. 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. Despite this, the adoption of physiology remains low.

As a result, non-invasive and automatic tools that reduce the heterogeneity of CAG interpretation are desirable. 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. We have recently developed AI models capable of accurate CAG segmentation. 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. Recently, we published the results of our validation study with an additional dataset of 117 images. 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.

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. %DSVE 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 %DSQCA 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 %DSVE/%DSQCA plotted in a scatter graphic in the Y-axis, against the %DSQCA 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,
generating a total of 984 %DSVE estimates. Most lesions had a %DSQCA of 50%–70% (Tables 1 and 2).

3.2 | Lesion severity assessment

3.2.1 | Overall results

Lesion severity by %DSVE was estimated to be higher with angiography than with segmentation. While there were significant differences between %DSVE and %DSQCA in both groups, the overall absolute difference in %DS was lower in the segmentation group. Additionally, no statistical difference was found between the segmentation %DSVE and the %DSQCA 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, both considering the overall sample and individual operators (Table 4).

3.3 | Results per QCA severity strata

For lesions with %DSQCA > 70%, there was a statistically significant higher %DSVE estimation in both the angiography and segmentation groups. In the segmentation group, %DSVE values were lower and closer to %DSQCA 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 %DSQCA of 50%–70%, there was a very large and significant difference between %DSVE with angiography, with median estimates of 80% or 90% across all operators and very low rates of strata agreement (0%–19%).

When %DSVE was undertaken with segmentation, differences with %DSQCA 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 %DSQCA < 50%, there were also significantly higher estimates %DSVE with angiography for all operators and low rates of strata agreement (11.6%–32.6%), but to a lesser degree than in the %DSQCA 50%–70% strata.

For %DSVE with segmentation, there was no statistically significant difference with %DSQCA 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 DSQCA > 70% strata for either the angiography or segmentation
### Table 3

Diameter stenosis assessed by visual estimation and QCA across operators, lesion severity strata and overall.

<table>
<thead>
<tr>
<th>%DS&lt;sub&gt;QCA&lt;/sub&gt; Stratum</th>
<th>Group / Parameter</th>
<th>%DS&lt;sub&gt;QCA&lt;/sub&gt; - Angiography</th>
<th>%DS&lt;sub&gt;QCA&lt;/sub&gt; - Segmentation</th>
<th>%DS&lt;sub&gt;QCA&lt;/sub&gt;</th>
<th>%DS&lt;sub&gt;VE&lt;/sub&gt; - Segmentation</th>
<th>%DS&lt;sub&gt;VE&lt;/sub&gt; - Angiography</th>
<th>%DS&lt;sub&gt;VE&lt;/sub&gt; - All</th>
<th>%DS&lt;sub&gt;QCA&lt;/sub&gt; - All</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases</td>
<td>%DS</td>
<td>84 +/- 15 90 (80 - 95)</td>
<td>78 +/- 19 80 (50 - 90)</td>
<td>71 +/- 21 70 (50 - 90)</td>
<td>76 +/- 21 80 (70 - 90)</td>
<td>77 +/- 20 80 (70 - 90)</td>
<td>64 +/- 17 60 (50 - 80)</td>
<td>54 +/- 23 60 (40 - 70)</td>
</tr>
<tr>
<td>%DS&lt;sub&gt;VE&lt;/sub&gt; vs %DS&lt;sub&gt;QCA&lt;/sub&gt; P Value*</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.327</td>
<td>0.998</td>
</tr>
<tr>
<td>Inter-Operator Difference**</td>
<td>&lt; 0.001</td>
<td>-</td>
<td>&lt; 0.001</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 70%</td>
<td>%DS</td>
<td>95 +/- 4 95 (90 - 99)</td>
<td>96 +/- 4 99 (90 - 99)</td>
<td>94 +/- 5 90 (90 - 99)</td>
<td>96 +/- 4 99 (90 - 99)</td>
<td>95 +/- 4 97 (90 - 99)</td>
<td>82 +/- 12 80 (70 - 90)</td>
<td>82 +/- 12 83 (70 - 90)</td>
</tr>
<tr>
<td>%DS&lt;sub&gt;VE&lt;/sub&gt; vs %DS&lt;sub&gt;QCA&lt;/sub&gt; P Value*</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.008</td>
<td>0.016</td>
<td>0.088</td>
</tr>
<tr>
<td>Inter-Operator Difference**</td>
<td>0.115*</td>
<td>0.334*</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-70%</td>
<td>%DS</td>
<td>88 +/- 8 90 (80 - 95)</td>
<td>83 +/- 12 80 (75 - 90)</td>
<td>76 +/- 15 80 (70 - 90)</td>
<td>84 +/- 12 90 (75 - 90)</td>
<td>83 +/- 13 85 (75 - 90)</td>
<td>66 +/- 14 60 (58 - 80)</td>
<td>60 +/- 17 60 (50 - 70)</td>
</tr>
<tr>
<td>%DS&lt;sub&gt;VE&lt;/sub&gt; vs %DS&lt;sub&gt;QCA&lt;/sub&gt; P Value*</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.002</td>
<td>0.707</td>
<td>0.837</td>
</tr>
<tr>
<td>Inter-Operator Difference**</td>
<td>&lt; 0.001</td>
<td>-</td>
<td>&lt; 0.001</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50%</td>
<td>%DS</td>
<td>71 +/- 17 80 (60 - 80)</td>
<td>62 +/- 20 65 (40 - 80)</td>
<td>53 +/- 17 50 (40 - 70)</td>
<td>57 +/- 20 50 (40 - 70)</td>
<td>61 +/- 20 60 (43 - 80)</td>
<td>52 +/- 13 50 (40 - 60)</td>
<td>33 +/- 12 30 (20 - 40)</td>
</tr>
<tr>
<td>%DS&lt;sub&gt;VE&lt;/sub&gt; vs %DS&lt;sub&gt;QCA&lt;/sub&gt; P Value*</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.116</td>
<td>0.538</td>
</tr>
<tr>
<td>Inter-Operator Difference**</td>
<td>&lt; 0.001</td>
<td>-</td>
<td>&lt; 0.001</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *Wilcoxon test paired samples; **Friedman test; significant p values in bold.

