Hybrid artificial intelligence outcome prediction using features extraction from stress perfusion cardiac magnetic resonance images and electronic health records

Ebraham Alskaf[1], **Richard Crawley**[1], **Cian M. Scannell**[1,2], **Avan Suinesiaputra**[1], **Alistair Young**[1], **Pier-Giorgio Masci**[1], **Divaka Perera**[1], **Amedeo Chiribiri**[1]

[1]School of Biomedical Engineering & Imaging Sciences, King's College London, St Thomas' Hospital, London, UK

[2]Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

Background—Prediction of clinical outcomes in coronary artery disease (CAD) has been conventionally achieved using clinical risk factors. The relationship between imaging features and outcome is still not well understood. This study aims to use artificial intelligence to link image features with mortality outcome.

Methods—A retrospective study was performed on patients who had stress perfusion cardiac magnetic resonance (SP-CMR) between 2011 and 2021. The endpoint was all-cause mortality. Convolutional neural network (CNN) was used to extract features from stress perfusion images, and multilayer perceptron (MLP) to extract features from electronic health records (EHRs), both networks were concatenated in a hybrid neural network (HNN) to predict study endpoint. Image CNN was trained to predict study endpoint directly from images. HNN and image CNN

Correspondence to: Ebraham Alskaf.

Correspondence to: Ebraham Alskaf, MD. School of Biomedical Engineering & Imaging Sciences, King's College London, St Thomas' Hospital, 4th Floor Lambeth Wing, Westminster Bridge Road, London, SE1 7EH, UK. ebraham.alskaf@kcl.ac.uk.

were compared with a linear clinical model using area under the curve (AUC), F1 scores, and McNemar's test.

Results—Total of 1,286 cases were identified, with 201 death events (16%). The clinical model had good performance (AUC =80%, F1 score =37%). Best Image CNN model showed AUC =72% and F1 score =38%. HNN outperformed the other two models (AUC =82%, F1 score =43%). McNemar's test showed statistical difference between image CNN and both clinical model (P<0.01) and HNN (P<0.01). There was no significant difference between HNN and clinical model (P=0.15).

Conclusions—Death in patients with suspected or known CAD can be predicted directly from stress perfusion images without clinical knowledge. Prediction can be improved by HNN that combines clinical and SP-CMR images.

Keywords

Introduction

Background

It has been more than two decades since the launch of coronary artery disease (CAD) risk stratification scores, which commonly aim to estimate the 10-year risk of developing CAD and other related health outcomes, such as major adverse cardiovascular events, heart failure and mortality (1). The concept of risk factors in CAD was first coined by the Framingham Heart Study (FHS), which published its findings in 1957, and demonstrated the epidemiologic relations of smoking, raised blood pressure and cholesterol levels to the incidence of CAD. The findings were truly revolutionary for it helped bring about a change in the way medicine is practiced (2). There are several important and potentially modifiable risk factors for cardiovascular disease (CVD), such as hypertension (HTN), dyslipidaemia, diabetes mellitus (DM), obesity, smoking, chronic kidney disease (CKD), anxiety and depression, social isolation, low physical activity and poor diet. Non-modifiable risk factors also exist in fewer numbers, such as ethnicity and family history of CVD disease (3).

Conventionally, these risk factors are used as inputs in different risk scoring algorithms to produce a quantifiable output, used by clinicians to predict long-term risk, clinical outcome and prognosis. Recent literature has shown strong predictive power of non-invasive imaging modalities of CAD, adding important prognostic value in predicting outcomes in patients with known or suspected CAD. Broadly, these non-invasive imaging techniques are either focused on functional/ischaemia assessment (e.g., stress perfusion cardiac magnetic resonance (SP-CMR) (4), dobutamine stress echocardiography (DSE) (5) and myocardial perfusion scanning (MPS) (6); or focused on imaging the coronary anatomy directly [e.g., coronary computed tomography angiography (CCTA) (7)]. Whilst several methods of combining functional and anatomical non-invasive imaging have been proposed, these remain in largely within the research-domain (8,9).

The revolution of artificial intelligence (AI) and neural networks within the medical domain over the last decade, has led to real world clinical applications with many automated medical tasks, including predictive analytics. Emerging studies have shown that AI can detect traditionally difficult to diagnose conditions, and empower outcome prediction, in addition to many other applications in treatment, safety, patient adherence, administration and precision medicine (10).

Rationale and knowledge gap

The utilisation of non-invasive imaging to assess patient risk and diagnose CAD is increasing our understanding of long-term patient outcomes. Contemporaneously, the use of AI algorithms in assessing cardiac risk factors and clinical data to predict outcome in at-risk patients is also being developed and adopted. Whilst findings from non-invasive imaging have been incorporated in such models, the algorithms largely rely on clinician interpretation of the imaging. Any direct relationship between the acquired images themselves and predicted outcome has not been investigated and is poorly understood.

Objective

This study aims to assess the probability of predicting patient outcomes from SP-CMR images using a novel AI approach for outcome prediction. We present this article in accordance with the TRIPOD reporting checklist (available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-1/rc).

