

The Potential of Radiomic Analysis for Enhancing the Diagnostic Ability of PET and CMR in Cardiac Sarcoidosis

Nouf Abdulla Mushari

Submitted in accordance with the requirements for the degree of Doctor of
Philosophy



UNIVERSITY OF LEEDS

University of Leeds

Faculty of Medicine and Health

Leeds Institute of Cardiovascular and Metabolic Medicine

January 2024

Intellectual Property Statement

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

- I. “Exploring the Utility of Radiomic Feature Extraction to Improve the Diagnostic Accuracy of Cardiac Sarcoidosis Using FDG PET” **Chapter 3**
Frontiers in Medicine (<https://doi.org/10.3389/fmed.2022.840261>)

- II. “Exploring the Utility of Cardiovascular Magnetic Resonance Radiomic Feature Extraction for Evaluation of Cardiac Sarcoidosis” **Chapter 4**
Diagnostics (<https://doi.org/10.3390/diagnostics13111865>)

- III. “An Assessment of PET and CMR Radiomic Features for Detection of Cardiac Sarcoidosis” **Chapter 5**
Frontiers in Nuclear Medicine (<https://doi.org/10.3389/fnume.2024.1324698>)

The authors and their respective contributions to these studies are provided below:

Authors: Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

Author contributions: NAM (the candidate) segmented all datasets, extracted the radiomic features, performed the statistical analysis and feature selection, analysed the results and wrote the manuscript. GS shared datasets, reviewed segmentations and helped in modifying code as well as in the guidance of the project. LD wrote python code and helped to modify it and provide essential guidance on how to perform the optimisation of the radiomic analysis and machine learning approaches. MGT facilitated the availability of data. MGT, ZAF, and PR contributed to reviewing the manuscript and the overall guidance of the project and data. CT supervised the specific study and helped in restructuring and reviewing the manuscript.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Nouf Abdulla Mushari to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

©2024 The University of Leeds and Nouf Abdulla Mushari

Acknowledgements

I would like to express my appreciation to all those who provided me the possibility to complete this project. Firstly, I would like to thank Saudi Government through Taif University for funding my PhD programme. I am grateful for the opportunities and resources available to me, which have been instrumental in the successful completion of my project. I am deeply grateful to my supervisor, Prof Charalampos Tsoumpas, for his exceptional guidance, constant support, and insightful supervision. I would like to thank my tutor, Prof David Buckley, for all of his support during this journey. I would also like to extend my appreciation to Dr Georgios Soultanidis and Dr Philip Robson at Mount Sinai Hospital in New York for their valuable contributions and active participation through frequent scientific discussions. I would also like to thank my colleague, Dr Lisa Duff, whose contributions throughout the project have been immensely valuable. I am grateful for her support, which has greatly influenced my growth and development.

Finally, I would like to express my heartfelt appreciation to my family who have played an invaluable role in my journey towards completing my PhD thesis. To my dear parents, who have been my unwavering pillars of support, I am forever indebted to your love, sacrifices, and belief in my dreams. Your prayers, encouragement, nurturing guidance and boundless love have shaped me into the person I am today. I carry your love in my heart, always. To my siblings, who have been my constant cheerleaders, your constant support and belief in me have been a source of strength and inspiration. My deepest appreciation goes to my beloved husband, Meshal, your constant presence, patience,

encouragement, understanding and belief in my abilities have strengthened me during the most challenging moments. I am forever grateful for your support. My dearest daughter, Joud, you have been my little source of light and inspiration. Your innocent smile and boundless love have lifted my spirits and reminded me of the true purpose behind my endeavours. Your understanding and patience, though beyond your years, have touched my heart profoundly. Thank you for being my constant reminder of what truly matters. To my upcoming baby, Juman, though you are yet to arrive, the anticipation of your presence has filled my heart with joy and renewed my determination to succeed. I eagerly await the day when I can share this achievement with you, my precious little one.

Thank you all for your support. This project would not have been possible without each and every one of you.

Abstract

Cardiac sarcoidosis (CS) is a granulomatous inflammatory disease whose aetiology is unknown, which features the existence of non-caseating granulomas. This thesis addresses the challenge of accurately diagnosing CS by enhancing the diagnostic capabilities of [^{18}F]fluorodeoxyglucose positron emission tomography ([^{18}F]FDG PET) and late gadolinium-enhanced cardiac magnetic resonance imaging (LGE-CMR). Independently, these modalities face limitations in isolating CS with high specificity and sensitivity. The thesis aimed to improve the diagnostic efficiency by integrating [^{18}F]FDG PET and LGE-CMR through advanced radiomic feature analysis. Radiomic analysis was conducted across various scenarios, encompassing comparisons between positive and negative CS groups, distinguishing between active and inactive disease states, and differentiating CS patients from those experiencing myocardial inflammation due to another cause (post-COVID-19 patients).

The thesis concludes that radiomic analysis can enhance the objectivity and complementarity of PET and CMR in identifying cardiac sarcoidosis. While PET-based analyses demonstrate high performance, the project underscores the essential role of CMR-based analysis in mitigating challenges associated with PET image preparation variability.

Contents

Intellectual Property Statement	i
Acknowledgements	iii
Abstract	v
Contents	vi
Table of Figures	ix
Table of Tables	xi
Abbreviations	xii
Chapter 1 Introduction	1
1.1 Overview of Thesis	1
1.2 Key Contributions	2
1.3 Publications	3
Chapter 2 Background	9
2.1 Overview of Cardiac Sarcoidosis	9
2.2 Diagnostic Tests for Cardiac Sarcoidosis	12
2.2.1 Positron Emission Tomography (PET)	12
2.2.1.1 Basic PET Principles	12
2.2.1.2 Image Quantification	19
2.2.1.3 Role of PET in Cardiac Sarcoidosis Diagnosis	21
2.2.2 Magnetic Resonance Imaging (MR)	23
2.2.2.1 MR Theory	23
2.2.2.2 Image Acquisition	28
2.2.2.3 Role of MR in Cardiac Sarcoidosis Diagnosis	30
2.2.3 PET/MR Imaging	31
2.2.3.1 Attenuation Correction	32
2.2.3.2 Motion Correction	33
2.2.3.3 Scatter Correction	34
2.3 Radiomic Analysis	34
2.3.1 Factors Affecting Radiomic Analysis	35
2.3.1.1 Image Acquisition and Protocol Variability	36
2.3.1.2 Image Reconstruction	37
2.3.1.3 Image Segmentation Accuracy	38
2.3.2 Feature Selection and Analysis	39
2.4 Project Justification	44

References.....	45
Chapter 3 Exploring the Utility of Radiomic Feature Extraction to Improve the Diagnostic Accuracy of Cardiac Sarcoidosis Using FDG PET	57
3.1 Introduction.....	58
3.2 Materials and Methods.....	60
3.2.1 Ethical Approval	60
3.2.2 Subject Selection	60
3.2.3 Imaging Protocol	61
3.2.4 Segmentations	61
3.2.4.1 Segmentation A	62
3.2.4.2 Segmentation B	63
3.2.5 Feature Extraction.....	63
3.2.6 Statistical Analysis	63
3.3 Results.....	64
3.3.1 Conventional Metrics Diagnostic Utility	64
3.3.2 Individual Radiomic Features Diagnostic Utility	65
3.3.3 Principal Component Analysis and Machine Learning	67
3.4 Discussion	70
3.5 Conclusion.....	73
References.....	74
Supplementary Materials	78
Chapter 4 Exploring the Utility of Cardiovascular Magnetic Resonance Radiomic Feature Extraction for Evaluation of Cardiac Sarcoidosis	81
4.1 Introduction.....	82
4.2 Materials and Methods.....	83
4.2.1 Ethical Approval	83
4.2.2 Subject selection	84
4.2.3 Imaging Protocols	84
4.2.4 Patient Classification.....	85
4.2.5 Segmentation	85
4.2.6 Feature Extraction.....	86
4.2.7 Statistical Analysis	87
4.3 Results.....	88
4.3.1 Individual Features—Diagnostic Utility.....	88
4.3.2 Signature Building and Machine Learning Performance	90

4.4 Discussion	94
4.5 Conclusion	97
References.....	98
Supplementary Materials	102
Chapter 5 An Assessment of PET and CMR Radiomic Features for Detection of Cardiac Sarcoidosis	103
5.1 Introduction.....	104
5.2 Materials and Methods.....	106
5.2.1 Ethical Approval	106
5.2.2 Subject Selection	107
5.2.3 Imaging Protocol	108
5.2.4 Segmentation	109
5.2.5 Feature Extraction.....	110
5.2.6 Statistical Analysis	111
5.3 Results.....	113
5.3.1 Individual Features Assessment.....	113
5.3.2 Feature Selection	114
5.4 Discussion	120
5.5 Conclusion	124
References.....	125
Supplementary Materials	130
Chapter 6 Discussion.....	133
6.1 Overview of Results	133
6.2 Limitations.....	136
6.3 Future Directions and Considerations	139
6.4 Conclusion	142
References.....	143
Thesis Addendum.....	147

Table of Figures

Figure 2.1 Photons coincidence events.....	19
Figure 2.2 The effect of the external magnetic field on the hydrogen nuclei	25
Figure 2.3 The representation of hydrogen nucleus spins in a static magnetic field B_0	25
Figure 3.1 Area under the curve (AUC) and accuracy with stratified five-fold cross-validation of the conventional metrics of (A) segmentation A and (B) segmentation B. SUV: standardised uptake value, TBR_{max} : maximum target-to-background ratio.....	65
Figure 3.2 Area under the curve (AUC) and accuracy with stratified five-fold cross-validation of the five best-performing radiomic features of (A) segmentation A and (B) segmentation B based on AUC values. GLSZM: Gray Level Size Zone Matrix, LAHGLE: Large Area High Gray Level Emphasis, GLCM: Gray Level Co-occurrence Matrix, MCC: Maximal Correlation Coefficient, GLCM C: Correlation, GLDM: Gray Level Dependence Matrix, LDHGLE: Large Dependence High Gray Level Emphasis, DV: Dependence Variance, DNU: Dependence Non-Uniformity, GLRLM: Gray Level Run Length Matrix, RLNU: Run Length Non-Uniformity, HGLZE: High Gray Level Zone Emphasis, NGTDM: Neighbouring Gray Tone Difference Matrix, NGTDM C: Complexity, SAHGLE: Small Area High Gray Level Emphasis.	67
Figure 3.3 Areas under the curve (AUC) and accuracies of machine learning classifiers for (A) Segmentation A and (B) Segmentation B. rf: Random Forest, lgr: Logistic Regression, svm: Support Vector Machine, dt: Decision Tree, gpc: Gaussian Process Classifier, sgd: Stochastic Gradient Descent, perc: Perceptron Classifier, pasagr: Passive Aggressive Classifier, nnet: Neural Network Classifier, kneigh: K-neighbors Classifier.....	68
Figure 3.4 The machine learning classifiers with high performance in (A) & (B) Segmentation A and (C) Segmentation B.	69
Figure 4.1 Radiomic features before and after removing correlated features with 95% confidence intervals. MAD: mean absolute deviation, IR: interquartile range, RMAD: robust mean absolute deviation, GLCM: gray level co-occurrence matrix, JE: joint entropy, SE: sum entropy, DA: difference average.....	91
Figure 4.2 Flowchart showing the performance of the LGE-CMR based radiomic feature gray level co-occurrence matrix joint entropy (GLCM_JE) in discriminating active from inactive disease in cardiac sarcoidosis (CS).	92
Figure 4.3 Areas under the curve (AUC) and accuracies of machine learning (ML) classifiers of both signatures with 95% confidence intervals. rf: random forest, lgr: logistic regression, svm: support vector machine, dt: decision tree, gpc: Gaussian process classifier, sgd: stochastic gradient descent, perc: perceptron classifier, pasagr: passive aggressive classifier, nnet: neural network classifier, kneigh: k-neighbors classifier.	93

