







Article

Development and External Validation of [18F]FDG PET-CT-Derived Radiomic Models for Prediction of Abdominal Aortic Aneurysm Growth Rate

Simran Singh Dhesi ^{1,†}, Pratik Adusumilli ^{1,†} , Nishant Ravikumar ² , Mohammed A. Waduud ³, Russell Frood ¹ , Alejandro F. Frangi ⁴, Garry McDermott ⁵, James H. F. Rudd ⁶, Yuan Huang ⁶, Jonathan R. Boyle ⁷ , Maysoon Elkhawad ⁷, David E. Newby ⁸, Nikhil Joshi ⁸, Jing Yi Kwan ^{3,9} , Patrick Coughlin ⁹, Marc A. Bailey ^{3,9} and Andrew F. Scarsbrook ^{1,10,*} 

¹ Department of Clinical Radiology, Leeds Teaching Hospitals NHS Trust, Leeds LS9 7TF, UK; s.dhesi2@nhs.net (S.S.D.); p.adusumilli@nhs.net (P.A.); russellfrood@nhs.net (R.F.)

² School of Computer Science, University of Leeds, Leeds LS2 9JT, UK; n.ravikumar@leeds.ac.uk

³ Leeds Institute of Cardiovascular & Metabolic Medicine, School of Medicine, University of Leeds, Leeds LS2 9JT, UK; m.a.waduud@leeds.ac.uk (M.A.W.); jingyi.kwan@nhs.net (J.Y.K.); m.a.bailey@leeds.ac.uk (M.A.B.)

⁴ Christabel Pankhurst Institute for Health Technology Research and Innovation, University of Manchester, Manchester M13 9PS, UK; a.frangi@manchester.ac.uk

⁵ Department of Medical Physics and Engineering, Leeds Teaching Hospitals NHS Trust, Leeds LS9 7TF, UK; garry.mcdermott@nhs.net

⁶ Department of Medicine, University of Cambridge, Cambridge CB2 0SP, UK; jhfr2@cam.ac.uk (J.H.F.R.); yh288@dpmms.cam.ac.uk (Y.H.)

⁷ Department of Vascular Surgery, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK; jonathan.boyle2@nhs.net (J.R.B.); maysoon.elkhawad@nhs.net (M.E.)

⁸ British Heart Foundation Centre of Research Excellence, University of Edinburgh, Edinburgh EH16 4SA, UK; d.e.newby@ed.ac.uk (D.E.N.); nikhil.joshi7@nhs.net (N.J.)

⁹ Leeds Vascular Institute, Leeds Teaching Hospitals NHS Trust, Great George Street, Leeds LS1 3EX, UK; patrick.coughlin1@nhs.net

¹⁰ Leeds Institute of Health and Research, Faculty of Medicine, University of Leeds, Leeds LS2 9LN, UK

* Correspondence: a.f.scarsbrook@leeds.ac.uk

† These authors contributed equally to this work.



Academic Editor: Carmelo Militello

Received: 22 December 2024

Revised: 27 January 2025

Accepted: 31 January 2025

Published: 5 February 2025

Citation: Dhesi, S.S.; Adusumilli, P.; Ravikumar, N.; Waduud, M.A.; Frood, R.; Frangi, A.F.; McDermott, G.; Rudd, J.H.F.; Huang, Y.; Boyle, J.R.; et al.

Development and External Validation of [18F]FDG PET-CT-Derived Radiomic Models for Prediction of Abdominal Aortic Aneurysm Growth Rate. *Algorithms* **2025**, *18*, 86.

<https://doi.org/10.3390/a18020086>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Objective (1): To develop and validate a machine learning (ML) model using radiomic features (RFs) extracted from [18F]FDG PET-CT to predict abdominal aortic aneurysm (AAA) growth rate. Methods (2): This retrospective study included 98 internal and 55 external AAA patients undergoing [18F]FDG PET-CT. RFs were extracted from manual segmentations of AAAs using PyRadiomics. Recursive feature elimination (RFE) reduced features for model optimisation. A multi-layer perceptron (MLP) was developed for AAA growth prediction and compared against Random Forest (RF), XGBoost, and Support Vector Machine (SVM). Accuracy was evaluated via cross-validation, with uncertainty quantified using dropout (MLP), standard deviation (RF), and 95% prediction intervals (XGBoost). External validation used independent data from two centres. Ground truth growth rates were calculated from serial ultrasound (US) measurements or CT volumes. Results (3): From 93 initial RFs, 29 remained after RFE. The MLP model achieved an MAE \pm SEM of $1.35 \pm 3.2e-4$ mm/year with the full feature set and $1.35 \pm 2.5e-4$ mm/year with RFE. External validation yielded $1.8 \pm 8.9e-8$ mm/year. RF, XGBoost, and SVM models produced comparable accuracies internally ($1.4-1.5$ mm/year) but showed higher errors during external validation ($1.9-1.97$ mm/year). The MLP model demonstrated reduced uncertainty with the full feature set across all datasets. Conclusions (4): An MLP model leveraging [18F]FDG PET-CT radiomics accurately predicted AAA growth rates and generalised well to external data. In the future, more sophisticated stratification could guide individualised patient care, facilitating risk-tailored management of AAAs.

Keywords: abdominal aortic aneurysm; positron emission tomography–computed tomography; machine learning; radiomics

1. Introduction

Abdominal aortic aneurysms (AAA) are defined as an aortic diameter over 3.0 cm, they are often asymptomatic prior to potentially fatal complications such as dissection or rupture [1]. Reported AAA prevalence varies depending on country and patient demographics, affecting up to 2.2% and 2.5% of the American and European populations, respectively [2]. The main risk factors for AAA include age, male gender, family history, hypertension, hypercholesterolaemia and smoking [3–5]. Risk of rupture is increased in patients with a vessel diameter greater than 5.5 cm, a growth rate above 1 cm/year and female gender [6–9].

Advanced age and male sex have led to the adoption of a national screening programme for men over 65 in the UK and US Preventive Services Task Force (USPSTF) recommending one-time ultrasound screening in men aged 65 to 75 years who have ever smoked [10,11]. Current AAA management guidelines advocate the use of diameter as a metric for risk stratification. AAAs are monitored with ultrasound until they reach the intervention threshold of 5.5 cm. In the UK, AAAs between 3.0–4.5 cm and 4.5–5.5 cm are monitored annually and every 3 months, respectively. Ultrasound provides anatomical information at fixed time points but does not predict potential growth rate. A biomarker to predict AAA growth and personalise surveillance regimens or intervention thresholds is lacking in the field.

Previous studies have evaluated AAA metabolic activity using fluorine-18-2-deoxy-D-glucose ([18F]FDG) positron emission tomography–computed tomography (PET-CT) as a marker of pathological wall weakening [12,13]. There are conflicting data on the utility of semi-quantitative analysis using standardised uptake value (SUV) metrics for prediction of future AAA growth and rupture [12–17]. More sophisticated imaging feature analysis may have greater potential to predict AAA growth. Radiomics is a process involving extraction of high dimensional features from medical imaging allowing quantitative analysis of the distribution and relationship of pixel grey levels [18]. This technique has been extensively studied in the oncology setting for outcome prediction modelling [19–24]. There is currently a paucity of data on the potential use of radiomic analysis for predictive modelling in non-oncological settings. In particular, the use of radiomic features derived from [18F]FDG PET-CT for prediction of AAA growth has not yet been reported to the best of our knowledge.

The aim of this study was to explore the utility of a machine learning (ML)-based regression model utilising radiomic features extracted from [18F]FDG PET-CT to predict future AAA growth rate.

2. Materials and Methods

A transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) study design was employed to assess the possible benefit of [18F]FDG PET-CT-derived radiomics in patients with AAA [25] (Supplementary Table S1). A sample size estimation was performed using a medium effect size (Cohen's $d = 0.5$), a significance level of 0.05, and a statistical power of 90%, determining that a minimum of 42 paired observations were required for adequate power.

2.1. Internal Patient Selection

An internal cohort consisting of data from consecutive patients with AAA who underwent [18F]FDG PET-CT for investigation of suspected malignancy between January

2009 and March 2020 who were subsequently enrolled in an AAA ultrasound surveillance programme were identified retrospectively from the institutional Radiology Information System (RIS) (CRIS, Wellbeing Software, <https://www.cdn-pacs.com/cris/> (accessed on 27 January 2025), Mansfield, UK). Additional independent cohorts from two external centres comprising prospectively identified patients diagnosed with AAA who had undergone PET-CT imaging were also collated. Exclusion criteria included patients who had undergone surgery or endovascular AAA repair (EVAR) prior to PET-CT, studies with significant motion artefact leading to mis-registration of the PET and CT components of the study and patients with tumours involving the aneurysmal segment of the aorta. Patient demographic details, clinical history, ultrasound follow-up data, treatment data, clinical outcome and follow-up duration were obtained from the institutional electronic patient record.

2.2. Imaging Acquisition and Reconstruction Parameters

All internal PET-CT studies were performed as part of routine clinical care using standardised departmental protocols. Imaging was acquired from skull base to upper thighs. No iodinated contrast media was administered. Serum blood glucose was routinely measured prior to imaging and if >10 mmol/L scanning was not performed. Patients fasted for six hours prior to intravenous injection of 4 MBq/kg of fluorine-18 [^{18}F]FDG.