Abbreviations: %DS<sub>QCA</sub>, diameter stenosis by QCA; %DS<sub>VE</sub>, diameter stenosis by visual estimation; QCA: quantitative coronary angiography.
groups. Considering the overall sample and the remaining strata, there were significant differences between operators in absolute %DSVE 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 %DSQCA 50%–70% strata. For the segmentation group, significant differences were only found in the %DSQCA < 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 %DSQCA > 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 %DSVE > 80%–90% in both groups. However, even in this stratum, absolute %DSVE values were less discrepant with %DSQCA in the segmentation group.

When considering lesions with %DSQCA of 51%–69%, overestimation was very frequent and pronounced in the angiography group, but not in the segmentation group. Indeed, the mean %DSVE 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 %DSQCA lesion severity strata.

The scatterplots (Figure 2, Figures S1 and S2) clearly illustrate these findings, as the discrepancy outside %DSQCA > 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 %DSQCA > 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 %DSQCA 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).
### Table 4

Agreement between %DSQCA strata and %DSVE strata between angiography and segmentation, across operators and overall.

<table>
<thead>
<tr>
<th>%DSQCA Strata/Group</th>
<th>Angiography All</th>
<th>Segmentation All</th>
<th>Angiography Operator 1</th>
<th>Segmentation Operator 1</th>
<th>Angiography Operator 2</th>
<th>Segmentation Operator 2</th>
<th>Angiography Operator 3</th>
<th>Segmentation Operator 3</th>
<th>Angiography Operator 4</th>
<th>Segmentation Operator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%DSVE/%DSQCA Strata Agreement (n/%)</td>
<td>148 (30,1)</td>
<td>297 (60,4)</td>
<td>27 (22,0)</td>
<td>65 (52,8)</td>
<td>35 (28,5%)</td>
<td>79 (64,2%)</td>
<td>47 (38,2)</td>
<td>80 (65,0)</td>
<td>39 (31,7)</td>
<td>73 (59,3)</td>
</tr>
<tr>
<td>P Value</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.007</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 70%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%DSVE/%DSQCA Strata Agreement (n/%)</td>
<td>88 (100)</td>
<td>78 (88,6%)</td>
<td>22 (100)</td>
<td>29 (90,9)</td>
<td>22 (100)</td>
<td>19 (86,4)</td>
<td>22 (100)</td>
<td>18 (81,8)</td>
<td>22 (100)</td>
<td>21 (95,5)</td>
</tr>
<tr>
<td>P Value</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-70%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%DSVE/%DSQCA Strata Agreement (n/%)</td>
<td>17 (7,3)</td>
<td>107 (46,1)</td>
<td>0 (0)</td>
<td>30 (51,7)</td>
<td>2 (3,4%)</td>
<td>23 (39,7%)</td>
<td>11 (19)</td>
<td>32 (55,2)</td>
<td>4 (6,9)</td>
<td>22 (37,9)</td>
</tr>
<tr>
<td>P Value</td>
<td>0.009</td>
<td>NA</td>
<td>0.243</td>
<td>0.008</td>
<td>0.113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%DSVE/%DSQCA Strata Agreement (n/%)</td>
<td>43 (25,0)</td>
<td>112 (65,1)</td>
<td>5 (11,6)</td>
<td>15 (34,9)</td>
<td>11 (25,6)</td>
<td>37 (86,0)</td>
<td>14 (32,6)</td>
<td>30 (69,8)</td>
<td>13 (30,2)</td>
<td>30 (69,8)</td>
</tr>
<tr>
<td>P Value</td>
<td>&lt; 0.001</td>
<td>0.001</td>
<td>0.590</td>
<td>0.003</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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 %DSVE vs. %DSQCA. The difference between the two %DSVE is visually clear. [Color figure can be viewed at wileyonlinelibrary.com]

FIGURE 3 Combined scatterplots of angiography and segmentation %DSVE/%DSQCA plotted in the Y-axis vs. %DSQCA plotted in the X-axis. The difference between the two %DSVE is visually clear. [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 and continue to be found today across the world, even in very large cohorts. 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 or scientific societies. However, in clinical practice, the estimation of stenosis severity continues to be undertaken mostly by visual estimation. The fact that correlation between QCA and FFR has not been demonstrated to be superior to that of visual estimation, or that clinical trials often don’t require QCA measurements for including patients, 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. 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 and is recommended in current guidelines, 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. 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, 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, Al-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. We too have previously published our results regarding the development of Al models capable of highly accurate segmentation, encompassing more than 500 images as a whole. 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. 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.

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 (% DSQCA 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


SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.