Figure 4.4 The performance of the best machine learning classifiers for both signatures, (a,b) for signature A, and (c) for signature B.	94
Figure 5.1 A flowchart of datasets selection	108
Figure 5.2 An example of the segmentation on PET/CMR image.	110
Figure 5.3 Statistical analysis workflow	113
Figure 5.4 Two cases of PET/CMR cardiac sarcoidosis (CS) and one case of post-COVID-19 (PC) associated with the best-performance features. PET/CMR CS (case 1) has significantly lower values than PC values, whereas (case 2) has values within the range of PC patients' values, potentially leading to a misdiagnosis. Display intensity of PET images is from 0 to 6.....	117

Table of Tables

Table 2.1 Common positron isotopes characteristics.....	15
Table 3.1 Conventional metrics and five best performing radiomic features for the different segmentations based on p -values. SUV: Standardised Uptake Value, TBR_{max} : maximum Target-to-Background Ratio, GLDM: Gray Level Dependence Matrix, GLCM: Gray Level Co-occurrence Matrix, GLSZM: Gray Level Size Zone Matrix, NGTDM: Neighbouring Gray Tone Difference Matrix.....	66
Table 4.1 The ten best-performing radiomic features based on the p -values. CS_{active} : active cardiac sarcoidosis, $CS_{inactive}$: inactive cardiac sarcoidosis, GLCM: gray level co-occurrence matrix, GLRLM: gray level run length matrix, GLDM: gray level dependence matrix, GLSZM: gray level size zone matrix.....	89
Table 5.1 Demographic information of the study population. PC: Post-COVID-19 Patients; CS: Cardiac Sarcoidosis Patients	108
Table 5.2 Five best radiomic features based on the p -values. GLSZM: gray level size zone matrix; GLRLM: gray level run length matrix; GLDM: gray level dependence matrix.....	114
Table 5.3 Areas under the curve (AUCs) and accuracies (ACC) of uncorrelated features. CI: confidence interval; GLSZM: gray level size zone matrix; GLDM: gray level dependence matrix; GLRLM: gray level run length matrix.....	115
Table 5.4 The machine learning classifiers performance of PET joint features with 95% confidence intervals (CI). ACC: Accuracy; AUC: area under the curve.....	116
Table 5.5 Areas under the curve (AUCs) and accuracies (ACC) of uncorrelated features using the Support Vector Machine (SVM) that used to create a CMR signature. CI: confidence interval; GLSZM: gray level size zone matrix; GLRLM: gray level run length matrix; GLDM: gray level dependence matrix.....	118
Table 5.6 The machine learning classifiers performance of LGE-CMR joint features with 95% confidence intervals (CI). ACC: Accuracy; AUC: area under the curve.....	119

Abbreviations

¹¹C	Carbon
¹³N	Nitrogen
¹⁵O	Oxygen
¹⁸F	Fluorine
3D	Three-dimensional
⁶⁸G	Gallium
ACC	Accuracy
AUC	Area Under the Curve
CI	Confidence Interval
CMR	Cardiovascular Magnetic Resonance
CS	Cardiac Sarcoidosis
CT	Computed Tomography
ECG	Electrocardiogram
FDG	Fluorodeoxyglucose
FLT	Fluorothymidine
GLCM	Gray Level Co-occurrence Matrix
GLDM	Gray Level Dependence Matrix
GLRLM	Gray Level Run Length Matrix
GLSZM	Gray Level Size Zone Matrix
HGLZE	High Gray Level Zone Emphasis
HRS	Heart Rhythm Society
IBSI	Image Biomarker Standardisation Initiative
JE	Joint Entropy
JMHW	Japanese Ministry of Health and Welfare
LAHGLE	Large Area High Gray Level Emphasis
LDHGLE	Large Dependence High Gray Level Emphasis
LGE	Late Gadolinium Enhancement

MR	Magnetic Resonance
NGTDM	Neighbouring Gray Tone Difference Matrix
OP-OSEM	Ordinary Poisson Ordered Subset Expectation Maximisation
PC	Post-COVID-19
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PSIR	Phase-Sensitive Inversion Recovery
PVE	Partial Volume Effect
RF	Radiofrequency
RLNU	Run Length Non-Uniformity
RMAD	Robust Mean Absolute Deviation
ROI	Region of Interest
SAHGLE	Small Area High Gray Level Emphasis
SPECT	Single-Photon Emission Computed Tomography
SUV	Standardised Uptake Value
SVM	Support Vector Machine
TBR_{max}	Maximum Target-to-Background Ratio
TOF	Time-of-Flight

Chapter 1 Introduction

1.1 Overview of Thesis

This thesis consists of six main chapters that collectively explore the theme of radiomic analysis in cardiac sarcoidosis (CS) to assist in the diagnosis. Each chapter provides a distinct perspective on the subject, as outlined below.

Chapter 2 Background

This chapter provides an overview of the fundamental knowledge related to this thesis. It covers the physics of the diagnostic tests that are used in the thesis and their role in the diagnosis of cardiac sarcoidosis. In addition, the hybrid PET/MR imaging and its considerations is provided. Moreover, the chapter discusses radiomic analysis and the process of feature selection.

Chapter 3 Exploring the Utility of Radiomic Feature Extraction to Improve the Diagnostic Accuracy of Cardiac Sarcoidosis Using FDG PET

In this chapter, the study aimed to examine the diagnostic efficacy of radiomic features extracted from [¹⁸F]FDG PET in comparison to conventional metrics. Additionally, the study sought to identify the most effective machine learning classifier for developing an automated model.

Chapter 4 Exploring the Utility of Cardiovascular Magnetic Resonance Radiomic Feature Extraction for Evaluation of Cardiac Sarcoidosis

The main focus of the study was to investigate the efficacy of radiomic features derived from LGE-CMR images in distinguishing between active and inactive

cardiac sarcoidosis. In addition, the study assessed the performance of machine learning classifiers on both individual features and principal components.

Chapter 5 An Assessment of PET and CMR Radiomic Features for Detection of Cardiac Sarcoidosis

The study investigated the utility of PET and LGE-CMR radiomic features in differentiating CS from other conditions characterised by myocardial inflammation. The primary emphasis was placed on patients who displayed cardiac-related symptoms following COVID-19 (referred to as PC patients).

Chapter 6 Discussion

This chapter provides a concise summary of the project's findings. The limitations encountered in this project are discussed. Furthermore, a comprehensive analysis is provided to highlight potential future directions for advancing this line of research.

1.2 Key Contributions

The key contributions of this thesis are listed as follows:

- I. The diagnostic utility of PET and LGE-CMR radiomic features has been explored from various perspectives. Numerous individual and combined radiomic features have demonstrated their potential value in diagnosing cardiac sarcoidosis across different scenarios.

- II. The methodology employed in this project has the potential to provide new possibilities for utilising radiomic analysis as a valuable source of information in the diagnosis of cardiac sarcoidosis.
- III. The effect of two manual segmentation approaches on the robustness of radiomic analysis among PET images were investigated and discussed. The factors that affected the approaches were also discussed.

1.3 Publications

Journal Papers

- I. “An assessment of PET and CMR radiomic features for detection of cardiac sarcoidosis” **Published**

Frontiers in Nuclear Medicine (<https://doi.org/10.3389/fnume.2024.1324698>)

Authors: Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

Author contributions: NAM (the candidate) segmented all datasets, extracted the radiomic features, performed the statistical analysis and feature selection, analysed the results and wrote the manuscript. GS shared datasets, reviewed segmentations and helped in modifying code as well as in the guidance of the project. LD wrote python code and helped to modify it and provide essential guidance on how to perform the optimisation of the radiomic analysis and machine learning approaches. MGT facilitated the availability of data. MGT, ZAF, and PR contributed to reviewing the manuscript and the overall guidance

of the project and data. CT supervised the specific study and helped in restructuring and reviewing the manuscript.

- II. “The clinical value of quantitative cardiovascular molecular imaging: a step towards precision medicine” **Published**

The British Journal of Radiology (<https://doi.org/10.1259/bjr.20230704>)

Authors: Hendrea Sanne Aletta Tingen, Gijs D. van Praagh, Pieter H. Nienhuis, Alwin Tubben, Nick D. van Rijsewijk, Derk ten Hove, **Nouf A. Mushari**, T. Samara Martinez-Lucio, Oscar I. Mendoza-Ibañez, Joyce van Sluis, Charalampos Tsoumpas, Andor W.J.M. Glaudemans and Riemer H.J.A. Slart

- III. “Exploring the Utility of Cardiovascular Magnetic Resonance Radiomic Feature Extraction for Evaluation of Cardiac Sarcoidosis” **Published**

Diagnostics (<https://doi.org/10.3390/diagnostics13111865>)

Authors: **Nouf A. Mushari**, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

Author contributions: NAM (the candidate) segmented all datasets, extracted the radiomic features, performed the statistical analysis and feature selection, analysed the results and wrote the manuscript. GS shared datasets, reviewed segmentations and helped in modifying code as well as in the guidance of the project. LD wrote python code and helped to modify it and provide essential

guidance on how to perform the optimisation of the radiomic analysis and machine learning approaches. MGT facilitated the availability of data. MGT, ZAF, and PR contributed to reviewing the manuscript and the overall guidance of the project and data. CT supervised the specific study and helped in restructuring and reviewing the manuscript.

- IV. “Exploring the utility of radiomic feature extraction to improve the diagnostic accuracy of cardiac sarcoidosis using FDG PET” **Published**

Frontiers in Medicine (<https://doi.org/10.3389/fmed.2022.840261>)

Authors: Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

Author contributions: NAM (the candidate) segmented all datasets, extracted the radiomic features, performed the statistical analysis and feature selection, analysed the results and wrote the manuscript. GS shared datasets, reviewed segmentations and helped in modifying code as well as in the guidance of the project. LD wrote python code and helped to modify it and provide essential guidance on how to perform the optimisation of the radiomic analysis and machine learning approaches. MGT facilitated the availability of data. MGT, ZAF, and PR contributed to reviewing the manuscript and the overall guidance of the project and data. CT supervised the specific study and helped in restructuring and reviewing the manuscript.

- V. “Assessment of different quantification metrics of [¹⁸F]-NaF PET/CT images of patients with abdominal aortic aneurysm” **Published**

Journal of Nuclear Cardiology (<https://doi.org/10.1007/s12350-020-02220-2>)

Authors: Mercy I Akerele, **Nouf A Mushari**, Rachael O Forsythe, Maaz Syed, Nicolas A Karakatsanis, David E Newby, Marc R Dweck, Charalampos Tsoumpas

Author contributions: MIA conceived the idea, reconstructed all the datasets using STIR software and supervised the study. NAM (the candidate) and MIA analysed the data using AMIDE software, performed the statistical analysis, analysed the results and wrote the manuscript. ROF, MRD and MS conducted the SoFIA3 study and supplied the patient data used for the study. NAK converted all the raw patient data from the scanner into STIR-readable format for reconstruction. DEN is the chief investigator of the SoFIA3 study. CT proposed and co-supervised the study and helped towards restructuring of the manuscript.

Conference abstracts

- I. “An assessment of the specificity of PET and CMR radiomic features for detection of cardiac sarcoidosis”

The Artificial Intelligence and Informatics in Nuclear Medicine (AINM), 2023, Groningen, Netherlands. **“Poster presentation”**

Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

- II. “Assessing the specificity of PET and MR radiomic features in cardiac sarcoidosis detection”

European Association of Nuclear Medicine (EANM), 2023, Vienna, Austria.