Internal imaging prior to June 2010 was performed using a 16-slice Discovery STE PET/CT scanner (General Electric (GE) Healthcare, Chicago, IL, USA) and from June 2010 to October 2015 on a 64-slice Philips Gemini TF64 scanner (Philips Healthcare, Best, Netherlands) or a Discovery 690 scanner (GE Healthcare, Chicago, IL, USA). After October 2015 a Discovery 710 scanner (GE Healthcare, Chicago, IL, USA) replaced the Philips scanner. All machines used iterative reconstruction, computed tomography (CT) for attenuation correction, applied scatter and randoms correction. Image reconstruction and acquisition parameters are outlined in Table 1.

Table 1. Breakdown of PET-CT Scanners Used and Associated Timeframes.

Dates	Scanner	Reconstruction	Scatter Correction	Randoms Correction	Matrix	Voxel Size (x, y, z)
January 2009–May 2010	GE Healthcare STE	OSEM ¹	Convolution subtraction	Singles	128	$4.6875 \times 4.6875 \times 3.27$
June 2010–October 2015	Philips Gemini TF64	BLOB-OS-TF ²	SS-SIMUL ⁴	DLYD ⁵	144 or 169	$4 \times 4 \times 4$
June 2010–March 2020	GE Healthcare Discovery 690	VPFX ³	Model based	Singles	192	$3.65 \times 3.65 \times 3.27$
May 2015–March 2020	GE Healthcare Discovery 710	VPFX	Model based	Singles	192	$3.65 \times 3.65 \times 3.27$

¹ OSEM: Ordered Subset Expectation Maximisation, ² BLOB-OS-TF: Spherical Symmetric Basis Functions Ordered Subset Time of Flight, ³ VPFX: Time of Flight Ordered Subset Expectation Maximisation, ⁴ SS-SIMUL: CT-based attenuation correction and scatter correction, and ⁵ DLYD: delayed event subtraction.

To ensure alignment between PET and CT images, standard clinical practices at our institution employed careful patient positioning and instructions to breath hold to minimise motion artefact. Whilst mis-registration is not infrequently encountered in some body parts such as within the head and neck or lower thoracic regions; we rarely experienced significant mis-registration between the PET and CT components of the abdominal aorta being analysed. However, if this was an issue, the data were excluded from further analysis to avoid inaccuracies. Furthermore, all examinations were reviewed by radiologists with a minimum of 3 years of experience under the supervision of a dual-certified radiologist/nuclear medicine physician with over 15 years of experience in PET-CT.

Average AAA growth rates were calculated from the first and last antero-posterior ultrasound measurements obtained from the AAA surveillance programme. Scan intervals

were dictated by AAA size as per national guidelines [10]. All scans were performed by experienced sonographers or radiologists.

2.3. Image Pre-Processing

PET-CT imaging was exported in Digital Imaging and Communications in Medicine (DICOM) format from the institutional picture archiving and communication system (PACS) and converted to Neuroimaging Informatics Technology Initiative (NIfTI) format using the dicom2nifti Python package (Icometrix, Chicago, IL, United States).

Volumes of interest (VOI) were manually drawn around the dilated segment of the aorta (Figure 1a) and aortic arch blood pool (Figure 1b) using ITK-SNAP (Version 3.8.0, <http://www.itksnap.org> (accessed on 27 January 2025)). The CT component of the PET-CT was used to perform segmentation to provide accurate anatomical localisation. AAA VOI segmentation included the wall and associated blood pool of the aneurysmal segment of the abdominal aorta, aortic arch blood pool VOI segmentation included only the blood pool with exclusion of the wall. Care was taken to avoid areas of tumour-related metabolic activity if present. Segmentations were performed by a Radiologist with 3 years' experience under the supervision of a dual-certified Radiologist & Nuclear Medicine Physician with >15 years' experience of PET-CT.

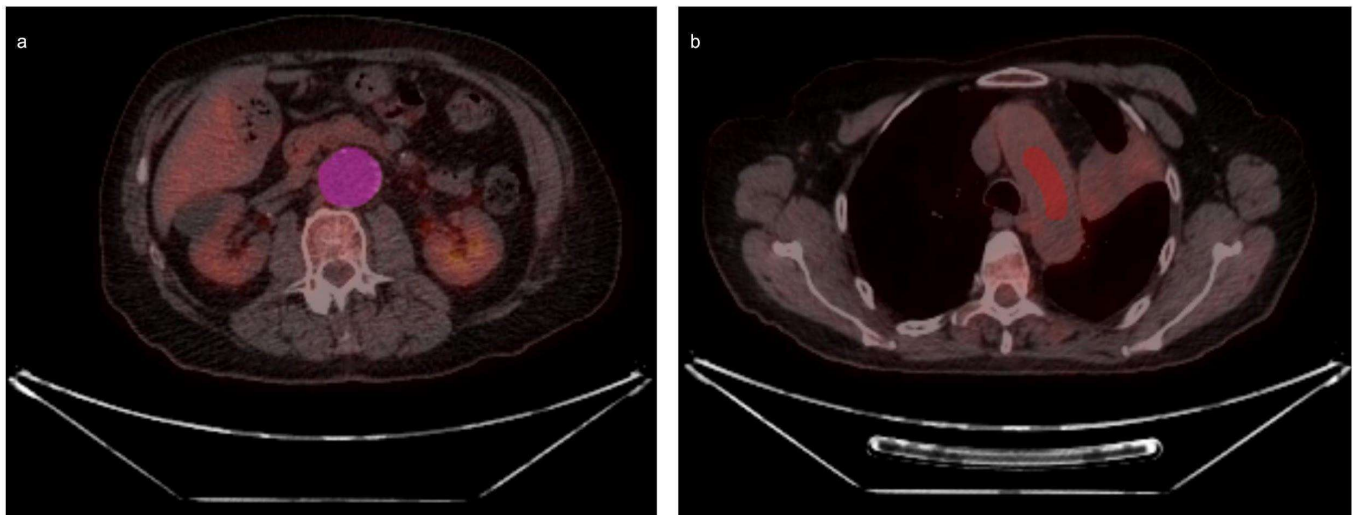


Figure 1. Axial fused [18F]FDG PET-CT images demonstrating a two-dimensional section of the three-dimensional volume of interest. (a) Volume of interest drawn around the abdominal aortic aneurysm. (b) Volume of interest drawn within the aortic arch blood pool.

2.4. Radiomic Feature Analysis

Radiomic features were extracted from the segmented VOI and analysed using PyRadiomics (Version 2.2.0, <https://github.com/AIM-Harvard/pyradiomics> (accessed on 27 January 2025)). Radiomic analysis was performed using the PET series, with the fused PET-CT series serving as a tool for accurate VOI placement through correct anatomical localisation (Figure 1a). PET-CT data were split into bin sizes of 0.5, previous studies have demonstrated better feature reproducibility and quantification when using fixed bin sizes over fixed bin counts [26–28]. Additionally, spatial resampling using voxel dimensions of $2 \times 2 \times 2$ mm was performed to eliminate voxel size variation. Otherwise, default hyperparameters were used unless stipulated in Supplementary Table S2.

A total of 18 first-order statistics and 75 textural features were extracted from the segmented VOIs. All features used in the study conformed to the definition provided by the Imaging Biomarker Standardization Initiative (IBSI) [29]. Among these, the “Maximum” first-order radiomic feature is mathematically equivalent to the commonly used

PET imaging metric, SUVmax, as it represents the highest voxel intensity within the segmented region of interest [29,30]. Supplementary Tables S3 and S4 detail all extracted radiomic features.

2.5. AAA Growth Rate Prediction Modelling

A multi-layer perceptron (MLP)-based regression model was developed to predict AAA growth rate based on radiomic features within the AAA sac from the whole AAA and aortic arch blood pool segmentations. Ground truth growth rate values for AAAs were calculated from serial ultrasound measurements.

Ratios of AAA to aortic arch blood pool values for radiomic features were calculated to compensate for blood pool uptake [30]. Normalising against the aortic arch blood pool, was conducted to reduce confounding factors and enhances the specificity of radiomic feature analysis. This approach is analogous to SUV normalisation commonly used in clinical PET to enable more reliable interindividual comparisons. A recursive feature elimination (RFE) algorithm based on Random Forest (RF) regression (available in Scikit-learn (Version 0.24) Python library) was used to select the most discriminative set of radiomic features for models predicting AAA growth rates [31,32]. Table 2 details the radiomic features of highest importance. Subsequently, a MLP was developed for regressing AAA growth rates using full and reduced sets of radiomic features. The MLP was implemented using TensorFlow-Keras (Version 2.3) [33].

Table 2. Features Included in the Reduced Dataset Derived from RF-RFE.

Feature Class	Feature Count	Feature Names
First Order	4	90Percentile, Maximum, Uniformity, Variance
Gray Level Co-Occurrence Matrix	2	DifferenceAverage, Imc1
Gray Level Run Length Matrix	2	ShortRunHighGrayLevelEmphasis, ShortRunLowGrayLevelEmphasis
		GrayLevelNonUniformity, GrayLevelNonUniformityNormalised, GrayLevelVariance, HighGrayLevelZoneEmphasis, LargeAreaEmphasis, LargeAreaHighGrayLevelEmphasis, LargeAreaLowGrayLevelEmphasis, LowGrayLevelZoneEmphasis, SizeZoneNonUniformity, SizeZoneNonUniformityNormalised, SmallAreaEmphasis, SmallAreaHighGrayLevelEmphasis, SmallAreaLowGrayLevelEmphasis, ZoneEntropy, ZonePercentage, ZoneVariance
Gray Level Size Zone Matrix	12	Busyness, Coarseness, Complexity, Contrast, Strength
Neighbouring Gray Tone Difference Matrix	5	

The internal dataset was initially spilt into 70% training/validation and 30% testing. K-fold cross-validation (K = 10) was subsequently used to split the former into 80% for training and 20% for validating the machine learning (ML) model. This enabled assessment of the reproducibility and generalisation capacity of the proposed model. The generated models (from the K-fold cross-validation experiments) were output into a Hierarchical Data Format 5 (HD5) file and evaluated against the test data. Figure 2 illustrates the ML framework.