Method

Study design and population

This was a retrospective observational cohort study. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Research Ethics Committee of King's College London Partnership (No. 20/ES/0005) and individual consent for this retrospective analysis was waived. Patients undergoing SP-CMR at a single centre (Guy's & St Thomas' NHS Foundation Trust) between April 2011 and March 2021 were screened, and only those with completed studies including full reports and available images were included.

The total number of all-cause mortality events was obtained for the whole population from electronic patients record. The follow-up period was variable depending on the date of the clinical scan, with data collection completed on 20 August, 2021.

Inclusion criteria

We included only patients with complete adenosine stress perfusion study and good quality images, who had complete reports. Exclusion criteria were: reports which were blinded for research purposes, conflicting reports description between main body text and summary findings, terminated stress study due to complications, reports with missing tissue characterisation information, contraindications to using stress agent, mass perfusion studies, dobutamine stress studies, lung perfusion studies, poor response to stress agent, and mis-

labelled reports which were originally highlighted as perfusion studies but after inspection found to be otherwise.

Data extraction

Clinical data—Clinical data extraction was performed using natural language processing (NLP) AI-based application called CogStack (11), which allows to extract information from unstructured data sources in multiple formats. Once extracted, harmonised and processed, multiple uses of this unstructured data become possible based around information retrieval and extraction. For the purpose of this study, an NLP model was trained to extract CVD risk factors from unstructured data using medical terms from Systematic Nomenclature of Medicine Clinical Terms (SNOMED CT), whereas baseline characteristics of the population were extracted from structured data using Application Programming Interface (API) search engine called Elastic Search (12). Samples of documents were ingested into the Medical Concept Annotation Tool (MedCAT), which is used to link electronic health records (EHRs) to biomedical ontologies such as SNOMED-CT and Unified Medical Language System (UMLS) and train NLP models. For this study, SNOMED-CT UK version was used for annotation. Text files were tokenized, lemmatised and pre-processed, then used as inputs into the network with the corresponding labels. Initial self-supervised model was trained using named-entity recognition + linking (NER + L) annotation, this algorithm is used to extract and locate name entities in unstructured text into a pre-defined categories for labelling before training the model. Fine tuning was achieved with supervised learning after a group of expert clinicians labelled a sample of reports with the relevant medical terminology. MedCAT trainer used multiple neural networks architecture (long-short-term-memory (LSTM), gated recurrent unit (GRU), and transformers), and the best performing model was deployed into CogStack. All data were anonymised. The AI-based data extraction pipeline is explained in Figure 1.

Image data—For image data extraction, SP-CMR images included three series of frames representing 3 levels of slices: basal, mid and apical left ventricular (LV) slices. These perfusion images were extracted in 2 stages and reviewed by a level 3 CMR reader. During stage 1 the frame of peak signal intensity within the LV cavity was selected using an automated pipeline based on sum and peak pixels per frame. An automated identification and discard of low resolution arterial input function frames utilised pixel gradient algorithms. The use of a four frames per slice allowed for visualisaton of myocardial contrast wash-in/wash-out. In stage 2 the images were cropped to include only the LV myocardium and cavity using a centre crop function. Late gadolinium enhancement (LGE) images were also extracted, including three long axis views (2-chamber, 3-chamber and 4-chamber views) and multi-slice short axis imaging of the whole ventricles. Unique case identification numbers (IDs) were used to link each image series with the corresponding clinical data.

Construction of neural networks

Image convolutional neural network (CNN)—CNN architecture was used when training image-based prediction models. Different experiments were performed with different CNN architectures. Training started from a simple design with LeNet (13) to more

complex networks, this included: AlexNet (14), VGG19 (15), ResNet50 (16) and GoogleNet (17). Images were resized to equal height and width of 224 pixels and all frames were stacked for each case with input shape of [224, 224, 25] to include 12 images for stress perfusion and 13 images for LGE datasets. The final layer of each neural network was a Dense layer of one node with 'sigmoid' activation function to prediction either 1 for mortality event or 0 for none. Binary cross-entropy loss function was used, and an early stopping was used after monitoring F1 score in the validation set. Adam optimiser was chosen with a learning rate of 0.001.

Hybrid neural network (HNN)—A HNN was developed to have mixed input data, both the CMR images and clinical information, in order to extract features from both data types and predict outcome. CNN architecture was used to extract features from stress perfusion and LGE images after removing the top prediction layer and flattening the output to a Dense shape of 4. A multi-layer perceptron (MLP) with 2 Dense layers was used to extract features from continuous and categorical clinical variables after removing the top prediction layer and flattening the output to a Dense shape of 4, to be compatible with the output of CNN. Both outputs were then concatenated and passed to 2 Dense layers with the prediction in the final Dense layer with one node and 'sigmoid' activation function to predict all-cause mortality. Five different CNN architectures were used in the experiments in a similar approach to the image CNN experiments.

The full pipeline is shown in Figure 2.