“Poster presentation”

Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

- III. “Exploring the Utility of Radiomic Feature Extraction to Improve the Diagnostic Accuracy of Cardiac Sarcoidosis Using FDG PET”

PET/MR and SPECT/MR & Total-body PET (PSMR-TBP), 2022, Isola d'Elba, Italy. **“Poster presentation”**

Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

- IV. “Exploring the Utility of Radiomic Feature Extraction to Improve the Diagnostic Accuracy of CS Using FDG PET”

European Molecular Imaging Meeting (EMIM), 2021, Göttingen, Germany.

“Oral presentation”

Nouf A. Mushari, George Soultanidis, Lisa Duff, Maria G. Trivieri, Zahi A. Fayad, Philip Robson, Charalampos Tsoumpas

V. “Assessment of different quantification metrics of [18F]-NaF PET/CT images of patients with abdominal aortic aneurysm”

European Molecular Imaging Meeting (EMIM), 2020, Virtual. **“Oral presentation”**

Nouf A Mushari, Mercy I Akerele, Rachael O Forsythe, Maaz Syed, Nicolas A Karakatsanis, David E Newby, Marc R Dweck, Charalampos Tsoumpas

Chapter 2 Background

2.1 Overview of Cardiac Sarcoidosis

Sarcoidosis comprises a granulomatous inflammatory disease involving multiple systems whose aetiology is unknown, which features the existence of non-caseating granulomas in the impacted organs (1, 2). The organs largely affected by sarcoidosis are the lungs. In more than 90% of documented cases, the disease is observed to develop in the pulmonary system (3, 4).

Nevertheless, extrapulmonary organs such as the heart can also be impacted (5). Cardiac sarcoidosis (CS) has a prevalence of 10 to 40 individuals per 100,000 in both the United States and Europe. Notably, the prevalence among Blacks is higher at 35.5 per 100,000 compared to Whites at 10.9 per 100,000 (6). In the clinical context, it is rare for cardiac involvement to occur, as it only manifests itself in around 5% of patients diagnosed with sarcoidosis, and symptoms may not be evident (i.e. 'clinically silent' disease), which is supported by the increased proportion of cardiac involvement identified in autopsy studies. A diagnosis cardiac involvement has been made in a minimum of 25% of sarcoid patients (7-9).

Despite the CS aetiology not being known, it has been recognised by scientists that the formation of the granuloma occurs as a result of a principal inciting event. Sarcoid lesions are characterised by discrete, compact non-caseating epithelioid. There are two different types of cells contained in the granuloma: lymphocytes and mononuclear macrophages. Additionally, numerous possible antigens exist that could facilitate the creation of granuloma lesions such as infections and environmental agents (10). It is possible for the formation of non-

caseating granulomas to form in distinct areas of the heart such as the left and right atria, right ventricle, left ventricular wall, basal ventricular septum, and papillary muscles (11). In cases involving medical symptoms, the disease can clinically manifest itself in different ways based on the site and profusion of the developed granulomas.

Patients diagnosed with CS are often determined to have complete heart block. There is an increased likelihood that this will manifest in younger CS patients, dissimilar to other conditions, and could lead to sudden death. As reported by Matsui, Y. et al. (12), 23% to 30% of patients diagnosed with myocardial sarcoidosis were determined to have complete heart block, whereas 12% to 32% of such cases involved bundle branch block. The causal factors include basal septum involvement, granulomas, or nodal artery involvement, ultimately causing ischaemia within the conduction system. Another condition that clinically manifests in cases of CS is ventricular arrhythmia. Around 25% to 65% of deaths caused by CS are attributed to sudden death resulting from ventricular arrhythmias or complete heart block (13). Around 40% of CS patients are observed to exhibit such manifestations as initial presentations of the disease. Additionally, approximately 25% to 75% of CS mortality as a result of cardiac issues is purely due to congestive heart failure (14). It is rare to observe pericardial effusion and valvular involvement, which manifests itself in 3% to 10% of CS patients, while 10% of patients have ventricular aneurysms resulting from extensive ventricular lesions (15). A further clinical manifestation is pulmonary hypertension which can be caused by inadequate functioning of the left ventricular and is an indicator of a low chance of recovery. In a study conducted by Shorr, A.F. et al. (16) involving 363 patients with CS, it was

reported that pulmonary hypertension was diagnosed in 74% of them for whom a lung transplant was required.

The diagnosis of CS can be difficult due to the absence of biomarkers or diagnostic tests with increased reliability. It has been shown that the specificity and sensitivity of frequently used invasive and non-invasive cardiac diagnostic tests, including endomyocardial biopsy, single-photon emission computed tomography (SPECT) and electrocardiogram (ECG) is reduced (17).

Furthermore, the criterion standards for diagnosing and screening CS are significantly limited. The Japanese Ministry of Health and Welfare (JMHW) (18) initially published CS diagnosis guidelines which were subsequently revised in 2006 (19). However, extensive validation of these guidelines has not occurred. One of the significant constraints of the guidelines is the necessity for cardiac involvement to be confirmed, whether it is through endomyocardial biopsy or by combining minor and major criteria. The Heart Rhythm Society published the criteria which comprise an expert consensus statement and incorporate the utilisation of state-of-the-art imaging techniques for the diagnosis of CS, including positron emission tomography (PET) and cardiovascular magnetic resonance imaging (CMR) (20). Nevertheless, one of the recommendations of these criteria is that extra-CS must be confirmed, which is not consistently feasible, particularly in cases of isolated CS. The most recent diagnostic guidelines for CS were published by the Japanese Circulation Society and its collaborative organisations (21). These guidelines provide criteria for diagnosing isolated CS, including confirmation of the disease through histological analysis using endomyocardial biopsy or clinical assessment based on abnormal uptake observed in [¹⁸F]FDG PET images of the heart. Additionally, the presence of at

least three other major criteria is required for diagnosis according to these guidelines, Table 2.1.

Table 2.1 Japanese Circulation Society guidelines for isolated cardiac sarcoidosis diagnosis

Criteria for cardiac involvement of sarcoidosis	
Major criteria	(a) High-grade atrioventricular block or fatal ventricular arrhythmia
	(b) Basal thinning of the ventricular septum or abnormal ventricular wall anatomy
	(c) Left ventricular contractile dysfunction
	(d) ⁶⁷ Ga citrate scintigraphy or [¹⁸ F]FDG PET reveals abnormally high uptake in the heart
	(e) Gadolinium-enhanced MRI reveals delayed contrast enhancement of the myocardium
Minor criteria	(f) Abnormal ECG findings
	(g) Perfusion defects on myocardial perfusion scintigraphy (SPECT)
	(h) Endomyocardial biopsy
Criteria for isolated cardiac sarcoidosis	
Histological diagnosis group	
Isolated cardiac sarcoidosis is diagnosed histologically when non-caseating epithelioid granulomas are detected in endomyocardial biopsy or surgical specimens.	
Clinical diagnosis group	
Isolated cardiac sarcoidosis is diagnosed clinically when the criterion (d) and at least three other criteria of the major criteria (a)-(e) are satisfied.	

The accuracy with which PET and CMR can effectively identify cardiac involvement is analogous to that of autopsy studies, while their prognostic ability is superior to classical clinical criteria (22, 23). The contribution of state-of-the-art imaging techniques to enhancing the identification and treatment of CS patients has been identified in the literature, which will be discussed in the following section. Such imaging techniques currently have critical importance for diagnosing cases at an early stage, predicting the disease and how it will progress, as well as the process of monitoring the response to therapy.

2.2 Diagnostic Tests for Cardiac Sarcoidosis

2.2.1 Positron Emission Tomography (PET)

2.2.1.1 Basic PET Principles

The PET camera is used to measure how positron-emitting radionuclides (tracers) are distributed inside an object. An advantage of PET imaging is that it offers metabolically significant radionuclides. Another distinctive tomographic ability is derived from the fact that two virtually back-to-back 511-keV gamma rays are simultaneously emitted after the annihilation of an electron, as well as the capability to facilitate precise quantitation of the uptake of tracers. PET has medical significance due to the existence of numerous beneficial positron-emitting tracers, including isotopes of oxygen (^{15}O), nitrogen (^{13}N), and carbon (^{11}C), which are fundamental elements of every living organism, as well as their physiologic processes. Therefore, it is possible to synthesise more tissue-specific and chemistry-specific tracers for administration into animals or humans in order to conduct in vivo studies of the physiology and pathophysiology. Fluorine (^{18}F) is an additional positron-emitting isotope that has importance, which is frequently employed for labelling the glucose analogue [^{18}F]FDG that is transported via the glucose pathway on its route between plasma and tissue cells. Nevertheless, dissimilar to glucose, FDG is constrained within the cell and it is not metabolised further; therefore, it can be utilised for the purpose of imaging (24).

Imaging techniques are generally grouped into two categories of anatomic or functional devices. Complementary data can be provided by these techniques that can be combined in order to facilitate the diagnostic process, as well as to plan, perform and evaluate therapy effectiveness. In fact, it is possible to enhance the process of interpreting functional images in the field of nuclear medicine by co-registering them with anatomic images.

The nucleus of an atom comprises neutrons and protons, collectively known as nucleons. While the electric charge of a proton is positive, a neutron possesses zero net electric charge. Two types of forces act on the neutrons and protons within the nucleus: a short-range powerful attractive force that serves to bind all nucleons together and a repulsive electric force that causes protons to repel each other. The stability of a nucleus rich in protons can be enhanced through a reduction in their surplus positive electric charge. A viable process of decay for a nucleus rich in protons is where a proton decays into a neutron, a neutrino, as well as a particle with a positive charge defined as a positron. As a proton's mass is less than that of a neutron, it is only possible for the process of decay to occur within the nucleus. Because a neutrino comprises a particle with almost no mass and no electric charge, there is no interaction in the tissue that surrounds it. The positron (β^+) is the electron's (β^-) antiparticle and although its mass is identical to an electron, the charge is opposite. It is possible to use an accelerator (e.g. cyclotron) to produce a proton-rich isotope in which a beam of high-energy deuterons or protons is generated for the purpose of penetrating the intended nuclei such that an increased number of protons can be embedded in them (25).

The kinetic energy produced when the nucleus decays is distributed between the neutrino and positron, meaning that the distribution of the emitted positrons' actual energies occurs along a continuous spectrum, ranging from close to zero, to the isotope's complete decay energy. The average kinetic energies of the emitted positrons for various frequently utilised radionuclides are shown in Table 2.2 (26, 27). Subsequent to the production of a positron, the distance the positron is capable of travelling before most of its kinetic energy is lost as a

result of being scattered to adjoining tissue is relatively small (0.2 – 2 mm for the majority of frequently employed tracers), at which point it is combined with an electron via a process of annihilation. This involves the annihilation of both the electron and positron as well as the conversion of their mass into energy ($E = mc^2$), which takes the form of a photon pair. According to the fundamental laws of physics, energy conservation, linear momentum and the energy of every gamma ray (photon) is 511 keV, which equates to a positron's or electron's mass. The emission of the two produced photons occurs virtually back-to-back (180 ± 0.6 degrees), meaning that two gamma rays travelling in opposite directions have a net momentum of zero. PET imaging has unique tomographic and quantitative imaging characteristics as result of the fact that the gamma pair is emitted in this back-to-back manner. Detection of the annihilation radiations occurs externally and they are utilised for determining the positron emitter's location and quantity. Additionally, PET scanners leverage the fact that the emission occurs back-to-back to 'electronically collimate' the pairs of photons, as well as to ascertain the trajectory of the annihilation process. As event localisation can be achieved without a physical collimator, the sensitivity of PET can be 100 times higher compared with SPECT. The image resolution is also enhanced by the absence of a collimator (24).