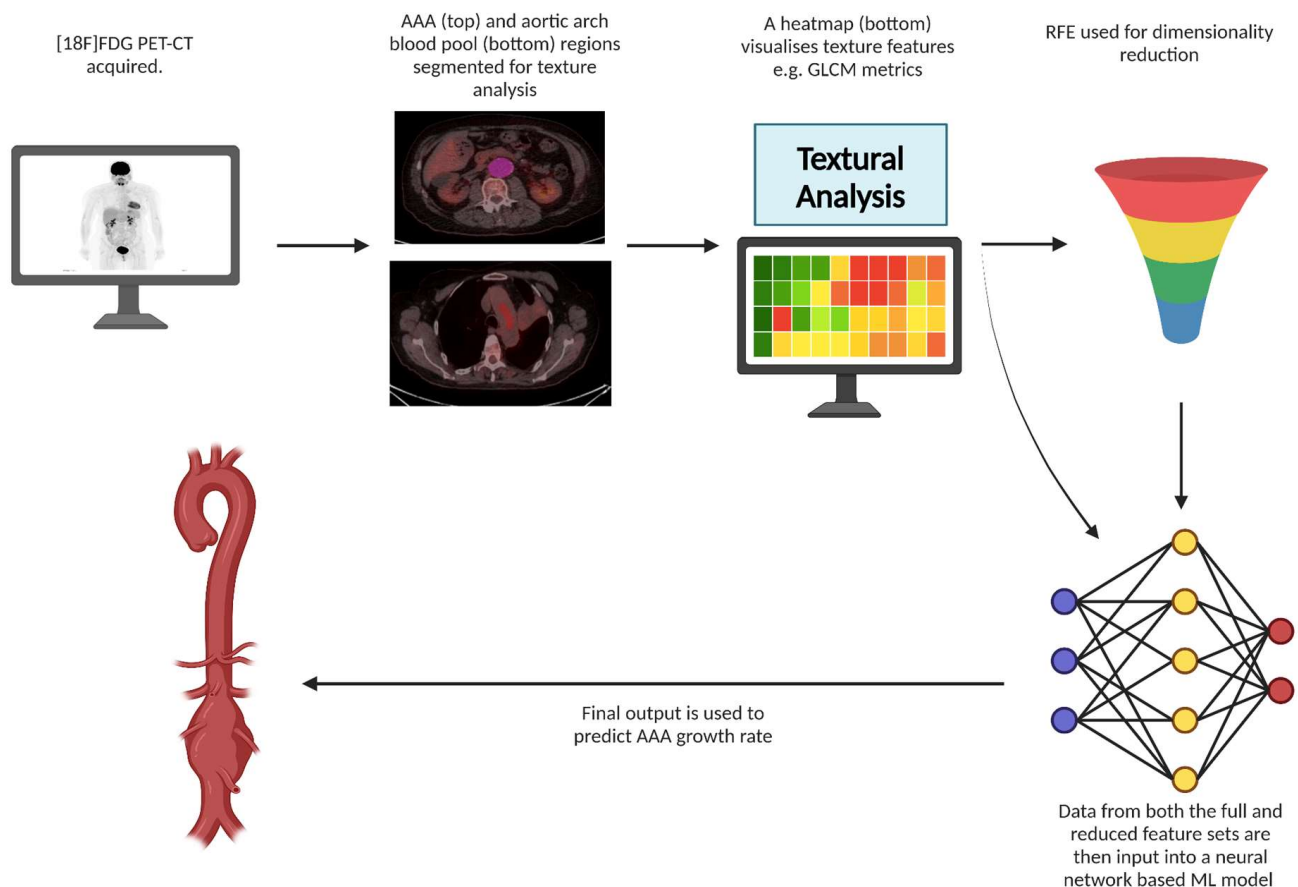


Figure 2. A flow diagram illustrating the machine learning (ML) framework: AAA and aortic arch blood pool regions are segmented for texture analysis, with radiomic features extracted from both. A heatmap (green, yellow and red grid) visualises texture features, such as Grey-Level Co-occurrence Matrix (GLCM) metrics. Dimensionality reduction (top right), as described in Section 2.5, uses Recursive Feature Elimination (RFE) based on Random Forest regression to select 29 features from an initial set of 93 radiomic features (details provided in Table 2). Both full and reduced feature sets are then input into a neural network-based ML model (bottom right) to predict AAA growth rate. The final outputs predict aneurysm growth rate.

2.6. Multi-Layer Perceptron (MLP) Architecture

The MLP architecture consisted of seven fully connected/dense layers, which alternated with dropout layers (i.e., dropout was applied before every weight layer in the MLP), as depicted in Figure 3. Dropout was used to regularise the network and prevent the latter overfitting to the training data [34]. The hidden units in all layers of the network (except the final layer) employed Rectified Linear Units (ReLU) as non-linear activation functions. Weights in all dense layers were initialised using the He normal method (`kernel_initializer = 'he_normal'` in Keras). Bias terms were initialised to small constant values, typically zero, which is standard practice for ReLU-based networks.

The final layer of the MLP comprised a single output unit with a sigmoid activation function, producing a single scalar value for the growth rate of each AAA. The regression-MLP was trained using the standard back-propagation algorithm, by minimising the mean-squared-error loss function (evaluated between the ground truth and predicted AAA growth rates) using the root-mean-squared gradient propagation (RMSprop) optimiser (with an initial learning rate = 0.01) [35]. This optimiser was chosen empirically as it provided the best performance throughout cross-validation. The same network architecture was employed for all experiments.

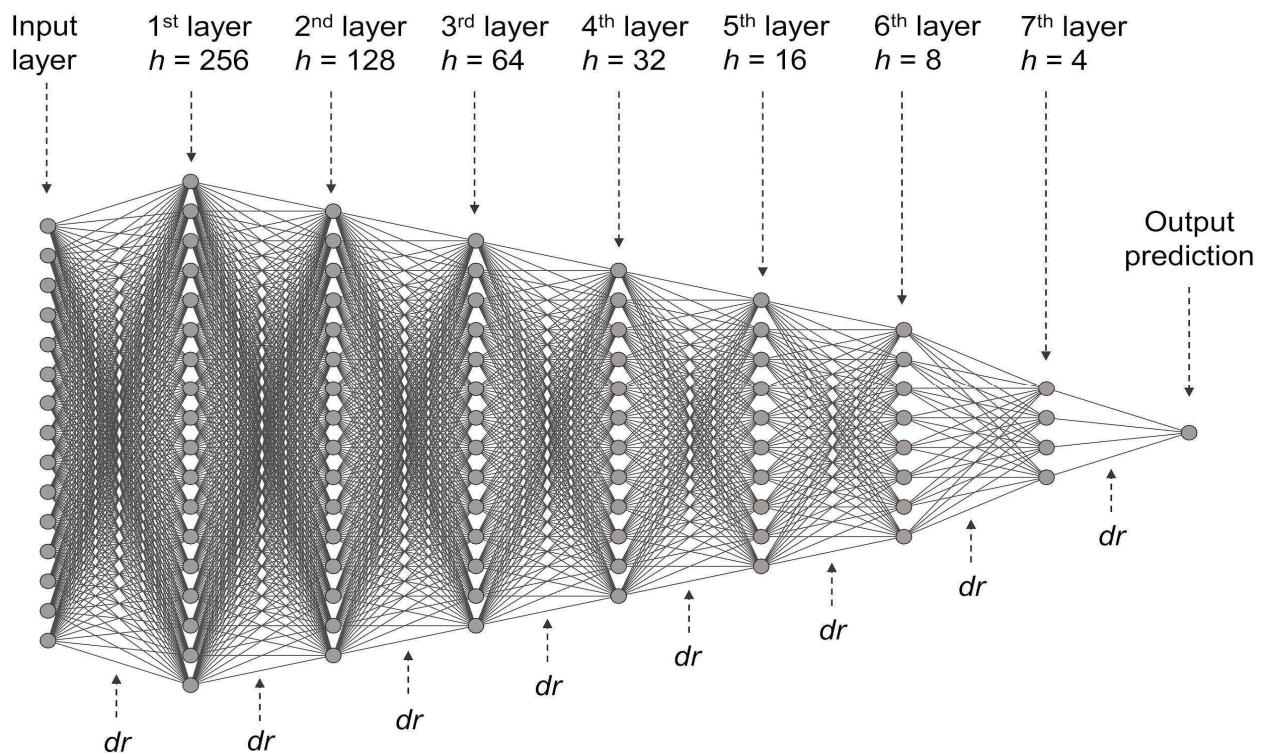


Figure 3. Schematic overview of the MLP used throughout this study. ‘ h ’ represents the number of hidden units in each layer, ‘ dr ’ represents intermediate dropout layers. All ‘ dr ’ applied a rate of 0.5 (50%). The features pass through seven fully connected dense layers of decreasing size: 256, 128, 64, 32, 16, 8, and 1 neuron in the output layer. Hidden layers use ReLU activations and are followed by ‘ dr ’ with a rate of 0.5. Each dense layer incorporates L2 reg (keras.regularisers.l2(0.01)). The final output node applies a sigmoid activation to produce a continuous prediction of AAA growth rate. Weight init is ‘he_normal’ for all dense layers. The model employs RMSprop ($lr = 1e-2$) and Huber loss ($\delta = 1.0$). Performance is evaluated using MAE (from keras.metrics).