Model training—Data was split into 60% for training, 15% for validation and 25% for testing. To overcome class imbalance, the initial bias was adjusted to reflect classes ratio and help training to make better initial guesses by setting the output layer's bias to reflect that, this can help with initial convergence. The initial bias $b0$ can be calculated as the following:

$$p0 = pos/(pos + neg) = 1/\left(1 + e^{-b0}\right)$$
$$b0 = -\log_e(1/p0 - 1) = \log_e(pos/neg)$$

[1]

This was followed by model building with the new initial bias, and outputting the bias after each epoch of training. Furthermore, class weights were calculated as:

$$w = N/(n \times 2)$$

[2]

Where w is the class weight, n is the number of class instances, N is the total size of the sample and 2 is the number of classes. Class weight was used as a parameter during fitting the model for training.

Training was monitored with early stopping based on validation precision/recall curve or F1 score, and binary cross-entropy was used as a loss function. After testing two optimisers

with various learning rates [Adam and Stochastic Gradient Descent (SGD)], Adam optimiser with a learning rate of 0.001 was used.

Statistical analysis

Categorical variables were expressed using number and percentage; continuous variables were expressed using mean and standard deviation. Follow-up was calculated as the mean time to all-cause mortality event, and all cases without events and with shorter duration from CMR date to collection date were excluded. The population was divided into three age subgroups given that different CVD risks vary based on adults age subgroups (18) (<65, 65–75, >75 years). The difference in baseline characteristics, clinical risk factors and CMR data between all subgroups was tested using the Chi-square test for categorical variables and a One-Way ANOVA for continuous variables.

Multivariate logistic regression was used as a baseline for comparison with neural networks performance using common CVD risk factors. Continuous variables were normalised using mean values, and categorical variables were used as binary (yes/no) input values, regardless of underlying types of classes. CVD risk factors included: age, gender, CKD, HTN, heart failure, smoking history, previous myocardial infarction, dyslipidaemia, DM and cerebrovascular accident (CVA). Similar approach was taken in building the MLP pipeline in HNN networks.

Testing model performance was evaluated using accuracy, precision, recall, area under the curve (AUC) and F1 score, and compared using McNemar's test. P value of <0.05 was considered statistically significant.

All statistical analysis and networks training was performed using Python programming language, version 3.10.

Results

Baseline characteristics

The extraction results and datasets used for models training are explained in Figure 3.

The total number of patients analysed was 1,286. The total number of patients who died was 201 (16%). Mean follow-up was 1,090 days (IQR, 351–1,749). The study population included males in around two thirds (66% male *vs* 34% female). More CMR studies were performed at 3 Tesla than 1.5 Tesla (62% *vs* 38%), using Siemens and Philips vendors. All stress perfusion imaging was conducted via the use of vasodilatory medication to achieve hyperaemia, with the vast majority performed using Adenosine, and a much smaller number receiving Regadenoson (88% *vs* 12%). Around a third of the population had inducible defects at peak hyperaemia (30%) and/or scar on LGE imaging (33%). All baseline characteristics, CMR data and clinical risk factors are shown in Table 1.

After dividing the population into three age subgroups, the older population (age >75 years) had a higher percentage of positive LGE and positive stress perfusion compared with the other age-defined subgroups, as shown in Figure 4.

Mortality prediction

The clinical model with logistic regression achieved good performance level (AUC =80%, F1 score =37%), with intercept of −4.17 and coefficients of CKD 0.84, HTN 0.53, male gender 0.59, heart failure 0.26, smoking 0.02, dyslipidaemia 0.13, DM 0.17 and age 4.10. Within image CNN, the best performing neural network was the AlexNet, which showed AUC =72% and F1 score =38%, however, there was poor convergence in training and signs of overfitting. For HNN, GoogleNet with three inception blocks was the best model for feature extraction from images. The performance of HNN was superior to both the image CNN and the clinical model (AUC =82%, F1 score =43%). Compared to image CNN, McNemar's test showed higher performance of both clinical model (P<0.01) and HNN (P<0.01). There was no significant difference between HNN and clinical model (P=0.15).

AUC and F1 scores for each image CNN model are shown in Table 2, for HNN models in Table 3, and for the clinical model in Table 4. Figure 5 shows the receiver operating characteristic (ROC) curves and the precision-recall curves comparison between image CNN, HNN and clinical model.

Discussion

Clinical practice relies on effective risk stratification to guide the management of patients with suspected or known CAD. Utilising clinical risk scores derived from large datasets and long periods of follow-up aids in clinical decision making, achieving an AUC of 80% in our study for the prediction of all-cause mortality based on clinical parameters. This highlights the high sensitivity and specificity of outcome prediction using conventional clinical risk factors and established prediction.

Recent literature emphasises the independent and additional prediction power of non-invasive imaging in CAD, with identified features linked to specific outcomes (4–7), such as high-risk plaque features identified on CCTA (19). Integrating non-invasive imaging into risk scoring algorithms is likely to enhance outcome prediction, encompassing mortality, ventricular arrhythmia, hospitalisation, and other related health outcomes.

The increasing availability and funding, along with technological advancements like higher imaging resolution, improved acceleration techniques, and lower radiation dosing, have made non-invasive imaging essential parts of daily clinical practice.

The prediction of mortality in CAD has been a crucial aspect of clinical practice, and non-invasive imaging, particularly stress perfusion CMR, has played a growing role in predicting treatment response and improving long term outcomes (20).