Table 2.2 Common positron isotopes characteristics

Positron isotopes	Half-life (min)	Average positron energy (MeV)
¹¹C	20.4	0.3
¹³N	9.96	0.4
¹⁵O	2.07	0.6
¹⁸F	109.7	0.2

The distance that emitted positrons are required to travel prior to the generation of a pair of gamma rays is minimal, which suggests that the site at which the positron is emitted, in other words the positron-labelled molecule site and the site where PET detects the generation of gamma rays differ slightly. The distance travelled by a positron prior to annihilation, known as the 'positron range', is dependent on its preliminary kinetic energy. The positron range and non-collinearity (the gamma ray pair are not precisely back-to-back) of the photon mean that an event cannot be positioned with complete certainty, which fundamentally limits PET resolution. Nevertheless, in standard PET employed for clinical scanning utilising FDG, the primary constraint to the spatial resolution of PET images is derived from the magnitude of the single small crystals utilised in the detection process.

PET detectors are responsible for detecting the annihilation photons emitted by positron-emitting radiotracers administered to patients. The most common type of PET detector is the scintillation detector. Scintillation detectors consist of scintillator crystals coupled with photodetectors, such as photomultiplier tubes (PMTs) or photodiodes. The emitted photons from the annihilation process interact with the scintillator crystal, causing it to emit flashes of light. The photodetectors convert this light into electrical signals that can be processed and analysed to reconstruct the PET image (28, 29). Some modern PET systems employ time-of-flight (TOF) detectors, which measure the time it takes for the annihilation photons to reach the detectors. TOF information allows for more precise localisation of the radiotracer, resulting in enhanced image quality and reduced image noise.

The surrounding detector system is not reached by all the photons that are emitted within the object (body). Interactions can occur between these photons and the bodily tissue, or they merely traverse it with no interactions occurring. Such interactions between photons and body tissue largely manifest as either Compton scattering or photoelectric interaction. Complete absorption of the photons engaged in photoelectric interaction occurs, meaning that they are incapable of reaching the detectors. Compton scattering constitutes the main type of interaction for photons at 511 keV in tissue, which occurs when a photon collides with an electron with low binding energy within an atom's outer shell. The interaction between an electron and a photon causes the latter to lose a portion of its energy and its trajectory is changed, thus causing the number of photons that would have potentially reached the detectors to be reduced. The process by which photon flux is lost as a result of interactions within body (or an object of any kind) is defined as attenuation. In PET systems, the likelihood that a pair of photons produced during a single event (positron annihilation) is capable of surviving the attenuation within the body is not dependent on the annihilation's location on the path that connects the two detectors. Thus, the attenuation effect can potentially be corrected if a transmission scan is performed utilising an external source.

PET exploits the characteristic that the creation of the two photons produced by positron annihilation occurs at the same time; therefore, every other event that does not meet the time-coincidence criterion is rejected. Coincidence occurs when two photons are detected inside a particular time period called the coincidence window, generally a time interval of about 10 ns. After detecting a gamma ray, if no another gamma ray accompanies it inside this interval, the

event is rejected. The origins of coincidence gamma pairs detected in this way are not always from the identical annihilation event. Events determined to be in coincidence are categorised as being true, random, or scatter coincidences. True coincidence events originate from the identical positron annihilation (Figure 2.1, Event 1). Random (also known as accidental) coincidence events are those where the origins of the two photons detected are two distinct annihilation events and are detected to be in coincidence by chance as a result of the coincidence window's finite nature (Figure 2.1, Event 2). Scatter coincidence events originate from an individual annihilation event, although scattering of a single or of both photons within the object occurs (Figure 2.1, Event 3). As a result of this process of scattering, the photon's energy and direction are altered, thus causing the loss of information about its position. It is only possible to detect one of the photon pairs in numerous annihilation events which are defined as 'singles'. Determining whether a given coincidence is a scatter, random or true event is not feasible with the detector. True events are desirable, whereas all other events (scatter and random) should be subject to measurement and estimation followed by removal from overall coincidence events (24). The energy window is typically set to a specific range of energies around the peak energy of the annihilation photons. This range is usually narrow to reduce the influence of scatter and random events, and other sources of noise.

When performing measurements with PET, background noise is generated by scatter and random events. After the image has been reconstructed from the tracer distribution, they cause the background to be generally blurred, which results in a reduction in the quality of the image. The size of scatter and

coincidence events are influenced by how the tracer activity is distributed (within and external to the field of view), the composition of the object being studied and how the PET scanner is designed.

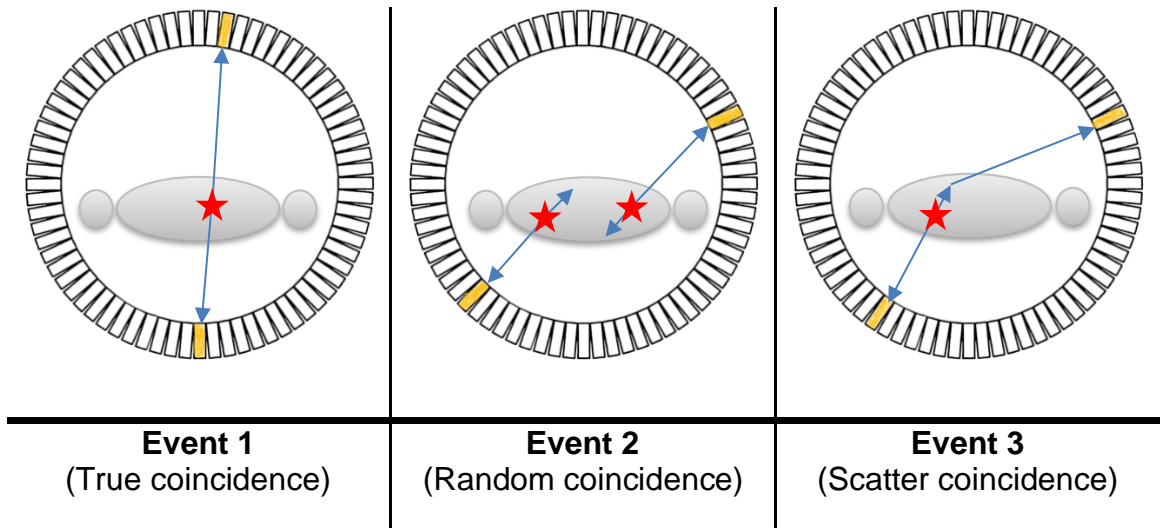


Figure 2.1 Photons coincidence events

2.2.1.2 Image Quantification

Theoretically, when imaging with PET, it is possible to accurately quantify the tracer uptake from the image data as a result of the attenuation-correction ability of this modality. Such information can enhance the ability of the observer to interpret the images. PET imaging using FDG which currently comprises the imaging tracer with the most clinical benefits for detecting cancer as well as many cardiovascular diseases can take advantage of the use of quantitative uptake measures to calculate the probability of a tumour being malignant according to the degree of metabolic activity, as well as to evaluate how the tumour responds to therapy.

Quantitation can be performed using various methods, including blood sampling, complex kinetic analysis with dynamic data acquisition, and visual assessment (30). The basic approach is where the images are visually assessed. FDG uptake is the most frequently assessed quantitative using the standardised uptake value (SUV), in which the fractional FDG uptake by the tumour is estimated (31, 32). This measure is also known as the differential uptake ratio (DUR) or dose uptake ratio. Measurement using SUV is frequently employed for the purpose of determining whether lesions are malignant or benign. The definition of SUV is considered to be the tissue concentration of tracer based on the measurement of a PET camera divided by the activity injected divided by the body weight:

$$SUV (g/ml) = \frac{\text{Tumor activity concentration } \left(\frac{MBq}{ml}\right)}{\text{Injected dose (MBq)}} \times \text{body weight (g)} \quad \text{Equation 2.1}$$

In addition to attenuation-correction, it necessary to precisely determine the dose being administered when using SUV, applying corrections for any residual activity that occurs in the tubing and syringe, while decay correction must be applied to the dose according to the imaging time. In the context of FDG, the uptake of tracer from 30 to 60 minutes subsequent to being injected is normally employed for tissue activity. Nevertheless, for a model independent assessment of glucose metabolism to be achieved, it is necessary to measure the SUV once the FDG tissue concentration has plateaued. The measurement can vary significantly according to how it is implemented exactly at the different clinical sites. In order for the variation to be minimised when making comparisons, identical quality control parameters should be followed when studying a given

site, including the time waited subsequent to injection, the imaging time, and as fasting protocols.

Although the accuracy of SUV quantitation is inferior to that of compartmental analysis, it can be easily implemented and blood sampling is unnecessary, thus giving it increased practicality in clinical settings. Although other quantitation methods offer greater reliability such as kinetic analysis with parameter optimisation, Patlak graphical analysis and simplified kinetic analysis, they still have increased complexity and due to practical limitations they are currently not employed other than for research purposes.

The image quantification affected by a phenomenon called partial volume effect (PVE) where the limited spatial resolution of the imaging system causes the mixing of signal from multiple tissue types within a single voxel. This effect becomes particularly relevant when small structures or lesions are imaged, as their boundaries may not be accurately represented due to the limited resolution. In cardiac PET imaging, PVE can affect the accuracy of quantitative measurements, such as myocardial blood flow, myocardial glucose utilisation or uptake values of radiotracers. The blurring and mixing of signal can lead to underestimation or overestimation of the true activity concentrations in the myocardium or lesions, leading to inaccurate in quantification.

2.2.1.3 Role of PET in Cardiac Sarcoidosis Diagnosis

The glucose analogue [^{18}F]FDG is capable of determining whether active CS is present. If the uptake within the macrophage-dense regions is increased,

[¹⁸F]FDG PET will be capable of visualising inflammation caused by sarcoidosis. Nonetheless, the diagnosis of CS cannot be made purely on the basis of [¹⁸F]FDG uptake, as accurate diagnoses cannot be made of certain regions with low intensity by applying this method (33). Additionally, it is not necessarily the case that regions with increased uptake denote that CS is present due to the fact that inflammation is caused by multiple heart diseases, including ischaemia in coronary artery disease (CAD) patients, as well as systemic rheumatic disorders with cardiac involvement (34, 35). Moreover, inflammation does not manifest in all CS patients; hence, it is not possible to eliminate CS, but the lack of active myocardial inflammation can be documented (36). Inflammation and perfusion can be assessed using staging systems in order to increase the diagnostic significance, although they do not offer outcome or histological validation (37). For [¹⁸F]FDG PET to perform more effectively in diagnosis, it is critical that the usage of glucose by normal cardiomyocytes is suppressed because this can enhance the specificity. Although various methods have been suggested, including adhering to a ketogenic diet (high-fat, low-carb), extended fasting, intravenous heparin and, most commonly, approaches where all these are combined, none is specifically applicable to CS (36). Nevertheless, methods aimed at enhancing the diagnostic performance are not beneficial for as many as 25% of patients, while they have the potential to produce false-positive results (38) because the physiological uptake of the myocardium continues to occur.

The benefits of PET imaging usage for diagnosis purposes have been documented. Patients with renal disorders or intracardiac devices can be exposed to PET without any concerns for their safety. Additionally, total-body

PET imaging can facilitate the process of detecting different kinds of extra-CS (33). For instance, the lymph nodes located next to the lungs are frequently affected as a result of effects on the lungs, which PET can detect. Although the diagnostic benefits of PET imaging for the identification of CS have not precisely been determined, attempts to establish its accuracy have been made by numerous researchers. In a meta-analysis involving 164 patients, rates of 78% and 89% were determined for the pooled specificity and sensitivity of PET in the diagnosis of CS (39). Kim, S.-J. et al. (40) found rates of 84% and 83%, respectively, for pooled sensitivity and specificity in a different meta-analysis involving 17 studies. Nevertheless, it is important to interpret such findings with caution as a result of the limited size of the studies pooled.