2.7. Model Uncertainty Quantification

In addition to evaluating the accuracy of the regression model at predicting AAA growth rates of samples in the held-out test set, the uncertainty in these predictions was quantified using dropout during inference, for all cross-validation experiments [36]. Quantifying model uncertainty provides a framework for identifying special/extreme cases, as models are likely to return predictions with high uncertainties. In a clinical setting, this would enable such special cases to be flagged for further investigation by the interpreting clinician before drawing any conclusions based on the predictions alone. Consequently, dropout during model inference was adopted as a Bayesian approximation of model uncertainty in this study to provide a sense of the model’s confidence when predicting growth rates of AAA samples. Model inference (keeping the dropout layers active) was repeated 1000 times for each unseen test sample, using all models trained across the ten-fold cross-validation experiments.

2.8. Evaluation Metrics

Prediction accuracy in a regression model unlike a classification model demonstrates how close the predicted value is to the ground truth. A perfect model would achieve a prediction error of 0. To quantitatively assess the uncertainty of models trained across ten-fold cross-validation experiments using both the full and reduced feature sets, mean absolute error (MAE) was employed. MAE quantifies the average absolute difference between predicted and observed growth rates of AAAs. The standard deviation (SD) of

MAE across folds was calculated to evaluate prediction consistency, while the standard error of the mean (SEM) was computed to quantify the precision of the average MAE. In addition, violin plots of the errors for the predicted growth rates were generated for each test sample, relative to the ground truth. These plots were created from predictions of all models with dropout during inference. Specifically, these plots summarise the variation in growth rate predictions for each test sample (due to model uncertainty), across a total of 10,000 evaluations (i.e., 1000 independent predictions from 10 models trained across the cross-validation experiments).

2.9. Additional Machine Learning Models

In addition to the MLP model, three other ML models: Random Forest (RF), XGBoost, and Support Vector Machine (SVM), were developed and evaluated for comparative predictive performance. These models are well-established machine learning techniques and have been extensively described in statistical learning literature [37]. These models followed consistent data handling procedures, with identical training and test data splits and the same K-fold cross-validation approach. RFE was applied across all models to select the most relevant radiomic features. Uncertainty quantification was implemented using dropout during inference for the MLP and RF models, while a 95% prediction interval (PI) was calculated for XGBoost. All models were evaluated using mean absolute error (MAE) as the primary metric. Among the tested models, the MLP model demonstrated the highest predictive accuracy, making it the primary focus of this study. Key parameters and results for RF, XGBoost, and SVM are provided in Supplementary Table S5.

2.10. External Validation

To evaluate model performance, an independent external validation dataset comprising 55 patients was used for testing. The dataset, sourced from Addenbrooke's Hospital (Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK) and Edinburgh Royal Infirmary (NHS Lothian, Edinburgh, UK), was used to assess the model's predictive accuracy, clinical utility, and applicability to a more diverse population [38,39].

Patients were recruited into two cohorts: those with asymptomatic AAA and an age and sex-matched control group with atherosclerosis but no aortic aneurysm disease, confirmed by non-contrast CT-derived aortic diameters. All control patients were recruited in Cambridge, while patients with aneurysms were recruited from both sites. Inclusion criteria for the AAA cohort included age over 50 years and aneurysm size between 3.0 and 5.5 cm on ultrasound. Inclusion criteria for the control cohort required clinically stable cardiovascular disease for more than six months, defined as previous myocardial infarction, stroke, or peripheral vascular disease. Exclusion criteria for both cohorts included type 1 or type 2 diabetes with fasting glucose above 11 mmol/L, renal impairment (serum creatinine > 250 µmol/L), contrast allergy, or inability to provide informed consent.

The external dataset used standardised image acquisition protocols (as per local Trust policies) and underwent identical processing, spatial resampling, VOI segmentation, and radiomic feature extraction (aligned to IBSI guidelines) as performed for the internal dataset. AAA growth rates were derived from serial ultrasound measurements, enabling direct comparison with internal dataset ground truths. Full details of image acquisition protocols can be provided upon request.

2.11. Hardware Specifications

All image processing and network training was performed on a personal computer (PC) running Microsoft Windows 10 (Microsoft Corporation, Redmond, WA, USA), AMD Ryzen 7 1700 central processing unit (CPU) (Advanced Micro Devices Inc., Santa Clara, CA, USA), 16 GB of 3333 MHz DDR4 SDRAM and a Zotac NVIDIA GeForce GTX 1070

AMP Edition graphics processing unit (GPU) with 8 GB of GDDR5 (Zotac International, Hong Kong) (NVIDIA Corporation, Santa Clara, CA, USA).

2.12. Statistical Analysis

Data preprocessing, descriptive statistics, and model evaluations were carried out using NumPy 1.18.0, Pandas 0.25.3, and SciPy 1.4.0. Univariate comparisons were performed using built-in functions in SciPy and statsmodels 0.10.2. Machine learning models, including Random Forest, Support Vector Machine, and XGBoost, as well as Recursive Feature Elimination (RFE) for feature selection, were implemented using scikit-learn 0.22 and XGBoost 0.90. Neural network modelling, including the multi-layer perceptron (MLP), was conducted using the Keras 2.3.1 API with a TensorFlow 2.1.0 backend. Model performance was evaluated using mean absolute error (MAE) and, where applicable, 95% prediction intervals (PI). Figures were created using Matplotlib 3.1.2 and Plotly 4.4.1. A significance level of $p < 0.05$ was applied for all hypothesis testing.

3. Results

The cohort comprised 153 patients, including 98 and 55 individuals in the internal and external datasets, respectively. Overall, 82.4% (126/153) were male, and 17.6% (27/153) were female, with a median age of 77 years (range: 58–105). For the internal dataset, median follow-up was 3.1 years (range: 0.48–14.8), with a median of 7 ultrasound scans per patient (range: 2–36). Among these patients, 3.3% (5/153) underwent open AAA repair, 0.7% (1/153) had EVAR, and 96.1% (147/153) had no intervention; no AAA ruptures occurred. Of the external dataset, only data relating to open AAA repair was available, reported in 21% (12/55) of cases. Additional details, including demographics and scanner usage, are listed in Table 3a,b.

Table 3. (a): Patient Demographics; (b) Scanner Breakdown Table (Internal Data Only).

(a)			
	Total	Internal	External
No. of Patients	153	98	55
Male/female	126/27	77/21	49/6
Median Age, years	77 (range: 58–105)	76 (range: 61–92)	80 (range: 58–105)
Median follow-up time, years	N/A ¹	3 years (range: 0–15)	N/A ¹
Median number of follow-up ultrasound scans	N/A ¹	7 (range: 2–36)	N/A ¹
Progression to open repair	17	5	12
Progression to EVAR	N/A ¹	1	N/A ¹
Stable—no intervention	N/A ¹	147	N/A ¹
(b)			
Scanner Model	Count		
GE Healthcare Discovery 710	54		
GE Healthcare Discovery 690	13		
GE Healthcare Discovery STE	9		
Philips Gemini TF64	22		

¹ N/A: not available.

The training dataset included 69 patients from the internal dataset, while the testing datasets comprised 29 internal and 55 external patients. Gender distributions were consistent across groups, showing no significant differences. However, there was a difference in age between the external testing dataset and the training dataset. No significant differences were found between the training and internal testing datasets for follow-up characteristics. Table 4 outlines the demographic differences between training and testing datasets.

Table 4. Training and Testing Demographics: Z-Scores and *p*-Values.

	Training and Validation	Internal Testing	External Testing	Statistical Test	Internal			External		
					Z-Score	<i>p</i> -Value	Significance	Z-Score	<i>p</i> -Value	Significance
Total Patients	69	29	55	Z-Test for Two Proportions Mann–Whitney U-Test Mann–Whitney U-Test Mann–Whitney U-Test	1.5024	0.1336	Not significant at $p < 0.05$	−1.57	0.1163	Not significant at $p < 0.05$
Male/Female	57/12	20/9	49/6							
Median Age (Years)	75.7 (60.0–91.8)	76.3 (60.7–90.7)	80.0 (58–105)		−0.1323	0.8966	Not significant at $p < 0.05$	−4.55	5.34×10^{-6}	Significant at $p > 0.05$ ²
Median Follow-Up Time (Years)	3.2 (0.57–12.6)	3.5 (0.48–14.8)	N/A ¹		−0.0642	0.9522	Not significant at $p < 0.05$	N/A ¹	N/A ¹	N/A ¹
Median Number of Ultrasound Scans	7 (2–30)	6 (2–36)	N/A ¹		1.5683	0.1164	Not significant at $p < 0.05$	N/A ¹	N/A ¹	N/A ¹

¹ N/A: not available; ² significant difference as indicated by $p > 0.05$

A total of 93 radiomic features were extracted per patient using PyRadiomics, forming the dataset used for model training and validation. The regression-MLP models trained across ten-fold cross-validation experiments were evaluated on unseen internal test samples. The statistical analysis, which assessed prediction accuracy, included calculations of the mean absolute error (MAE) and its standard error of the mean (SEM), with the model achieving an MAE of $1.35 \pm 3.2e-4$ mm/year. The average training and validation curves for loss and accuracy, with prediction error standard deviations (computed across all experiment folds) are shown in Figure 4a.