Our study pioneers the incorporation of image pixel data for predicting clinical outcomes using deep learning and AI technology. To the best of our knowledge, this is the first application of using AI to link image pixels to prognosis in stress perfusion CMR. While prediction from clinical risk factors and CMR findings outperformed image CNN, the HNN combining both types of data achieved the best AUC and F1 scores. This suggests

the potential clinical utility of AI, which may identify subtle features missed by human interpretation.

The integration of AI in risk stratification, as demonstrated in this study, holds promise for improving prognostic assessment in patients with suspected CAD. Future research could explore larger and multi-centre populations, and novel AI models using unsupervised learning might generalise predictions on unseen data.

Limitations

This study used retrospective cohort data with heavy class imbalance, only 16% of the population had an event, this can have an impact on the generalisation of the model to external datasets.

As there was no significant statistical difference between HNN and clinical model on McNemar's test, the redundant information in raw stress perfusion and LGE images are likely to have caused detrimental effects. More refined images with segmentation and quantification of the areas of interests are likely to improve the results in the future.

Conclusions

Direct prediction of mortality in patients with suspected CAD is achievable through analysis of stress perfusion images, even in the absence of clinical information. The utilisation of specific features or characteristics within stress perfusion images contributes to the accuracy of this direct prediction. Furthermore, the integration of HNN, which combines both clinical and stress perfusion CMR images demonstrates a notable improvement in mortality prediction.

This advancement in predictive capabilities has the potential to revolutionise clinical decision making in CAD. The ability to directly predict mortality, coupled with the refinement and combination of clinical and imaging data, positions HNNs as valuable tools in guiding treatment strategies and improving overall patient outcomes.

Acknowledgments

Data Sharing Statement

Available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-1/dss

References

1. Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. Circulation. 1998; 97: 1837–47. [PubMed: 9603539]

2. Hajar R. Risk Factors for Coronary Artery Disease: Historical Perspectives. Heart Views. 2017; 18: 109–14. [PubMed: 29184622]

3. Visseren FLJ, Mach F, Smulders YM, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J. 2021; 42: 3227–37. [PubMed: 34458905]

4. Pezel T, Unterseeh T, Kinnel M, et al. Long-term prognostic value of stress perfusion cardiovascular magnetic resonance in patients without known coronary artery disease. J Cardiovasc Magn Reson. 2021; 23: 43. [PubMed: 33827603]

5. Sicari R, Pasanisi E, Venneri L, et al. Stress echo results predict mortality: a large-scale multicenter prospective international study. J Am Coll Cardiol. 2003; 41: 589–95. [PubMed: 12598070]

6. Schinkel AF, Boiten HJ, van der Sijde JN, et al. Prediction of 9-year cardiovascular outcomes by myocardial perfusion imaging in patients with normal exercise electrocardiographic testing. Eur Heart J Cardiovasc Imaging. 2012; 13: 900–4. [PubMed: 22588207]

7. Deseive S, Shaw LJ, Min JK, et al. Improved 5-year prediction of all-cause mortality by coronary CT angiography applying the CONFIRM score. Eur Heart J Cardiovasc Imaging. 2017; 18: 286–93. [PubMed: 28363203]

8. Hajhosseiny R, Rashid I, Bustin A, et al. Clinical comparison of sub-mm high-resolution non-contrast coronary CMR angiography against coronary CT angiography in patients with low-intermediate risk of coronary artery disease: a single center trial. J Cardiovasc Magn Reson. 2021; 23: 57. [PubMed: 33993890]

9. Pontone G, Baggiano A, Andreini D, et al. Dynamic Stress Computed Tomography Perfusion With a Whole-Heart Coverage Scanner in Addition to Coronary Computed Tomography Angiography and Fractional Flow Reserve Computed Tomography Derived. JACC Cardiovasc Imaging. 2019; 12: 2460–71. [PubMed: 31005531]

10. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019; 6: 94–8.

11. CogStack. Available online: https://cogstack.org/about/

12. Kraljevic Z, Searle T, Shek A, et al. Multi-domain clinical natural language processing with MedCAT: The Medical Concept Annotation Toolkit. Artif Intell Med. 2021; 117 102083 [PubMed: 34127232]

13. LeNet Architecture: A Complete Guide. Available online: https://www.kaggle.com/code/blurredmachine/lenet-architecture-a-complete-guide

14. AlexNet Architecture: A Complete Guide. Available online: https://www.kaggle.com/code/blurredmachine/alexnet-architecture-a-complete-guide

15. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. doi: 10.48550/arXiv:1409.1556

16. Kaiming H, Xiangyu Z, Shaoqing R, et al. Deep Residual Learning for Image Recognition. doi: 10.48550/axrXiv:1512.03385

17. Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions. doi: 10.48550/arXiv:1409.4842

18. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular Risks Associated with Gender and Aging. J Cardiovasc Dev Dis. 2019; 6: 19. [PubMed: 31035613]

19. Lee SE, Sung JM, Andreini D, et al. Differences in Progression to Obstructive Lesions per High-Risk Plaque Features and Plaque Volumes With CCTA. JACC Cardiovasc Imaging. 2020; 13: 1409–17. DOI: 10.21037/jmai-24-1 [PubMed: 31734214]

20. Sammut EC, Villa ADM, Di Giovine G, et al. Prognostic Value of Quantitative Stress Perfusion Cardiac Magnetic Resonance. JACC Cardiovasc Imaging. 2018; 11: 686–94. [PubMed: 29153572]

Highlight box

Key findings

- Outcome prediction can be achieved directly from stress perfusion cardiac magnetic resonance images using artificial intelligence (AI).