2.2.2 Magnetic Resonance Imaging (MR)

2.2.2.1 MR Theory

As a medical imaging approach, MR imaging is non-ionising and non-invasive and draws on the nuclear magnetic resonance (NMR) principle. Dissimilar to other techniques in which radioactive isotopes or X-rays are employed, ionising radiation is not used in MR. Additionally, as a result of its excellent soft tissue contrast, it is more effective than different modalities and is specifically beneficial for diagnosing issues impacting tendons, cartilage, ligaments, and joints.

Independent studies were conducted in which identifying the principle of MR caused by exciting protons immersed in a static magnetic via a radiofrequency field at the resonance frequency (41). Subsequently, it was found that the

excitation of normal tissues and tumours occurred at identical frequencies, but that their relaxation times differed. Such findings laid the foundation for the use of MR for the purpose of diagnosing diseases and monitoring treatment.

MR functions according to the principle of NMR, which manifests as a result of the interaction between nuclei and a powerful magnetic field. It is necessary for these nuclei to have both a spin angular momentum and intrinsic magnetic momentum for this to occur (i.e. a non-zero spin must be possessed) (42). As a result of this characteristic, a nucleus can be positioned within a magnetic field for the purpose of excitation via radio waves, which results in the emission of a radiofrequency signal that can be detected. This property is possessed by the nucleus of a hydrogen atom, and due to the fact that around 70% of the human body is comprised of water (which contains nuclei of hydrogen), hydrogen-based resonance is the basis of MR functioning. Both the spin and charge of the hydrogen nucleus are fixed and the magnetic moment that is produced is oriented in a random direction when no external magnetic field is present, Figure 2.2. When a uniform magnetic field B_0 exists, the precession of the spins around B_0 will occur in a trajectory way that describes a cone as a result of the spin angular momentum, Figure 2.3.

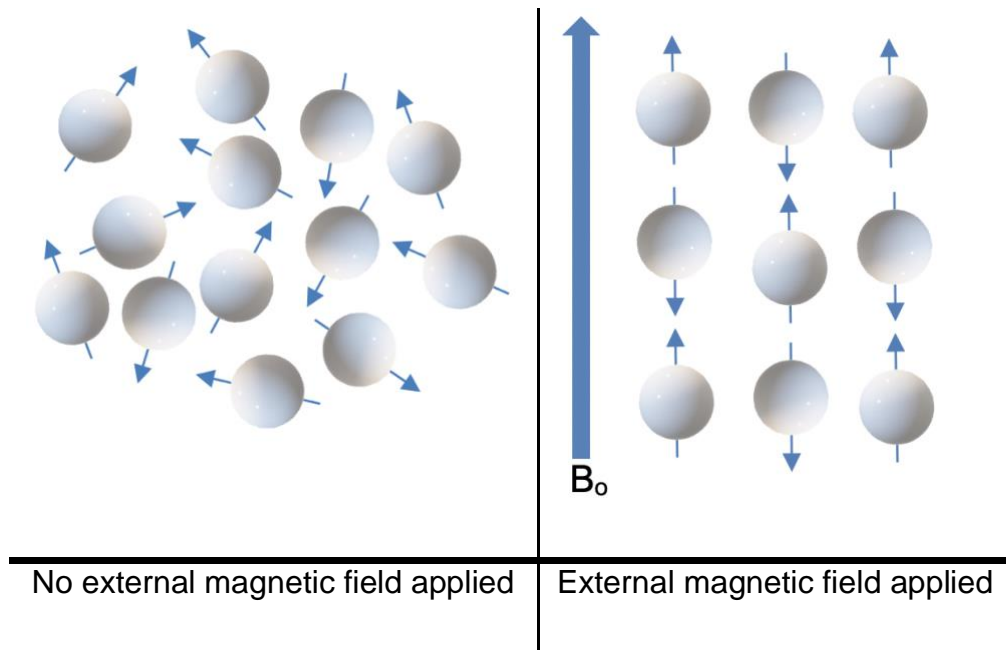


Figure 2.2 The effect of the external magnetic field on the hydrogen nuclei

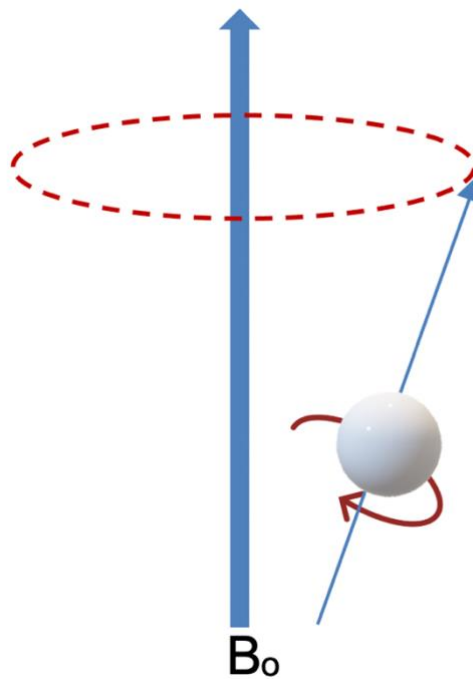


Figure 2.3 The representation of hydrogen nucleus spins in a static magnetic field B_0

Larmor's equation is used to calculate the angular frequency, ω_0 , of the precession:

$$\omega_0 = \gamma |B_0| \quad \text{Equation 2.2}$$

Furthermore, the Bloch equation describes the precession:

$$\frac{d\mu}{dt} = \omega_0 \mu \quad \text{Equation 2.3}$$

Where;

γ denotes the gyromagnetic constant (42.58 MHz/T for hydrogen).

$|B_0|$ denotes the strength of the magnetic field.

μ is the magnetic moment.

As the temperature of the human body is finite, the spin precession becomes anisotropic, thus making it more probable that the orientation of the magnetic moments will occur with reduced magnetic energy parallel with B_0 than increased magnetic energy anti-parallel with B_0 . This results in a net magnetisation, M , which is zero in the absence of an external magnetic field (41).

When there is no radiofrequency (RF) pulse, the alignment of the initial net magnetisation, M_0 , is with the magnetic field B_0 . However, when the application

of an RF pulse occurs at the Larmor frequency, this produces a resonance and causes M_o to tilt in the opposite direction to B_o . The M_o is comprised of two elements: the transverse element, M_{xy} , whose rotation occurs on the xy plane, and the longitudinal element, M_z , which is on the z-axis. As M_{xy} rotates around the z-axis, an electromagnetic field is generated that is able to induce a current through the receiver coil. After amplification, it is possible to detect the generated signal which is defined as the MR signal or free-induction decay.

The application of an RF pulse causes the spin system to remain in a state of excitement; however, upon termination of the RF pulse, energy is lost, thus causing the spin to revert to its state of equilibrium. Such a process is defined as relaxation, while its duration is referred to as the relaxation time (T1 and T2). Relaxation occurs via one of two mechanisms: spin-lattice and spin-spin relaxation. The former comprises the longitudinal relaxation of the M_z element. This is represented by the time constant T1 which is defined as the rate of energy transferral between the nuclear spin system and the adjacent molecules, called the lattice. The M_{xy} component's transverse relaxation is the spin-spin relaxation which is characterised by the time constant T2.

Subsequent to a 90° pulse, precession of the spins occurs in a coherent manner in a uniform direction; however, loss of such coherence progressively occurs as a result of field inhomogeneity, as well as the fact that the spins interact directly with each other. Nevertheless, no energy is transferred to the lattice during this kind of relaxation. Dephasing that occurs as a result of the joint effects of T2 decay and field inhomogeneity is denoted as T2* and the duration is less than T2 (41).

It is important to note that the spin-spin and spin-lattice relaxation are related, as it is impossible for the z-magnetisation to occur if the xy-magnetisation does not decrease in tandem. Therefore, $T_2 \leq T_1$ ($T_2 \approx T_1$ in liquids, but $T_2 \ll T_1$ in solids).

Phase-Sensitive Inversion Recovery (PSIR) is a sequence used in CMR. It is a specialised imaging technique designed to enhance the visualisation of late gadolinium enhancement (LGE) in the myocardium. PSIR sequences are particularly useful for identifying areas of scar tissue or fibrosis in the heart. It involves acquiring images after administering a contrast agent, typically gadolinium-based, to the patient. It utilises an inversion recovery pulse sequence, where the inversion time is carefully chosen to null the signal from normal myocardium while retaining the signal from areas with delayed gadolinium enhancement. By employing the PSIR technique, CMR can provide detailed information about myocardial viability, scar tissue, and areas of cardiac pathology. This valuable imaging tool aids in the diagnosis, assessment, and treatment planning for various cardiovascular conditions.

2.2.2.2 Image Acquisition

In the acquisition process, spatial organisation of the raw data is performed such that a signal can be located in the matrix through the application of a method of spatial encoding grounded on the magnetic field gradient. At the initial step, a slice is selected for imaging. As the Larmor frequency will alter in this direction, it is possible to excite the individual slices through the application

of varying RF pulses, thus instigating resonance. According to the sequence, the selection of slices can be undertaken utilising the X, Y or Z gradients, which denote the sagittal, coronal, and axial slices, respectively (43). Nevertheless, the thickness of the slice is dependent on the amalgamation of two factors: the gradient's steepness or strength, and the bandwidth or frequency range within the RF pulse (44). The slice generated is thicker when the gradient is shallow, whereas it is thinner when the gradient is steep.

After the slice positions and thicknesses have been selected, it is then possible to identify the MR signal's spatial position through the use of spatial encoding. The phase encoding process involves the switching on of the gradient in a specific orthogonal direction, which alters the frequencies of the spins in relation to where they are located on the gradient. Resultantly, the excited spins whose position is higher possess more powerful magnetic energy and gain phase in comparison with spins whose position is lower. This produces a phase shift of the spins, whereas the identification of the individual slices can be made according to their specific phase. In frequency encoding, the application of the gradient occurs in the third direction, which produces a magnetic field where the strength rises in a right to left direction. The frequency variation that is generated results in the left side spin to have slower precession compared with the right side spins. Subsequent to the acquisition of a broad spectrum of frequencies, it is possible to characterise the individual slices according to their individual frequencies. Every volume element (voxels) can be spatially identified by combining the phase and frequency results. Several iterations of the encoding process are performed at various gradient strengths, while the information that contains the MR signal is stored into K-space. The number of

pixels within the image is determined by the number of phase encoding steps. Both the image and resolution quality are improved when the number of encoding steps is increased, although this causes the overall duration of the scan to be extended (43). This is followed by the application of a Fourier transform to each of the raw k-space values, which ultimately produces the final MR image.

2.2.2.3 Role of MR in Cardiac Sarcoidosis Diagnosis

CMR comprises a non-invasive imaging technique that has considerable importance in the diagnosis and screening of CS patients. It is capable of detecting indicators of CS, including perfusion defects, myocardial oedema, and scarring; additionally, it is employed for assessing the function and geometry of the two ventricles. In certain studies, reports indicated that the CMR specificity and sensitivity ranged between 75% and 100% and 76% and 78%, respectively, for the detection of CS (45, 46). It may be possible to evaluate the inflammatory element of CS and oedema through the addition of parameters, including T2-weighted imaging and T2 mapping. In other studies, it was shown that T2-weighted imaging is a feasible option compared with [¹⁸F]FDG PET for inflammation assessment; nevertheless, this method has shortcomings as a result of the reduced signal-to-noise ratio (47, 48). LGE imaging can be used to evaluate myocardial scarring. In comparison with normal myocardium, the extracellular contrast agent gadolinium exhibits slow washout in regions characterised by fibrosis. While LGE can facilitate the identification of CS, according to the distribution and pattern of LGE (45, 49), it could be non-specific. Furthermore, patients diagnosed with advanced renal disease must not be injected with gadolinium contrast (50).