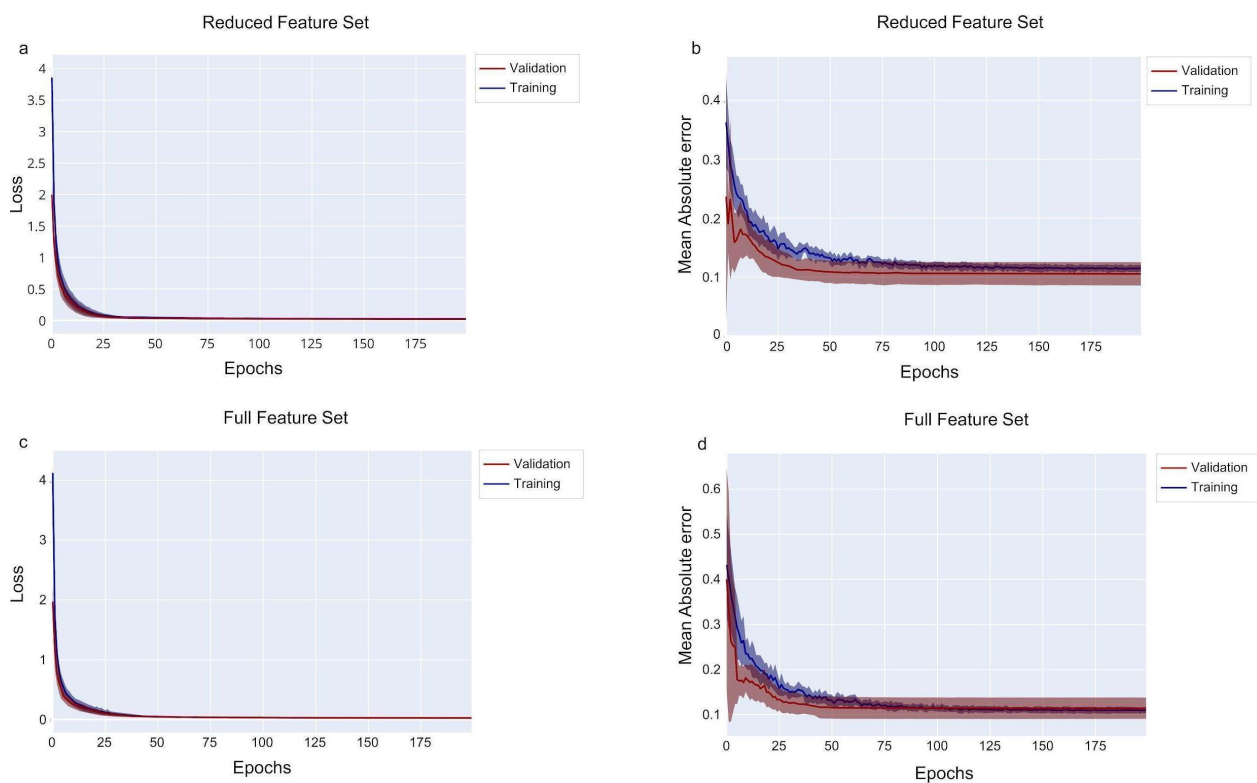


Figure 4. (a) Average training and validation curves for loss and accuracy for the reduced feature set, (b) prediction error standard deviations for the reduced feature set, (c) average training and validation curves for loss and accuracy for the full feature set, and (d) prediction error standard deviations for the full feature set.

To refine the models, features with least significance were eliminated using RF-based RFE which yielded 29 features. As before, regression-MLP models were trained across

ten-fold cross-validation experiments using these reduced sets of features. The models were evaluated on the same set of unseen internal test samples as previously, resulting in a near identical average accuracy of $1.35 \pm 2.5\text{e-}4$ mm/year. The average training and validation curves for loss and accuracy, with prediction error standard deviations (computed across all experiment folds) are illustrated in Figure 4c.

The MLP model was subsequently evaluated on unseen external test samples, achieving an accuracy of $1.8 \pm 8.9\text{e-}8$ mm/year. Of the other models tested, internal full feature, internal RFE, and external validation accuracies were, respectively, as follows: for Random Forest (RF), $1.5 \pm 3.8\text{e-}4$ mm/year (uncertainty: ± 0.05 mm/year), $1.48 \pm 3.7\text{e-}4$ mm/year (uncertainty: ± 0.04 mm/year), and $1.95 \pm 6.9\text{e-}4$ mm/year (uncertainty: ± 0.06 mm/year); for eXtreme Gradient Boosting, $1.42 \pm 4.2\text{e-}4$ mm/year (95% PI: [1.38, 1.46] mm/year), $1.40 \pm 4.0\text{e-}4$ mm/year (95% PI: [1.36, 1.44] mm/year), and $1.97 \pm 7.1\text{e-}4$ mm/year (95% PI: [1.93, 2.01] mm/year); and for Support Vector Machine (SVM), $1.45 \pm 4.1\text{e-}4$ mm/year, $1.40 \pm 3.9\text{e-}4$ mm/year, and $1.90 \pm 7.2\text{e-}4$ mm/year.

Overall, the results from the study demonstrate that the model was able to predict the growth rate of a AAA to an accuracy of between 1.35 and 1.8 mm/year, with minimal variation in prediction values across internal and external datasets.

Using dropout during MLP model inference, uncertainty was quantified and its impact on the variation in growth rate prediction accuracy, for each unseen internal test sample is summarised by the violin plots shown in Figure 5 for models trained using the full and reduced feature sets. Model uncertainty was generally higher when trained using the reduced feature set (across all cross-validation experiments), relative to the full feature set. This is verified by the higher median values and interquartile ranges for the MAE, visible for all 29 internal and 55 external test samples when the full set of features were used to train the models. Additionally, the kernel density estimates for the MAE for each test sample show that predictive growth rate errors are most likely to occur in the range of 1.30–1.40 mm/year when the full feature set is used, and 1.40–1.50 mm/year when the reduced feature set is used.

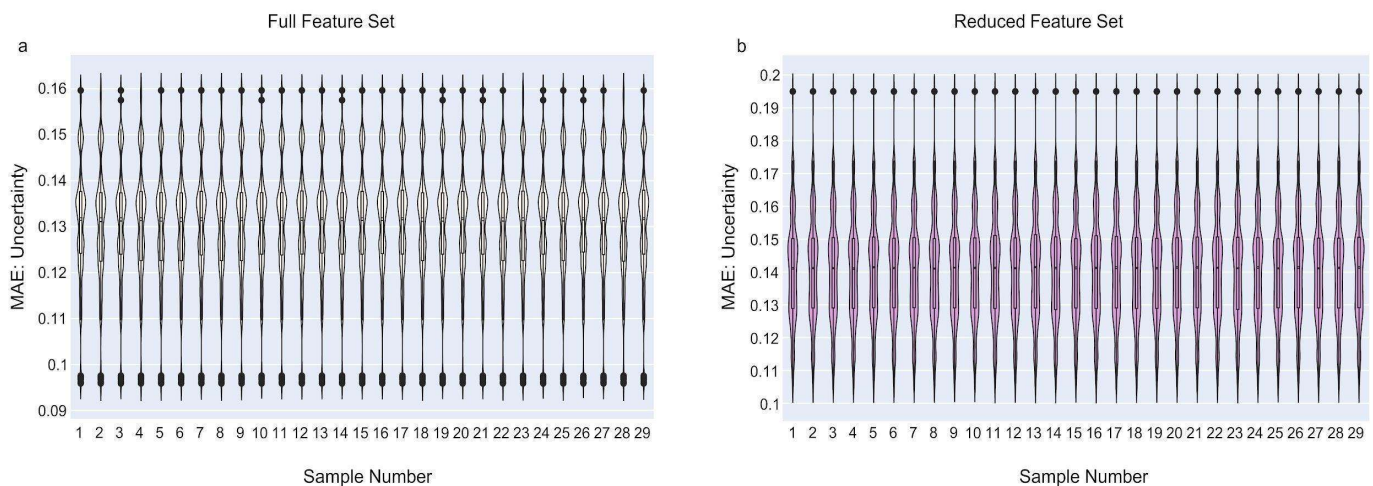


Figure 5. Violin plots demonstrating uncertainty on the model trained with full feature set (a) and reduced feature set (b). Model uncertainty was generally higher when trained using the reduced feature set relative to the full feature set. Note: MAE: mean absolute error.

4. Discussion

This study tested several machine learning models using radiomic features derived from [18F]FDG PET-CT to predict the growth rate of abdominal aortic aneurysms. Overall, predictive accuracy was highest for a regression-MLP model. To the best of our knowledge,

there are no previous studies to have investigated the use of [18F]FDG PET-CT to predict the growth rate of abdominal aortic aneurysms.

Conventional PET-CT metrics, such as SUVmax, MTV, and TLG, are well-established in oncological imaging for quantifying metabolic activity. SUVmax is frequently used to assess tumour aggressiveness and response to therapy in a variety of cancers [40–42]. Similarly, MTV and TLG have been shown to predict treatment outcomes in an oncological setting, reflecting tumour burden and metabolic activity, respectively [43]. However, their application in vascular disease, including AAA prediction, remains limited and under-explored. Moreover, studies investigating SUV metrics for predicting AAA growth and rupture have reported conflicting results, likely due to variability in imaging protocols and methodologies, which complicates their clinical utility in this context [12–17]. By contrast, there is growing evidence supporting the predictive capability of textural radiomic parameters in various cardiovascular disorders, offering nuanced insights into tissue heterogeneity that conventional metrics may overlook [44,45]. Radiomic features may therefore offer an advanced alternative by providing a detailed analysis of texture and structural patterns, capturing subtle imaging biomarkers that reflect underlying pathophysiology. This enables a deeper understanding of aneurysm behaviour and improves predictive accuracy compared to traditional metrics such as SUV.