- Outcome prediction using AI can be enhanced by mixed data types from image pixels and electronic health records.

**What is known and what is new?**

- Non-invasive cardiac imaging has additional and independent prognostic values.

- Outcome prediction using imaging data still relies on expert human interpretation of image findings.

- This study adds an insight into how to link image pixel data with prognosis using AI.

What is the implication, and what should change now?

- This study will open the door for a novel approach in outcome prediction using non-invasive imaging without former knowledge of patients' data.

- It also introduces a novel hybrid AI prognostic tool, which has the potential to overcome conventional clinical risk scoring.

- More research on larger and multi-centre datasets, and more refined imaging data have to be performed to achieve higher performance and novel clinical applications.
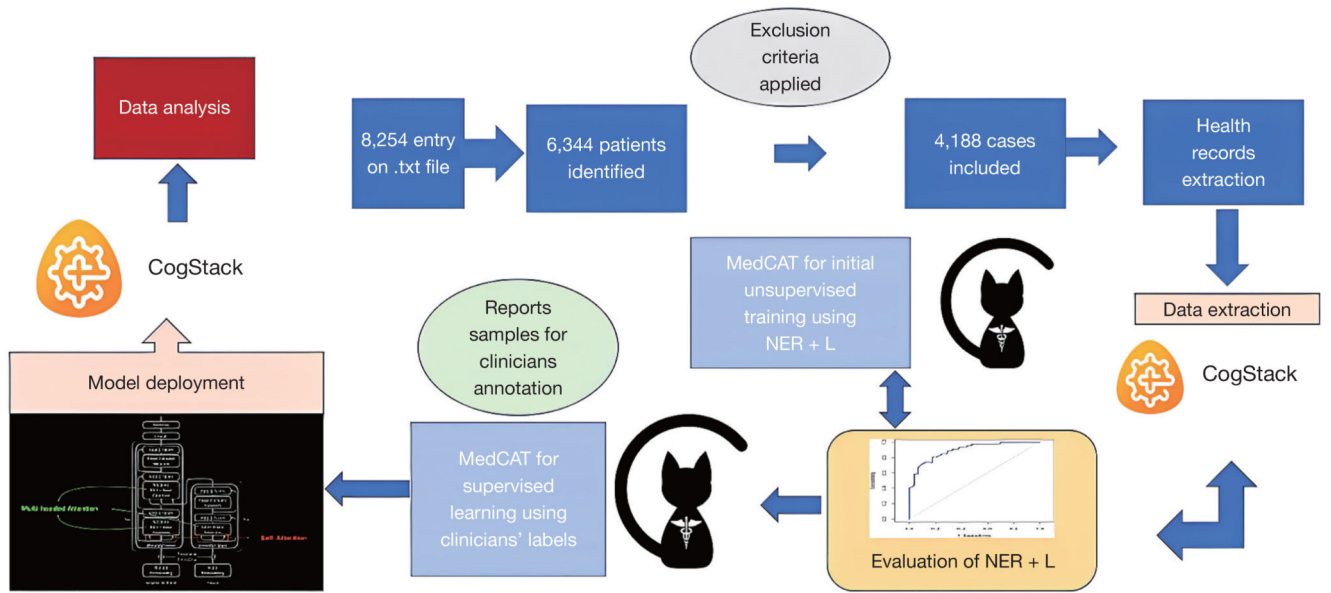
Figure 1.
A diagram showing the data extraction process using CogStack. MedCAT, Medical Concept Annotation Tool; NER + L, named-entity recognition + linking.
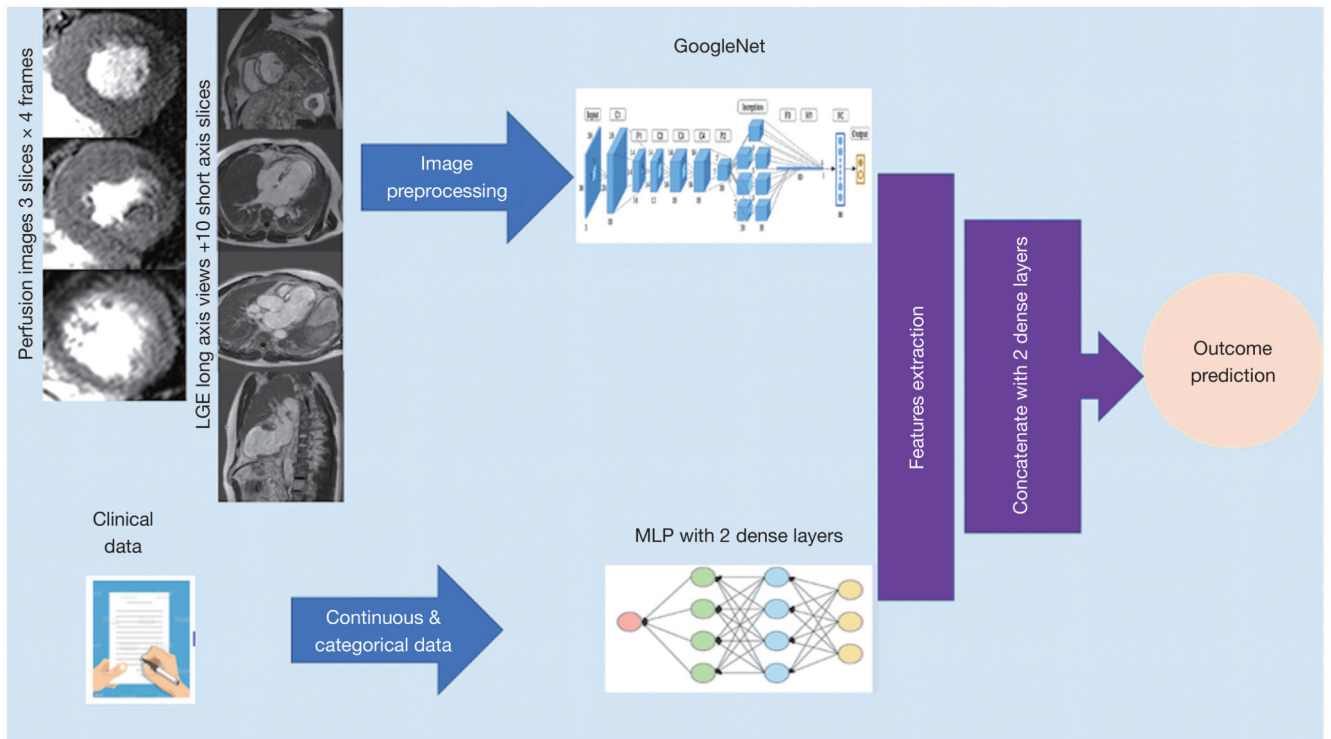
Figure 2.
Full pipeline showing the HNN architecture combining CNN (top) and MLP (bottom). CNN, convolutional neural network; HNN, hybrid neural network; LGE, late gadolinium enhancement; MLP, multi-layer perceptron.
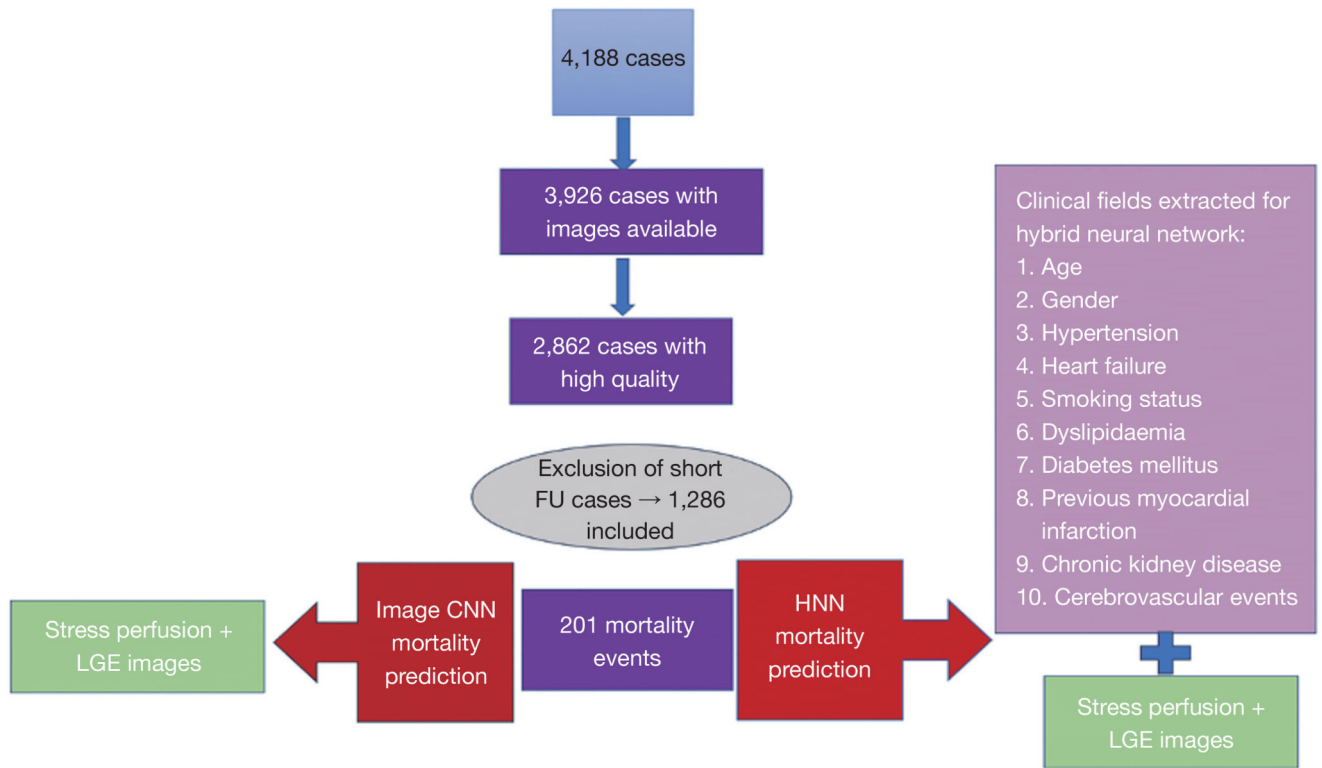
Figure 3.
A diagram showing sample selection for training both types of models, and the clinical variables used by the hybrid neural network. CNN, convolutional neural network; FU, follow-up; HNN, hybrid neural network; LGE, late gadolinium enhancement.
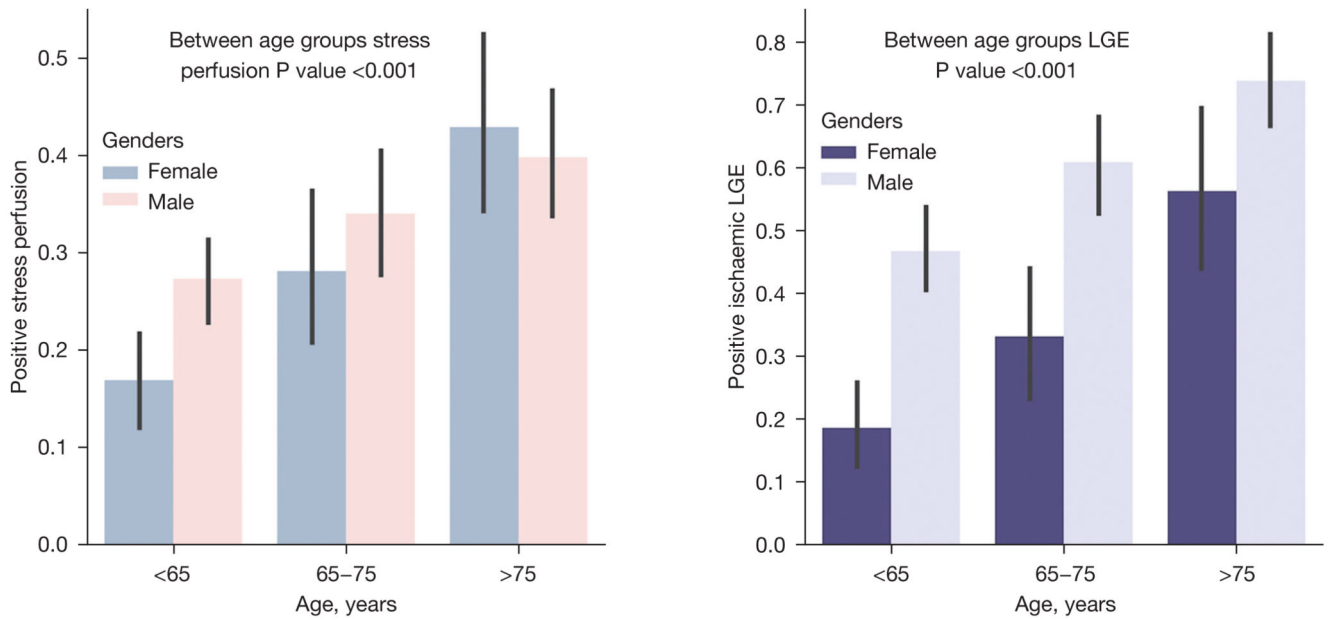
Figure 4.