2.2.3 PET/MR Imaging

Several comprehensive reviews and guidelines have extensively covered the various clinical uses of cardiac PET/MR imaging, as evidenced by a range of scholarly articles (51-56). These sources have reported significant advancements in the areas of diagnosis, prognosis, and therapy monitoring for a wide array of cardiac conditions, including myocardial ischemia, sarcoidosis, myocarditis, and cardio-oncology. It is worth noting that fully integrated PET/MR systems represent some of the most intricate and sophisticated devices employed in the field of medical imaging. The development and implementation of such systems have encountered substantial challenges, with the process spanning several decades from initial preclinical stages to their integration into clinical practice (57, 58). PET/MR imaging offers several advantages in cardiology, particularly for the evaluation of cardiac sarcoidosis. Compared to PET/CT, PET/MR imaging significantly reduces radiation exposure to patients. This is especially important for patients who require multiple imaging studies over time, such as monitoring disease progression or treatment response. In addition, MR provides excellent soft tissue contrast, allowing for the detailed assessment of myocardial tissue characteristics. This is particularly beneficial for evaluating myocardial inflammation, fibrosis, and scar tissue associated with cardiac sarcoidosis. However, PET/MR scanners are not as widely available as PET/CT or standalone MR systems. This can restrict access to this imaging modality for some patients and healthcare facilities. In addition, the magnetic field strength and inhomogeneities may influence the positron range leading to reduced PET spatial resolution. Furthermore, the presence of metal implants,

patient motion, and respiratory or cardiac motion can introduce artefacts and affect image quality in PET/MR.

2.2.3.1 Attenuation Correction

Attenuation correction is an important aspect to consider when discussing PET/MR imaging in cardiology, including the evaluation of cardiac sarcoidosis. Attenuation correction is the process of accounting for the attenuation or weakening of PET signals as they pass through different tissues in the body. In PET/CT imaging, attenuation correction is performed using CT data, which provides information about tissue density for accurate correction of PET images.

In PET/MR, attenuation correction becomes more challenging since MR images do not directly provide information about tissue attenuation. However, several approaches have been developed to address this limitation. MR-based attenuation correction methods which use MR images to estimate tissue attenuation by assigning attenuation values to different tissue types. The 3D breath-hold Dixon sequence is an MR imaging technique that allows for the acquisition of multiple sets of images with different echo times. These images are used to estimate tissue properties, such as fat and water content, which can be utilised for various applications, including attenuation correction. These tissue property maps are then used to generate an attenuation map for PET attenuation correction in PET/MR imaging. By providing an estimation of tissue attenuation, the Dixon technique helps improve the accuracy of PET quantification and interpretation in the presence of different tissue types and densities (59). Several studies involving cardiac PET/MR imaging with both

phantoms and cardiovascular patients have demonstrated a notable similarity between PET/CT and PET/MR in terms of SUV values. These findings highlight the validity and reliability of Dixon-based attenuation correction in PET/MR imaging (60-62).

2.2.3.2 Motion Correction

Respiratory, cardiac or patient motion artefacts may occur during image acquisition. These artefacts can lead to image blurring and misalignment, that can directly influence the process of attenuation correction for the corresponding PET images. Subsequently, affecting the accuracy of quantitative analysis. Various respiratory patterns during Dixon and PET acquisitions can result in misalignment between attenuation correction maps and PET images, thereby compromising the accuracy of the correction. PET/MR imaging presents complementary strategies to tackle this challenge, such as prospective or retrospective gating (63-65). In addition, by acquiring images during breath-holding, the 3D Dixon sequence enables the acquisition of images at different points in the respiratory cycle when motion is minimised (64). This reduces the impact of respiratory motion artefacts and improves the alignment of images, resulting in sharper and more accurate images for analysis. To correct motion in attenuation maps, these approaches make use of dedicated MR sequences or PET counts distributions to capture the myocardial motion model during image acquisition. With the integration of motion information, PET/MR imaging can be aligned more accurately and attenuation correction can be performed more accurately.

2.2.3.3 Scatter Correction

In addition to the attenuation of photons, the presence of scattered photons also introduces potential errors in accurately estimating the concentration of activity, thereby impacting the quantification of PET data (66). The PET/MR system faces specific challenges due to its relatively small PET ring diameter, large axial field of view, and wide PET coincidence window. Consequently, the likelihood of random and scatter coincidences in PET/MR is relatively high. To address this issue, PET/MR systems employ narrower energy windows while maintaining comparable energy resolution (59). This approach effectively manages the scattering fraction, indicating that the overall performance of PET detectors within MR scanners remains largely unaffected.

2.3 Radiomic Analysis

Radiomic analysis has emerged as a transformative and highly promising approach in the field of medical imaging, commonly known as radiomic features. In recent years, the field of radiomics has garnered increasing attention and recognition due to its ability to extract a wealth of quantitative data from medical images (67). Unlike traditional qualitative assessments, radiomic analysis delves deep into the pixel-level intricacies of these images, generating an extensive array of quantitative features. This wealth of information encompasses not only fundamental parameters such as size, volume, and intensity but also intricate details regarding texture, shape, and spatial relationships within the imaged regions (68). Such comprehensive data analysis offers significant potential for aiding in the diagnosis, prognosis, and treatment planning of various medical conditions.

By applying advanced computational techniques, radiomics enables the extraction of a multitude of quantitative features from medical images, thereby opening new avenues for improving decision making. The key applications of radiomics in medical imaging include cancer diagnosis and characterisation, treatment response assessment, prognosis and survival prediction, and neurological and cardiovascular disorders. Radiomics plays a pivotal role in oncology compared to other applications (69). It can aid in distinguishing between benign and malignant lesions, assessing tumour heterogeneity, and predicting the aggressiveness of cancer. This not only facilitates early detection but also assists in tailoring personalised treatment plans. While in neuroimaging and cardiovascular imaging, radiomics has the potential to identify disease-specific biomarkers and help with early detection of conditions such as cardiac disorders (70). However, Rainey, C. et al. (71) highlighted the necessity of structured education for healthcare practitioners in artificial intelligence to equip the current and future workforce for the imminent clinical integration of artificial intelligence in healthcare. Some educational institutions have initiated this incorporation to stay updated with the evolving landscape (72).

2.3.1 Factors Affecting Radiomic Analysis

The processes prior to feature extraction play a crucial role in modifying the input image and should not be underestimated. These procedures encompass the image acquisition, encompassing all scanner settings; addressing attenuation and scatter issues; image reconstruction and image segmentation. Often, these steps have been fine-tuned for visual interpretation, which may not

necessarily yield the most favourable radiomic outcomes (73). Nevertheless, the imperative for maintaining consistent methodologies dictates that these visually optimised images must be employed.

2.3.1.1 Image Acquisition and Protocol Variability

The process of acquiring images involves multiple aspects that can introduce variability. For instance, the duration per bed position does have some impact, but it is relatively minor when compared to other factors (74), such as the time between the injection and the actual scan. Specific steps in the image acquisition process may vary based on the application or the manufacturer of the scanning equipment. For instance, longer scan times can potentially reduce image noise, although this may not always be practically feasible in a clinical setting (75).

If the data under study are originated from different scanners/centres, a harmonisation technique must be applied. Harmonisation is the process of standardising or aligning data from different sources or formats to ensure compatibility and consistency. It aims to overcome issues arising from variations in data collection methodologies or measurement scales, which can hinder accurate and reliable analysis. The widely utilised technique of ComBat harmonisation has demonstrated its effectiveness in multiple independent studies (76, 77). It performs retrospective standardisation of radiomic features acquired using diverse protocols. ComBat achieves this by mitigating the centre effect, thereby ensuring consistency across datasets (78).

2.3.1.2 Image Reconstruction

The process of image reconstruction in medical imaging may have an impact on radiomic analysis, influencing the quality, reliability, and interpretability of radiomic features (74, 79, 80). However, some studies showed that radiomic features are robust to different reconstruction algorithms (81). Image reconstruction transforms raw data acquired from imaging modalities into clinically useful images (82). However, it should take into account many considerations that may be affected by image reconstruction and thereby radiomic analysis. Firstly, the image reconstruction affects the spatial resolution, which affects the ability to discern small structures or subtle variations within the imaged region (83). Higher spatial resolution can lead to more detailed radiomic features, potentially revealing important information for diagnosis and treatment planning. In addition, the reconstruction process can introduce or exacerbate noise and imaging artefacts (84, 85). High noise levels can compromise the precision of radiomic features, while artefacts may lead to incorrect or misleading findings. Image denoising techniques and artefact correction methods are often used to mitigate these issues (86). The stability of radiomic features is a key consideration. In some cases, different reconstruction algorithms or settings can yield substantially different radiomic results. This can impact the reproducibility and reliability of radiomic analysis, especially in multi-centre studies. Furthermore, one of the strengths of radiomics is the ability to assess tissue heterogeneity. Image reconstruction can influence the quantification of heterogeneity-related features (87), making it vital to standardise reconstruction procedures for meaningful comparisons. Iterative reconstruction algorithms, which are increasingly used in modern imaging, can

enhance radiomic analysis by reducing noise and improving feature stability (88).

In PET/MR scanners, while Fourier transform-based methods are commonly used for MR image reconstruction, specialised algorithms such as iterative techniques are typically employed for PET image reconstruction. In this thesis, PET images were reconstructed using one of the iterative reconstruction approaches called ordinary Poisson ordered subset expectation maximisation (OP-OSEM). This algorithm is an extension of the OSEM algorithm, which is a widely adopted iterative reconstruction method in PET. In OP-OSEM, the algorithm iteratively updates the estimated activity distribution in the image space based on the acquired projection data, following the principles of maximum likelihood estimation. It takes into account the Poisson nature of PET data, which arises from the random nature of positron annihilation events. OP-OSEM can indirectly mitigate the impact of PVE to some extent through the iterative reconstruction process. By iteratively updating the estimated activity distribution based on the acquired projection data, OP-OSEM can help recover sharper boundaries and enhance spatial details, thereby potentially reducing the blurring caused by PVE. However, additional techniques specifically targeting PVE, such as partial volume correction algorithms, may be necessary.

2.3.1.3 Image Segmentation Accuracy

When employing radiomic analysis, it is customary to delineate a region of interest (ROI) for examination. While it is possible to analyse the entire imaging volume, this approach is often computationally demanding. Given that

segmentation defines the region to be studied in radiomics, it is imperative that the segmentation method be both precise and consistently replicable. In a PET/CT study using a phantom was discovered that altering the segmentation method resulted in stability for only 20% of the features (89). Similarly, another study demonstrated that whether an automatic or manual segmentation method was used, only 13% of the features remained unaffected (90).

Numerous considerations come into play when choosing between manual and automated segmentation approaches. Manual methods, while potentially more accurate, tend to be slow, susceptible to inter/intra-observer variation, and demand substantial expert input. Despite these drawbacks, manual segmentation can prove superior in scenarios involving low resolution or non-contrast enhanced imaging. On the other hand, automated segmentation is generally faster and more reproducible, facilitating the analysis of extensive datasets and yielding more substantial results (91). However, challenges may arise when dealing with images lacking high resolution or contrast.

2.3.2 Feature Selection and Analysis

Feature selection is a crucial step in machine learning and data analysis, aimed at identifying the most relevant and informative features from a given dataset and eliminating noise or irrelevant information. The goal is to reduce the dimensionality of the data while retaining the most discriminative features, which can lead to improved computational efficiency and reduced model complexity, thereby improving model performance and interpretability. In addition, feature selection can mitigate the risk of overfitting by reducing the

chances of the model learning from random or spurious correlations within the data as well as minimise the risk of Type I errors (i.e., rejecting a null hypothesis when it is true) (92).