The regression-MLP model was trained to minimise the mean squared error loss function, evaluated between predicted and ground truth aneurysm growth rates whilst avoiding overfitting of the training data. Each fully connected layer of the MLP is preceded by a dropout layer, which acts as a regulariser and prevents the model from overfitting to the training data. The models trained with the full set of [18F]FDG PET-CT features extracted and the reduced set identified using RF-RFE, both resulted in very similar average growth rate prediction accuracy of 1.35 mm/year on the internal dataset. Dimensionality reduction with RF-RFE showed no tangible improvement in the growth rate prediction accuracy, although the uncertainty quantification experiments revealed that using a reduced set of features increased model uncertainty (relative to using the full feature set). Based on these results, it can be inferred that models trained using the full set of features are more likely to predict AAA growth rates to a higher degree of confidence, than those trained using the reduced feature set.

When evaluated on an unseen external dataset, the model achieved an accuracy of 1.8 mm/year, a result that falls within the 2 mm margin of accuracy typically considered acceptable for technical variability [46,47]. Consistency across internal and external datasets suggests the model's potential for real-world application. Given that more rapid AAA growth is associated with a higher rupture rate this model could provide a useful risk stratification metric [8,9,48].

The RESCAN study, a meta-analysis of surveillance intervals for small AAAs, reported an average annual growth rate of 1–2 mm per year [49]. While the reported error margin in our study is comparable to this observed yearly growth rate, potentially limiting its utility for precise short-term predictions, the model remains promising for longer-term predictions, such as over 5 or 10 years. Cumulative predictions over such extended periods could enable stratified monitoring strategies, focusing resources on patients at higher risk of rapid growth and optimising clinical resource allocation.

Previous studies have explored automated segmentation methods to model AAA growth using morphometric characteristics, biomarkers and computational fluid dynamics (CFD) in conjunction with ML methods [7–9]. Hirata et al. described the use of a ML model to predict AAA growth using aneurysm area and major axis diameter derived from serial computed tomography angiography (CTA) studies [50]. Lee et al. developed a ML model able to predict AAA diameter to within 2 mm in 85% and 71% of patients at 12 and

24 months, respectively. They used ultrasound to explore flow-mediated dilatation (FMD) of the brachial artery as a biomarker in combination with AAA diameter [51]. Rengarajan et al. investigated a ML algorithm encompassing auto-segmentation of AAA on CTA and prediction of rupture risk (9). This used a generalised additive model (GAM) with six geometric and one biomechanical marker yielding an accuracy of 87%, 78% sensitivity and 92% specificity for AAA rupture risk classification [52]. More recently, Chandrashekar et al. demonstrated that a linear regression model derived from geometric features, automatically calculated from deep learning created segmentations, which could predict AAA growth to within 2 mm in 87% of cases on CT [53]. However, due to the novel nature of the current model, no meaningful direct comparison was able to be made with prior work.

Work by Lv et al. [54] investigated the use of radiomic features from perivascular adipose tissue (PVAT) to enhance predictive performance for AAA growth. Their study demonstrated that growing AAAs post-endovascular aneurysm repair (EVAR) exhibited significantly higher surface area-to-volume ratios (0.70 vs. 0.63, $p = 0.04$) and more heterogeneous PVAT texture, with differences in dependence variance (29.88 vs. 26.96, $p = 0.01$) and long-run emphasis (4.76 vs. 4.20, $p = 0.01$). Reported findings mirror previous research in coronary artery disease, where PVAT radiomic features have been shown to improve risk assessment for atherosclerotic progression. In CAD, heterogeneous PVAT texture has been associated with increased inflammatory activity and vascular remodelling, leading to the development of novel risk stratification methods [55,56]. By analogy, for incidentally detected AAAs, as examined in this cohort, PVAT may reflect similar inflammatory or mechanical changes in the vascular environment that are not captured by aneurysm sac analysis alone. Integrating texture features from surrounding tissues like PVAT could therefore provide a more comprehensive understanding of aneurysm behaviour and progression. Furthermore, texture analysis, particularly metrics derived from grey-level co-occurrence matrix (GLCM) and grey-level dependence matrix (GLDM), has also been increasingly recognised for its role in predicting aneurysm behaviour [54,57]. For instance, studies have shown that texture heterogeneity in thrombi and PVAT can reflect inflammatory activity and vascular remodelling, which may contribute to aneurysm progression. Integrating texture features from surrounding tissues like PVAT could therefore provide a more comprehensive understanding of aneurysm behaviour beyond traditional morphometric measures.

The study by Lv et al. further highlighted the advantages of combining radiomic features with clinical data in a multi-modal predictive model, which outperformed models using radiomic or clinical data alone. Their clinico-radiological model achieved an area under the curve (AUC) of 0.78 (95% CI: 0.65–0.91), significantly higher than models using radiomic or clinical features alone (AUC = 0.69 for each) [58]. These findings may suggest a merit to incorporating clinical parameters, such as cholesterol levels, blood pressure, and patient history, to enhance the predictive accuracy of future models.

Whilst these preliminary results are promising, there are limitations to the study. The ground truth was based on serial ultrasound measurements, which despite being the standard-of-care investigation used for screening and follow-up for AAAs, are operator dependent and unidimensional. [18F]FDG PET-CT is not routinely used to assess AAAs unless there is concern for an infective or inflammatory process. The cohort in this study was derived from patients with an AAA who were undergoing diagnostic evaluation for suspected malignancy which may not reflect the general population with AAAs. Although cases with tumours overlapping the aorta were excluded, the potential for systemic malignancy to subtly influence aneurysm wall tracer uptake cannot be entirely ruled out. A future study could incorporate patient demographics, morphometric features, automatic data-driven feature extraction from volumetric data and biomechanical finite element analysis to enhance growth rate prediction accuracy and facilitate evaluation of risk of

aneurysm rupture. The sample size of 153 and retrospective nature limits the power of the study. For instance, accuracy would typically increase to a certain degree, with the training dataset size, while external validation with a larger multi-centre dataset would improve confidence in the prediction accuracy of the proposed approach, on unseen test cases. In our study, the feature-to-sample ratio after dimensionality reduction resulted in approximately 5.3 samples per feature. While this may be seen as a limitation under the traditional ‘One in Ten Rule’ for mitigating overfitting, more recent literature suggests that this rule may be overly conservative for advanced machine learning models [59]. The application of ten-fold cross-validation and external validation in our study likely reduced the risk of overfitting, but future studies with larger datasets are needed to fully address this issue. Data augmentation was not employed in this study due to its retrospective nature, which limited the feasibility of introducing synthetic or augmented data. The focus was on analysing radiomic features directly extracted from existing PET-CT datasets. While this approach ensured the integrity of the dataset, the absence of data augmentation may have limited the model’s ability to generalise further. Future prospective studies incorporating data augmentation techniques could help to mitigate overfitting and improve model performance. Furthermore, the absence of harmonisation techniques to address potential variability between PET/CT scanners was not performed in this study. While consistent imaging protocols were used, inter-scanner variability cannot be entirely ruled out. Future studies incorporating harmonisation strategies, such as phantom calibration or scanner-specific corrections, may enhance reproducibility and standardisation. Finally, manual segmentation of the AAA and aortic blood pool is time consuming and not suitable for routine integration into clinical reporting workflow. An automatic segmentation algorithm is under development, and this should aid further validation and potential clinical translation.

5. Conclusions

In this initial exploratory study, an MLP-based regression model utilising [18F]FDG PET-CT-derived radiomic features demonstrated the ability to predict AAA growth rate with an accuracy of 1.35 mm/year. These promising initial results warrant validation in a larger prospective multi-centre cohort. In the future, more sophisticated stratification could guide individualised patient care facilitating tailored management of AAA.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/a18020086/s1>, Table S1: Tabulated TRIPOD Checklist; Table S2: PyRadiomic Customised Settings; Table S3: Radiomic Feature Classes; Table S4: All Radiomic Features; Table S5: Summary of Machine Learning Models and Performance.

Author Contributions: Conceptualisation, P.A., N.R., M.A.W., R.F., A.F.F., J.H.F.R., M.A.B. and A.F.S.; methodology, P.A., N.R., M.A.W., R.F., A.F.F., J.H.F.R., M.A.B. and A.F.S.; software, N.J., G.M., J.Y.K., Y.H. and P.C.; validation, P.A., S.S.D., N.R., N.J., G.M., J.Y.K., Y.H. and P.C.; formal analysis, P.A., S.S.D., N.J., N.R., G.M., J.Y.K., Y.H. and P.C.; investigation, S.S.D., P.A., N.R., M.A.W., R.F., A.F.F., M.A.B. and A.F.S.; resources, A.F.S.; data curation, P.A., S.S.D., N.J., N.R., G.M., J.Y.K., Y.H. and P.C.; writing—original draft preparation, S.S.D., P.A., N.R., M.A.W., R.F., A.F.F., M.A.B. and A.F.S.; writing—review and editing, all authors; visualisation, P.A., S.S.D., N.J., N.R., G.M., Y.H. and P.C.; supervision, A.F.S.; project administration, A.F.S. All authors have read and agreed to the published version of the manuscript.