Categorical bar plot showing different age groups with gender categories and comparison based on positive stress perfusion (left) and positive ischaemic LGE (right) (with 95% confidence intervals). LGE, late gadolinium enhancement.
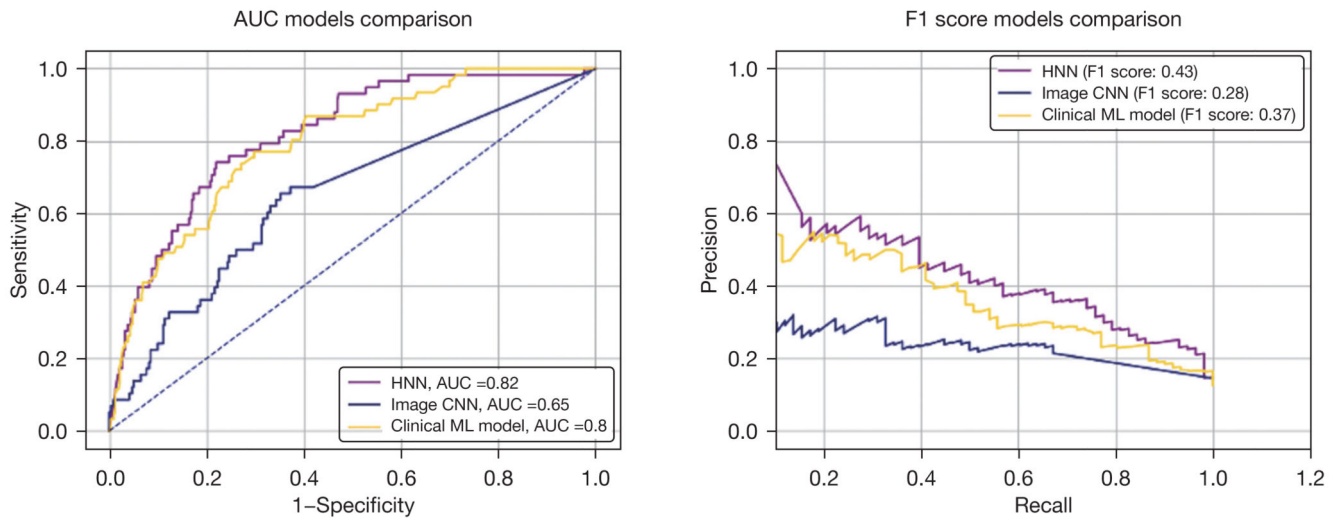
Figure 5.