There are various methods and techniques available for feature selection, ranging from simple statistical measures to more advanced algorithms. These methods can be broadly categorised into filter, wrapper, and embedded approaches. Filter methods assess the relevance of features based on statistical measures such as correlation, mutual information, or chi-square tests. They independently evaluate each feature without considering the predictive model. Filter methods are computationally efficient and can be easily applied to large datasets. However, they may overlook interactions between features and rely solely on statistical properties. Unlike filter methods, wrapper methods assess feature subsets by using a specific machine learning algorithm as a black box. They create multiple feature subsets, train and evaluate the model on each subset, and select the one that yields the best performance. Wrapper methods consider the interaction between features and the performance of the specific model being used. However, they tend to be computationally expensive and may be prone to overfitting if the dataset is small. Embedded methods incorporate feature selection within the model training process itself. They typically utilise regularisation techniques, such as Lasso or Ridge regression, which impose penalties on the model coefficients, effectively shrinking less important features to zero. Embedded methods strike a balance between the efficiency of filter methods and the effectiveness of wrapper methods. They perform feature selection during model training, reducing both computation time and the risk of overfitting (92).

The choice of feature selection method depends on various factors, such as the dataset size, dimensionality, and the specific problem. It is essential to carefully evaluate the trade-offs between computational cost, model performance, and interpretability when selecting a feature selection technique. Furthermore, it is crucial to validate the selected features and assess their stability and generalisability using appropriate evaluation metrics and validation strategies. Cross-validation and independent test datasets can help ensure the robustness and reliability of the selected features, improving the credibility of the results. However, it is suggested not to split small datasets into separate training and testing sets. Instead, it is advised to utilise the entire small dataset for model development and training with some form of cross-validation. Cross-validation involves splitting the dataset into multiple subsets, training the model on one subset, and evaluating its performance on the remaining subsets (93). This helps estimate the generalisability of the selected features and reduce the risk of selecting features that are specific to the training set.

Overfitting and Type I error are common concern in feature selection. To address this, Bonferroni correction, dimensionality reduction techniques and cross-validation techniques are often employed. The Bonferroni correction technique is a statistical method used to adjust the significance level or p -values in hypothesis testing when multiple comparisons are conducted simultaneously. When conducting multiple statistical tests, the probability of obtaining a significant result by chance alone increases with the number of comparisons. This inflation of the Type I error rate can lead to a higher likelihood of false positives. To apply the Bonferroni correction, the desired significance level is

divided by the number of comparisons being made. This adjusted significance level is then used as the new threshold for determining statistical significance for each individual test. However, it is important to note that the Bonferroni correction is conservative, meaning that it can be overly stringent and increase the risk of Type II errors (i.e., failing to reject a null hypothesis when it is false) (94). This is because the correction adjusts for multiple comparisons without considering any potential dependencies or correlations between the tests.

Principal component analysis (PCA) is considered a most common dimensionality reduction method. It is commonly used to reduce the dimensionality of a dataset while retaining as much information as possible. PCA works by transforming the original features into a new set of uncorrelated variables called principal components. These principal components are linear combinations of the original features, and they are ordered in such a way that the first component captures the maximum variance in the data, the second component captures the second highest variance, and so on (95). The reduced-dimensional representation can be used as input to other machine learning algorithms.

After reducing the radiomic feature dataset, the next step involves evaluating the diagnostic and predictive value of the remaining extracted features. This evaluation might be conducted alongside the feature selection process or separately. Traditionally, statistical analysis techniques, such as the Mann Whitney U test, correlation test and Cox Regression, have been employed for this purpose. However, there has been a rise in popularity of machine learning classifiers for these tasks like linear or logistic regression. These classifiers,

along with others, have gained widespread acceptance in constructing diagnostic or predictive models. Machine learning classifiers offer a versatile and powerful toolset for analysing radiomic features and their relationship with clinical data. They encompass regression analysis, clustering, and decision tree-based approaches, enabling comprehensive exploration and interpretation of the data.

Various evaluation metrics are used to assess the importance of features and machine learning classifiers during and after the selection process. Area under the curve (AUC) and accuracy are both evaluation metrics commonly used in machine learning, particularly for classification tasks. However, they measure different aspects of model performance. AUC refers to the area under the receiver operating characteristic (ROC) curve. The ROC curve is a graphical representation of the trade-off between the true positive rate (sensitivity) and the false positive rate ($1 - \text{specificity}$) as the classification threshold is varied. AUC is particularly useful when dealing with imbalanced datasets. On the other hand, accuracy is a straightforward measure that calculates the proportion of correctly classified instances out of the total number of instances. However, accuracy can be misleading in situations where the class distribution is imbalanced. In cases where one class dominates the dataset, a model that predicts the majority class for all instances can achieve high accuracy despite not capturing the minority class well. Therefore, accuracy may not adequately reflect the model's performance in situations with imbalanced classes (96). Correlation coefficients, p -values and feature importance scores are other examples of evaluation metrics. In this thesis, AUC was used as a ranking metric while accuracy as a confirming metric.

2.4 Project Justification

The independent use of [¹⁸F]FDG PET and LGE-CMR to evaluate, respectively, myocardial inflammation and fibrosis as markers of CS may fail to identify isolated CS with high specificity and sensitivity. Focal myocardial uptake of [¹⁸F]FDG, highly indicative of active CS, can be confounded by high physiological [¹⁸F]FDG background signal from healthy cardiomyocytes. LGE-CMR alone cannot distinguish acute from chronic scarring, indicating, respectively, active or quiescent disease.

The aim of this project was to increase the diagnostic efficiency of CS by using [¹⁸F]FDG PET and LGE-CMR. The data were analysed using advanced radiomic features analysis. In particular, it was hypothesised that the combined use of [¹⁸F]FDG PET and LGE-CMR will support and confirm each other's findings and contribute to a more accurate assessment and diagnosis compared to the use of either technique alone.

References

1. Birnie, D.H., Nery, P.B., Ha, A.C. and Beanlands, R.S. Cardiac sarcoidosis. *Journal of the American College of Cardiology*. 2016, **68**(4), pp.411-421.
2. Hulten, E., Aslam, S., Osborne, M., Abbasi, S., Bittencourt, M.S. and Blankstein, R. Cardiac sarcoidosis—state of the art review. *Cardiovascular diagnosis and therapy*. 2016, **6**(1), p.50.
3. Ginelliová, A., Farkaš, D., Iannaccone, S.F. and Vyhnálková, V. Sudden unexpected death due to severe pulmonary and cardiac sarcoidosis. *Forensic Science, Medicine, and Pathology*. 2016, **12**(3), pp.319-323.
4. Petek, B.J., Rosenthal, D.G., Patton, K.K., Behnia, S., Keller, J.M., Collins, B.F., Cheng, R.K., Ho, L.A., Bravo, P.E. and Mikacenic, C. Cardiac sarcoidosis: diagnosis confirmation by bronchoalveolar lavage and lung biopsy. *Respiratory Medicine*. 2018, **144**, pp.S13-S19.
5. Deng, J.C., Baughman, R.P. and Lynch Iii, J.P. Cardiac involvement in sarcoidosis. In: *Seminars in respiratory and critical care medicine*: Copyright© 2002 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, 2002, pp.513-528.
6. Hussain, K. and Shetty, M. Cardiac sarcoidosis. 2022.
7. Iwai, K., Tachibana, T., Takemura, T., Matsui, Y., Kitalchi, M. and Kawabata, Y. Pathological studies on sarcoidosis autopsy. I. Epidemiological features of 320 cases in Japan. *Pathology International*. 1993, **43**(7-8), pp.372-376.
8. Perry, A. and Vuitch, F. Causes of death in patients with sarcoidosis. A morphologic study of 38 autopsies with clinicopathologic correlations. *Archives of pathology & laboratory medicine*. 1995, **119**(2), p.167.
9. Kim, J.S., Judson, M.A., Donnino, R., Gold, M., Cooper Jr, L.T., Prystowsky, E.N. and Prystowsky, S. Cardiac sarcoidosis. *American heart journal*. 2009, **157**(1), pp.9-21.

10. Doughan, A.R. and Williams, B.R. Cardiac sarcoidosis. *Heart*. 2006, **92**(2), pp.282-288.
11. Roberts, W.C., McAllister Jr, H.A. and Ferrans, V.J. Sarcoidosis of the heart: a clinicopathologic study of 35 necropsy patients (group I) and review of 78 previously described necropsy patients (group II). *The American journal of medicine*. 1977, **63**(1), pp.86-108.
12. Matsui, Y., Iwai, K., Tachibana, T., Fruie, T., Shigematsu, N., Izumi, T., Homma, A.H., Mikami, R., Hongo, O. and Hiraga, Y. Clinicopathological study on fatal myocardial sarcoidosis. *Annals of the New York Academy of Sciences*. 1976, **278**(1), pp.455-469.
13. Sekiguchi, M., NUMAO, Y., IMAI, M., FURUIE, T. and MIKAMI, R. Clinical and histopathological profile of sarcoidosis of the heart and acute idiopathic myocarditis concepts through a study employing endomyocardial biopsy I. sarcoidosis: symposium on secondary myocardial disease. *Japanese circulation journal*. 1980, **44**(4), pp.249-263.
14. Yazaki, Y., Isobe, M., Hiroe, M., Morimoto, S.-i., Hiramitsu, S., Nakano, T., Izumi, T., Sekiguchi, M. and Group, C.J.H.S. Prognostic determinants of long-term survival in Japanese patients with cardiac sarcoidosis treated with prednisone. *The American journal of cardiology*. 2001, **88**(9), pp.1006-1010.
15. Sekhri, V., Sanal, S., DeLorenzo, L.J., Aronow, W.S. and Maguire, G.P. Cardiac sarcoidosis: a comprehensive review. *Archives of medical science: AMS*. 2011, **7**(4), p.546.
16. Shorr, A.F., Davies, D.B. and Nathan, S.D. Predicting mortality in patients with sarcoidosis awaiting lung transplantation. *Chest*. 2003, **124**(3), pp.922-928.
17. Freeman, A.M., Curran-Everett, D., Weinberger, H.D., Fenster, B.E., Buckner, J.K., Gottschall, E.B., Sauer, W.H., Maier, L.A. and Hamzeh,

- N.Y. Predictors of cardiac sarcoidosis using commonly available cardiac studies. *The American journal of cardiology*. 2013, **112**(2), pp.280-285.
18. Hiraga, H. The guides for the diagnosis of cardiac sarcoidosis. *Report of the Japanese Research Committee for Diffuse Lung Disease of Japan Ministry Welfare*. 1993, pp.23-24.
 19. Hiraga, H., Yuwai, K. and Hiroe, M. Diagnostic standard and guidelines for sarcoidosis. *Jpn J Sarcoidosis Granulomatous Disord*. 2007, **27**, pp.89-102.
 20. Birnie, D.H., Sauer, W.H., Bogun, F., Cooper, J.M., Culver, D.A., Duvernoy, C.S., Judson, M.A., Kron, J., Mehta, D. and Nielsen, J.C. HRS expert consensus statement on the diagnosis and management of arrhythmias associated with cardiac sarcoidosis. *Heart rhythm*. 2014, **11**(7), pp.1304-1323.
 21. Terasaki, F. and Yoshinaga, K. New guidelines for diagnosis of cardiac sarcoidosis in Japan. *Annals of Nuclear Cardiology*. 2017, **3**(1), pp.42-45.
 22. Mehta, D., Lubitz, S.A., Frankel, Z., Wisnivesky, J.P., Einstein, A.J., Goldman, M., Machac, J. and Teirstein, A. Cardiac involvement in patients with sarcoidosis: diagnostic and prognostic value of outpatient testing. *Chest*. 2008, **133**(6), pp.1426-1435.
 23. Greulich, S., Deluigi, C.C., Gloekler, S., Wahl, A., Zürn, C., Kramer, U., Nothnagel, D., Bültel, H., Schumm, J. and Grün, S. CMR imaging predicts death and other adverse events in suspected cardiac sarcoidosis. *JACC: Cardiovascular Imaging*. 2013, **6**(4), pp.501-511.
 24. Kim, E.E., Lee, M.-C., Inoue, T. and Wong, W.-H. *Clinical PET: principles and applications*. Springer Science & Business Media, 2013.
 25. Fowler, J. Positron emitter-labeled compounds: priorities and problems. *Positron Emission Tomography and Autoradiography: Principles and Application for the Brain and Heart*. 1986, pp.391-450.