Funding: M.W. and M.B. were supported by personal fellowships from the British Heart Foundation during the time of the work. J.H.F.R. is part-supported by the NIHR Cambridge Biomedical Research Centre (BRC), the British Heart Foundation, HEFCE, the EPSRC, and the Wellcome Trust. A.S. is part-supported by Cancer Research UK (C19942/A28832) and NIHR Leeds BRC (NIHR203331). DEN is supported by the British Heart Foundation (CH/09/002, RG/F/22/110093, RE/24/130012). The

APC was funded by the authors personally. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR Leeds and Cambridge BRCs, or the Department of Health and Social Care. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki. Prospective informed written consent was obtained from all patients at the time of imaging for use of their anonymised [18F]FDG PET-CT imaging data in research and service development projects. Formal ethics committee approval was waived for this study, as it was considered by the institutional review board to represent an evaluation of a routine clinical service.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to institutional data-sharing restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AAA	Abdominal Aortic Aneurysm
APC	Article Processing Charge
BRC	Biomedical Research Centre
CAD	Coronary Artery Disease
CFD	Computational Fluid Dynamics
CT	Computed Tomography
EVAR	Endovascular Aneurysm Repair
FDG	Fluorine-18-2-Deoxy-D-Glucose
FMD	Flow-Mediated Dilation
GAM	Generalised Additive Model
GLCM	Gray-Level Co-Occurrence Matrix
GLDM	Gray-Level Dependence Matrix
IBSI	Imaging Biomarker Standardization Initiative
MAE	Mean Absolute Error
MLP	Multi-Layer Perceptron
ML	Machine Learning
PET-CT	Positron Emission Tomography–Computed Tomography
PI	Prediction Interval
PVAT	Perivascular Adipose Tissue
RFE	Recursive Feature Elimination
RF	Random Forest
SD	Standard Deviation
SVM	Support Vector Machine
TRIPOD	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
UK	United Kingdom
US	Ultrasound
VOI	Volume of Interest

References

1. Sakalihasan, N.; Michel, J.B.; Katsargyris, A.; Kuivaniemi, H.; Defraigne, J.O.; Nchimi, A.; Hultgren, R. Abdominal aortic aneurysms. *Nat. Rev. Dis. Primers* **2018**, *4*, 34. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Li, X.; Zhao, G.; Zhang, J.; Duan, Z.; Xin, S. Prevalence and Trends of the Abdominal Aortic Aneurysms Epidemic in General Population—A Meta-Analysis. *PLoS ONE* **2013**, *8*, e81260. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Larsson, E.; Granath, F.; Swedenborg, J.; Hultgren, R. A population-based case-control study of the familial risk of abdominal aortic aneurysm. *J. Vasc. Surg.* **2009**, *49*, 47–50. [\[CrossRef\]](#) [\[PubMed\]](#)

4. Blanchard, J.F.; Armenian, H.K.; Friesen, P.P. Risk factors for abdominal aortic aneurysm: Results of a case-control study. *Am. J. Epidemiol.* **2000**, *151*, 575–583. [CrossRef]
5. Kent, K.C.; Zwolak, R.M.; Egorova, N.N.; Riles, T.S.; Manganaro, A.; Moskowitz, A.J.; Gelijns, A.C.; Greco, G. Analysis of risk factors for abdominal aortic aneurysm in a cohort of more than 3 million individuals. *J. Vasc. Surg.* **2010**, *52*, 539–548. [CrossRef]
6. Shum, J.; Xu, A.; Chatnuntawe, I.; Finol, E. A framework for the automatic generation of surface topologies for abdominal aortic aneurysm models. *Ann. Biomed. Eng.* **2010**, *39*, 249–259. [CrossRef]
7. Thompson, S.; Brown, L.; Sweeting, M.; Bown, M.; Kim, L.; Glover, M.; Buxton, M.; Powell, J. Systematic review and meta-analysis of the growth and rupture rates of small abdominal aortic aneurysms: Implications for surveillance intervals and their cost-effectiveness. *Health Technol. Assess* **2013**, *17*, 1–118. [CrossRef]
8. Brown, P.M.; Zelt, D.T.; Sobolev, B. The risk of rupture in untreated aneurysms: The impact of size, gender, and expansion rate. *J. Vasc. Surg.* **2003**, *37*, 280–284. [CrossRef]
9. Aggarwal, S.; Qamar, A.; Sharma, V.; Sharma, A. Abdominal aortic aneurysm: A comprehensive review. *Exp. Clin. Cardiol.* **2011**, *16*, 11–15.
10. GOVUK. AAA Screening: Professional Guidance. GOVUK. 2020. Available online: <https://www.gov.uk/government/collections/aaa-screening-supporting-documents> (accessed on 27 November 2024).
11. US Preventive Services Task Force; Owens, D.; Davidson, K.; Krist, A. Screening for Abdominal Aortic Aneurysm: US Preventive Services Task Force Recommendation Statement. *JAMA* **2019**, *322*, 2211. [CrossRef]
12. Lee, H.; Paeng, J.C.; Kim, K.H.; Cheon, G.J.; Lee, D.S.; Chung, J.-K.; Kang, K.W. Correlation of [18F]FDG PET/CT Findings with Long-Term Growth and Clinical Course of Abdominal Aortic Aneurysm. *Nucl. Med. Mol. Imaging* **2018**, *52*, 46–52. [CrossRef] [PubMed]
13. Reeps, C.; Essler, M.; Pelisek, J.; Seidl, S.; Eckstein, H.-H.; Krause, B.-J. Increased 18F-fluorodeoxyglucose uptake in abdominal aortic aneurysms in positron emission/computed tomography is associated with inflammation, aortic wall instability, and acute symptoms. *J. Vasc. Surg.* **2008**, *48*, 417–423. [CrossRef] [PubMed]
14. Akerele, M.I.; Karakatsanis, N.A.; Forsythe, R.O.; Dweck, M.R.; Syed, M.; Aykroyd, R.G.; Tsoumpas, C. Iterative reconstruction incorporating background correction improves quantification of [18F]-NaF PET/CT images of patients with abdominal aortic aneurysm. *J. Nucl. Cardiol.* **2019**, *28*, 1875–1886. [CrossRef] [PubMed]
15. Kotze, C.W.; Groves, A.M.; Menezes, L.J.; Harvey, R.; Endozo, R.; Kayani, I.A.; Yusuf, S.W. What is the relationship between 18F-[18F]FDG aortic aneurysm uptake on PET/CT and future growth rate? *Eur. J. Nucl. Med. Mol. Imaging* **2011**, *38*, 1493–1499. [CrossRef]
16. Kotze, C.W.; Menezes, L.J.; Endozo, R.; Groves, A.M.; Ell, P.J.; Yusuf, S.W. Increased metabolic activity in abdominal aortic aneurysm detected by 18F-fluorodeoxyglucose (18F-[18F]FDG) positron emission tomography/computed tomography (PET/CT). *Eur. J. Vasc. Endovasc. Surg.* **2008**, *38*, 93–99. [CrossRef]
17. Gandhi, R.; Bell, M.; Bailey, M.; Tsoumpas, C. Prospect of positron emission tomography for abdominal aortic aneurysm risk stratification. *J. Nucl. Cardiol.* **2021**, *28*, 2272–2282. [CrossRef]
18. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [CrossRef]
19. Shi, Z.; Traverso, A.; Soest, J.; Dekker, A.; Wee, L. Technical Note: Ontology-guided radiomics analysis workflow (O-RAW). *Med. Phys.* **2019**, *46*, 5677–5684. [CrossRef]
20. Fornacon-Wood, I.; Faivre-Finn, C.; O'Connor, J.P.B.; Price, G.J. Radiomics as a personalized medicine tool in lung cancer: Separating the hope from the hype. *Lung Cancer* **2020**, *46*, 197–208. [CrossRef]
21. Peng, H.; Dong, D.; Fang, M.-J.; Li, L.; Tang, L.-L.; Chen, L.; Li, W.-F.; Mao, Y.-P.; Fan, W.; Liu, L.-Z.; et al. Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma. *Clin. Cancer Res.* **2019**, *25*, 4271–4279. [CrossRef]
22. Carvalho, S.; Leijenaar, R.T.H.; Troost, E.G.C.; van Timmeren, J.E.; Oberije, C.; van Elmpt, W.; de Geus-Oei, L.-F.; Bussink, J.; Lambin, P. 18F-fluorodeoxyglucose positron-emission tomography ([18F]FDG-PET)-Radiomics of metastatic lymph nodes and primary tumor in non-small cell lung cancer (NSCLC)—A prospective externally validated study. *PLoS ONE* **2018**, *13*, e0192859. [CrossRef] [PubMed]
23. Nougaret, S.; Tibermacine, H.; Tardieu, M.; Sala, E. Radiomics: An Introductory Guide to What It May Foretell. *Curr. Oncol. Rep.* **2019**, *21*, 70. [CrossRef] [PubMed]
24. Lee, J.W.; Lee, S.M. Radiomics in Oncological PET/CT: Clinical Applications. *Nucl. Med. Mol. Imaging* **2018**, *52*, 170–189. [CrossRef] [PubMed]
25. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G.M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ* **2015**, *350*, g7594. [CrossRef]