The ROC curves plot (left) and the precision-recall curves plot (right) showing the high performance of HNN over clinical prediction model and image CNN for mortality prediction using GoogleNet architecture. AUC, area under the curve; CNN, convolutional neural network; ML, machine learning; HNN, hybrid neural network; ROC, receiver operating characteristic.

Table 1

Baseline characteristics by age subgroups

Variables	Total (n=1,286)	<65 years (n=577)	65–75 years (n=383)	>75 years (n=326)	P value
Death	201 [16]	28 [5]	63 [16]	110 [34]	<0.001 *
Sex					0.241
Male	845 [66]	370 [64]	250 [65]	225 [69]	
Female	441 [34]	207 [36]	133 [35]	101 [31]	
Clinical risk factors					
Smoking	147 [11]	47 [8]	56 [15]	44 [13]	0.004 *
DM	61 [5]	27 [5]	19 [5]	15 [5]	0.759
HTN	515 [40]	197 [34]	174 [45]	144 [44]	<0.001 *
Dyslipidaemia	281 [22]	117 [20]	106 [28]	58 [18]	0.022 *
CVA	116 [9]	37 [6]	47 [12]	32 [10]	0.004 *
CKD	81 [6]	15 [3]	30 [8]	36 [11]	<0.001 *
Previous MI	319 [25]	144 [25]	104 [27]	71 [22]	0.309
Heart failure	226 [18]	82 [14]	73 [19]	71 [22]	0.005 *
Arrhythmia					
AF	194 [15]	45 [8]	72 [19]	77 [24]	<0.001 *
Atrial flutter	63 [5]	22 [4]	26 [7]	15 [5]	0.059
VT	110 [9]	40 [7]	34 [9]	36 [11]	0.010 *
VF	15 [1]	7 [1]	6 [2]	2 [1]	0.724
Field strength					0.723
1.5 T	492 [38]	220 [38]	146 [38]	126 [39]	
3 T	794 [62]	356 [62]	234 [61]	204 [63]	
Stress agent					0.041 *
Adenosine	1,130 [88]	518 [90]	339 [89]	273 [84]	
Regadenoson	156 [12]	59 [10]	44 [11]	53 [16]	
LVEF	55±14	58±12	55±14	51±15	<0.001 *
RVEF	59±10	59±09	59±10	58±12	0.303
+ve ischaemia	384 [30]	137 [24]	118 [31]	129 [40]	<0.001 *
+ve LGE	424 [33]	109 [19]	141 [37]	174 [53]	<0.001 *

Values are presented as number [%] for categorical variables; mean ± standard deviation for continuous variables.

*
P<0.05. DM, diabetes mellitus; HTN, hypertension; CVA, cerebrovascular accident; CKD, chronic kidney disease; MI, myocardial infarction; AF, atrial fibrillation; VT, ventricular tachycardia; VF, ventricular fibrillation; T, Tesla; LVEF, left ventricular ejection fraction; RVEF, right ventricular ejection fraction; LGE, late gadolinium enhancement; +ve, at least one positive myocardial segment.

Table 2

Comparison of performance metrics for all image CNN models

Image CNN	Accuracy	Precision	Recall	AUC	F1 score
AlexNet	0.59	0.25	0.69	0.72	0.38
GoogleNet	0.81	0.28	0.40	0.65	0.28
LeNet	0.60	0.20	0.62	0.63	0.26
ResNet50	0.52	0.19	0.61	0.63	0.22
VGG19	0.52	0.24	0.59	0.61	0.27

CNN, convolutional neural network; AUC, area under the curve.

Table 3

Comparison of performance metrics for all HNN models

HNN	Accuracy	Precision	Recall	AUC	F1 score
AlexNet	0.75	0.24	0.70	0.74	0.36
GoogleNet	0.70	0.26	0.77	0.82	0.43
LeNet	0.72	0.25	0.73	0.76	0.37
ResNet50	0.67	0.22	0.74	0.75	0.25
VGG19	0.70	0.25	0.73	0.76	0.35

HNN, hybrid neural network; AUC, area under the curve.

Table 4

Comparison of performance metrics for the clinical model

Accuracy	Precision	Recall	AUC	F1 score
0.81	0.57	0.27	0.80	0.37

AUC, area under the curve.