26. Leijenaar, R.T.; Nalbantov, G.; Carvalho, S.; Van Elmpt, W.J.; Troost, E.G.; Boellaard, R.; Lambin, P. The effect of SUV discretization in quantitative [18F]FDG-PET Radiomics: The need for standardized methodology in tumor texture analysis. *Sci. Rep.* **2015**, *5*, 11075. [CrossRef]
27. Leijenaar, R.T.; Carvalho, S.; Velazquez, E.R.; Van Elmpt, W.J.; Parmar, C.; Hoekstra, O.S.; Lambin, P. Stability of [18F]FDG-PET Radiomics features: An integrated analysis of test-retest and inter-observer variability. *Acta Oncol.* **2013**, *52*, 1391–1397. [CrossRef]
28. Pyradiomics. Frequently Asked Questions—Pyradiomics v3.0.post5+gf06ac1d Documentation. *Pyradiomics*. 2020. Available online: <https://pyradiomics.readthedocs.io/en/latest/faq.html> (accessed on 27 November 2024).
29. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **2020**, *295*, 328–338. [CrossRef]
30. Saleem, B.R.; Beukinga, R.J.; Boellaard, R.; Glaudemans, A.W.J.M.; Reijnen, M.M.P.J.; Zeebregts, C.J.; Slart, R.H.J.A. Textural features of 18F-fluorodeoxyglucose positron emission tomography scanning in diagnosing aortic prosthetic graft infection. *Eur. J. Nucl. Med. Mol. Imaging* **2017**, *44*, 886–894. [CrossRef]
31. Github. GitHub—Scikit-Learn/Scikit-Learn: Scikit-Learn: Machine Learning in Python. 2020. Available online: <https://github.com/scikit-learn/scikit-learn> (accessed on 27 November 2024).
32. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
33. Keras. Keras: The Python Deep Learning API. 2020. Available online: <https://keras.io/> (accessed on 27 November 2024).
34. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1958.
35. Team, K. Keras Documentation: RMSprop. 2020. Available online: <https://keras.io/api/optimizers/rmsprop/> (accessed on 27 November 2024).
36. Gal, Y.; Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *Int. Conf. Mach. Learn.* **2016**, *48*, 1050–1059.
37. James, G.; Witten, D.; Hastie, T.; Tibshirani, R.; James, G.; Witten, D.; Hastie, T.; Tibshirani, R. Linear model selection and regularization. *Introd. Stat. Learn. Appl. R* **2021**, *2021*, 225–288.
38. Collins, G.S.; Dhiman, P.; Ma, J.; Schluskel, M.M.; Archer, L.; Van Calster, B.; Harrell, F.E.; Martin, G.P.; Moons, K.G.; Van Smeden, M.; et al. Evaluation of clinical prediction models (part 1): From development to external validation. *BMJ* **2024**, *8*, 384. [CrossRef] [PubMed]
39. Riley, R.D.; Archer, L.; Snell, K.I.; Ensor, J.; Dhiman, P.; Martin, G.P.; Bonnett, L.J.; Collins, G.S. Evaluation of clinical prediction models (part 2): How to undertake an external validation study. *BMJ* **2024**, *15*, 384. [CrossRef]
40. Higashi, K.; Ueda, Y.; Ayabe, K.; Sakurai, A.; Seki, H.; Nambu, Y.; Oguchi, M.; Shikata, H.; Taki, S.; Tonami, H.; et al. FDG PET in the evaluation of the aggressiveness of pulmonary adenocarcinoma: Correlation with histopathological features. *Nucl. Med. Commun.* **2000**, *21*, 707–714. [CrossRef]
41. Fathinul, F.; Nordin, A.J.; Lau, W.F. 18 [F] FDG-PET/CT is a useful molecular marker in evaluating tumour aggressiveness: A revised understanding of an in-vivo FDG-PET imaging that alludes the alteration of cancer biology. *Cell Biochem. Biophys.* **2013**, *66*, 37–43. [CrossRef]
42. Kitagawa, Y.; Sano, K.; Nishizawa, S.; Nakamura, M.; Ogasawara, T.; Sadato, N.; Yonekura, Y. FDG-PET for prediction of tumour aggressiveness and response to intra-arterial chemotherapy and radiotherapy in head and neck cancer. *Eur. J. Nucl. Med. Mol. Imaging* **2003**, *30*, 63–71. [CrossRef]
43. Dholakia, A.S.; Chaudhry, M.; Leal, J.P.; Chang, D.T.; Raman, S.P.; Hacker-Prietz, A.; Su, Z.; Pai, J.; Oteiza, K.E.; Griffith, M.E.; et al. Baseline metabolic tumor volume and total lesion glycolysis are associated with survival outcomes in patients with locally advanced pancreatic cancer receiving stereotactic body radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2014**, *89*, 539–546. [CrossRef]
44. Badesha, A.S.; Frood, R.; Bailey, M.A.; Coughlin, P.M.; Scarsbrook, A.F. A Scoping Review of Machine-Learning Derived Radiomic Analysis of CT and PET Imaging to Investigate Atherosclerotic Cardiovascular Disease. *Tomography* **2024**, *10*, 1455–1487. [CrossRef]
45. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [CrossRef]
46. Jaakkola, P.; Hippeläinen, M.; Farin, P.; Rytönen, H.; Kainulainen, S.; Partanen, K. Interobserver variability in measuring the dimensions of the abdominal aorta: Comparison of ultrasound and computed tomography. *Eur. J. Vasc. Endovasc. Surg.* **1996**, *12*, 230–237. [CrossRef] [PubMed]
47. Singh, K.; Jacobsen, B.K.; Solberg, S.; Bønna, K.H.; Kumar, S.; Bajic, R.; Arnesen, E. Intra- and interobserver variability in the measurements of abdominal aortic and common iliac artery diameter with computed tomography. The Tromsø study. *Eur. J. Vasc. Endovasc. Surg.* **2003**, *25*, 399–407. [CrossRef] [PubMed]

48. Limet, R.; Sakalihassan, N.; Albert, A. Determination of the expansion rate and incidence of rupture of abdominal aortic aneurysms. *J. Vasc. Surg.* **1991**, *14*, 540–548. [[CrossRef](#)] [[PubMed](#)]
49. Powell, J.T.; Sweeting, M.J.; Brown, L.C.; Gotensparre, S.M.; Fowkes, F.G.; Thompson, S.G. Systematic review and meta-analysis of growth rates of small abdominal aortic aneurysms. *J. Br. Surg.* **2011**, *98*, 609–618. [[CrossRef](#)]
50. Hirata, K.; Nakaura, T.; Nakagawa, M.; Kidoh, M.; Oda, S.; Utsunomiya, D.; Yamashita, Y. Machine Learning to Predict the Rapid Growth of Small Abdominal Aortic Aneurysm. *J. Comput. Assist. Tomogr.* **2020**, *44*, 37–42. [[CrossRef](#)]
51. Lee, R.; Jarchi, D.; Perera, R.; Jones, A.; Cassimjee, I.; Handa, A.; Clifton, D.; Bellamkonda, K.; Woodgate, F.; Killough, N.; et al. Applied Machine Learning for the Prediction of Growth of Abdominal Aortic Aneurysm in Humans. *EJVES Short Rep.* **2018**, *39*, 24–28. [[CrossRef](#)]
52. Rengarajan, B.; Wu, W.; Wiedner, C.; Ko, D.; Muluk, S.C.; Eskandari, M.K.; Menon, P.G.; Finol, E.A. A Comparative Classification Analysis of Abdominal Aortic Aneurysms by Machine Learning Algorithms. *Ann. Biomed. Eng.* **2020**, *48*, 1419–1429. [[CrossRef](#)]
53. Chandrashekar, A.; Handa, A.; Lapolla, P.; Shivakumar, N.; Ngetich, E.; Grau, V.; Lee, R. Prediction of abdominal aortic aneurysm growth using geometric assessment of computerized tomography images acquired during the aneurysm surveillance period. *Ann. Surg.* **2023**, *277*, e175–e183. [[CrossRef](#)]
54. Lv, R.; Hu, G.; Zhang, S.; Zhang, Z.; Chen, J.; Wang, K.; Wang, Z.; Jin, Z. Assessing abdominal aortic aneurysm growth using radiomic features of perivascular adipose tissue after endovascular repair. *Insights Into Imaging* **2024**, *15*, 232. [[CrossRef](#)]
55. Kotanidis, C.P.; Antoniadis, C. Perivascular fat imaging by computed tomography (CT): A virtual guide. *Br. J. Pharmacol. Nov.* **2021**, *178*, 4270–4290. [[CrossRef](#)]
56. Oikonomou, E.K.; Williams, M.C.; Kotanidis, C.P.; Desai, M.Y.; Marwan, M.; Antonopoulos, A.S.; Thomas, K.; Thomas, S.; Akoumianakis, I.; Fan, L.M.; et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. *Eur. Heart J.* **2019**, *40*, 3529–3543. [[CrossRef](#)] [[PubMed](#)]
57. Ding, N.; Hao, Y.; Wang, Z.; Xuan, X.; Kong, L.; Xue, H.; Jin, Z. CT texture analysis predicts abdominal aortic aneurysm post-endovascular aortic aneurysm repair progression. *Sci. Rep.* **2020**, *10*, 12268. [[CrossRef](#)] [[PubMed](#)]
58. Rezaeitalshmahalleh, M.; Mu, N.; Lyu, Z.; Zhou, W.; Zhang, X.; Rasmussen, T.E.; McBane, R.D.; Jiang, J. Radiomic-based textural analysis of intraluminal thrombus in aortic abdominal aneurysms: A demonstration of automated workflow. *J. Cardiovasc. Trans Res.* **2020**, *16*, 1123–1134. [[CrossRef](#)] [[PubMed](#)]
59. Vittinghoff, E.; McCulloch, C.E. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am. J. Epidemiol.* **2007**, *165*, 710–718. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.