26. Derenzo, S.E. Precision measurement of annihilation point spread distributions for medically important positron emitters. 1979.
27. Phelps, M.E., Mazziotta, J. and Schelbert, H.R. Positron emission tomography and autoradiography: principles and applications for the brain and heart. 1985.
28. Berger, A. How does it work?: Positron emission tomography. *BMJ: British Medical Journal*. 2003, **326**(7404), p.1449.
29. Humm, J.L., Rosenfeld, A. and Del Guerra, A. From PET detectors to PET scanners. *European journal of nuclear medicine and molecular imaging*. 2003, **30**, pp.1574-1597.
30. Graham, M., Peterson, L. and Hayward, R. Comparison of simplified quantitative analyses of FDG uptake. *Nuclear medicine and biology*. 2000, **27**(7), pp.647-655.
31. Hamberg, L.M., Hunter, G.J., Alpert, N.M., Choi, N.C., Babich, J.W. and Fischman, A.J. The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification? *Journal of Nuclear Medicine*. 1994, **35**(8), pp.1308-1312.
32. Keyes, J.W. SUV: standard uptake or silly useless value? *Journal of Nuclear Medicine*. 1995, **36**(10), pp.1836-1839.
33. Ramirez, R., Trivieri, M., Fayad, Z.A., Ahmadi, A., Narula, J. and Argulian, E. Advanced imaging in cardiac sarcoidosis. *Journal of nuclear medicine*. 2019, **60**(7), pp.892-898.
34. Sarikaya, I. Cardiac applications of PET. *Nuclear medicine communications*. 2015, **36**(10), pp.971-985.
35. Zhuang, H. and Codreanu, I. Growing applications of FDG PET-CT imaging in non-oncologic conditions. *Journal of biomedical research*. 2015, **29**(3), p.189.
36. Chareonthitawee, P., Beanlands, R.S., Chen, W., Dorbala, S., Miller, E.J., Murthy, V.L., Birnie, D.H., Chen, E.S., Cooper, L.T. and Tung, R.H. Joint SNMMI–ASNC expert consensus document on the role of ¹⁸F-FDG

- PET/CT in cardiac sarcoid detection and therapy monitoring. *Journal of nuclear medicine*. 2017, **58**(8), pp.1341-1353.
37. Slart, R.H., Glaudemans, A.W., Lancellotti, P., Hyafil, F., Blankstein, R., Schwartz, R.G., Jaber, W.A., Russell, R., Gimelli, A. and Rouzet, F. A joint procedural position statement on imaging in cardiac sarcoidosis: from the Cardiovascular and Inflammation & Infection Committees of the European Association of Nuclear Medicine, the European Association of Cardiovascular Imaging, and the American Society of Nuclear Cardiology. *European Heart Journal-Cardiovascular Imaging*. 2017, **18**(10), pp.1073-1089.
 38. Osborne, M.T., Hulten, E.A., Murthy, V.L., Skali, H., Taqueti, V.R., Dorbala, S., DiCarli, M.F. and Blankstein, R. Patient preparation for cardiac fluorine-18 fluorodeoxyglucose positron emission tomography imaging of inflammation. *Journal of Nuclear Cardiology*. 2017, **24**(1), pp.86-99.
 39. Youssef, G., Leung, E., Mylonas, I., Nery, P., Williams, K., Wisenberg, G., Gulenchyn, K.Y., deKemp, R.A., DaSilva, J. and Birnie, D. The use of 18F-FDG PET in the diagnosis of cardiac sarcoidosis: a systematic review and metaanalysis including the Ontario experience. *Journal of nuclear medicine*. 2012, **53**(2), pp.241-248.
 40. Kim, S.-J., Pak, K. and Kim, K. Diagnostic performance of F-18 FDG PET for detection of cardiac sarcoidosis; A systematic review and meta-analysis. *Journal of Nuclear Cardiology*. 2019, pp.1-13.
 41. McRobbie, D.W., Moore, E.A., Graves, M.J. and Prince, M.R. *MRI from Picture to Proton*. Cambridge university press, 2017.
 42. Lipton, M.L. *Totally accessible MRI: a user's guide to principles, technology, and applications*. Springer Science & Business Media, 2010.
 43. Weishaupt, D., Köchli, V.D., Marincek, B., Froehlich, J.M., Nanz, D. and Pruessmann, K.P. *How does MRI work?: an introduction to the physics and function of magnetic resonance imaging*. Springer, 2006.

44. Sprawls, P. *Magnetic resonance imaging: principles, methods, and techniques*. Medical Physics Publishing Madison, 2000.
45. Smedema, J.-P., Snoep, G., van Kroonenburgh, M.P., van Geuns, R.-J., Dassen, W.R., Gorgels, A.P. and Crijns, H.J. Evaluation of the accuracy of gadolinium-enhanced cardiovascular magnetic resonance in the diagnosis of cardiac sarcoidosis. *Journal of the American College of Cardiology*. 2005, **45**(10), pp.1683-1690.
46. Ohira, H., Tsujino, I., Ishimaru, S., Oyama, N., Takei, T., Tsukamoto, E., Miura, M., Sakaue, S., Tamaki, N. and Nishimura, M. Myocardial imaging with 18 F-fluoro-2-deoxyglucose positron emission tomography and magnetic resonance imaging in sarcoidosis. *European Journal of Nuclear Medicine and Molecular Imaging*. 2008, **35**(5), pp.933-941.
47. Amano, Y., Tachi, M., Tani, H., Mizuno, K., Kobayashi, Y. and Kumita, S. T2-weighted cardiac magnetic resonance imaging of edema in myocardial diseases. *The Scientific World Journal*. 2012, **2012**.
48. Crouser, E.D., Ono, C., Tran, T., He, X. and Raman, S.V. Improved detection of cardiac sarcoidosis using magnetic resonance with myocardial T2 mapping. *American journal of respiratory and critical care medicine*. 2014, **189**(1), pp.109-112.
49. Ichinose, A., Otani, H., Oikawa, M., Takase, K., Saito, H., Shimokawa, H. and Takahashi, S. MRI of cardiac sarcoidosis: basal and subepicardial localization of myocardial lesions and their effect on left ventricular function. *American Journal of Roentgenology*. 2008, **191**(3), pp.862-869.
50. Russo, R.J., Costa, H.S., Silva, P.D., Anderson, J.L., Arshad, A., Biederman, R.W., Boyle, N.G., Frabizzio, J.V., Birgersdotter-Green, U. and Higgins, S.L. Assessing the risks associated with MRI in patients with a pacemaker or defibrillator. *New England Journal of Medicine*. 2017, **376**(8), pp.755-764.

51. Torigian, D.A., Zaidi, H., Kwee, T.C., Saboury, B., Udupa, J.K., Cho, Z.-H. and Alavi, A. PET/MR imaging: technical aspects and potential clinical applications. *Radiology*. 2013, **267**(1), pp.26-44.
52. Quick, H.H. Integrated PET/Mr. *Journal of magnetic resonance imaging*. 2014, **39**(2), pp.243-258.
53. Bergquist, P.J., Chung, M.S., Jones, A., Ahlman, M.A., White, C.S. and Jeudy, J. Cardiac applications of PET-MR. *Current cardiology reports*. 2017, **19**, pp.1-10.
54. Robson, P.M., Dey, D., Newby, D.E., Berman, D., Li, D., Fayad, Z.A. and Dweck, M.R. MR/PET imaging of the cardiovascular system. *JACC: Cardiovascular Imaging*. 2017, **10**(10 Part A), pp.1165-1179.
55. Nensa, F., Bamberg, F., Rischpler, C., Menezes, L., Poeppel, T.D., La Fougère, C., Beitzke, D., Rasul, S., Loewe, C. and Nikolaou, K. Hybrid cardiac imaging using PET/MRI: a joint position statement by the European Society of Cardiovascular Radiology (ESCR) and the European Association of Nuclear Medicine (EANM). *European Radiology*. 2018, **28**, pp.4086-4101.
56. Cardoso, R. and Leucker, T.M. Applications of PET-MR imaging in cardiovascular disorders. *PET clinics*. 2020, **15**(4), pp.509-520.
57. Buchanan, M., Marsden, P., Mielke, C. and Garlick, P. A system to obtain radiotracer uptake data simultaneously with NMR spectra in a high field magnet. *IEEE Transactions on Nuclear Science*. 1996, **43**(3), pp.2044-2048.
58. Garlick, P.B., Marsden, P.K., Cave, A.C., Parkes, H.G., Slates, R., Shao, Y., Silverman, R.W. and Cherry, S.R. PET and NMR dual acquisition (PANDA): applications to isolated, perfused rat hearts. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*. 1997, **10**(3), pp.138-142.
59. Bogdanovic, B., Solari, E.L., Asiares, A.V., McIntosh, L., van Marwick, S., Schachoff, S. and Nekolla, S.G. PET/MR technology: advancement and

- challenges. In: *Seminars in Nuclear Medicine*: Elsevier, 2022, pp.340-355.
60. Lau, J.M., Laforest, R., Sotoudeh, H., Nie, X., Sharma, S., McConathy, J., Novak, E., Priatna, A., Gropler, R. and Woodard, P. Evaluation of attenuation correction in cardiac PET using PET/MR. *Journal of Nuclear Cardiology*. 2017, **24**, pp.839-846.
61. Oldan, J.D., Shah, S.N., Brunken, R.C., DiFilippo, F.P., Obuchowski, N.A. and Bolen, M.A. Do myocardial PET–MR and PET–CT FDG images provide comparable information? *Journal of Nuclear Cardiology*. 2016, **23**, pp.1102-1109.
62. Vontobel, J., Liga, R., Possner, M., Clerc, O.F., Mikulicic, F., Veit-Haibach, P., Ter Voert, E.E., Fuchs, T.A., Stehli, J. and Pazhenkottil, A.P. MR-based attenuation correction for cardiac FDG PET on a hybrid PET/MRI scanner: comparison with standard CT attenuation correction. *European journal of nuclear medicine and molecular imaging*. 2015, **42**, pp.1574-1580.
63. Küstner, T., Schwartz, M., Martirosian, P., Gatidis, S., Seith, F., Gilliam, C., Blu, T., Fayad, H., Visvikis, D. and Schick, F. MR-based respiratory and cardiac motion correction for PET imaging. *Medical image analysis*. 2017, **42**, pp.129-144.
64. Kolbitsch, C., Neji, R., Fenchel, M., Schuh, A., Mallia, A., Marsden, P. and Schaeffter, T. Joint cardiac and respiratory motion estimation for motion-corrected cardiac PET-MR. *Physics in Medicine & Biology*. 2018, **64**(1), p.015007.
65. Munoz, C., Neji, R., Kunze, K.P., Nekolla, S.G., Botnar, R.M. and Prieto, C. Respiratory-and cardiac motion-corrected simultaneous whole-heart PET and dual phase coronary MR angiography. *Magnetic Resonance in Medicine*. 2019, **81**(3), pp.1671-1684.

66. Zaidi, H. and Karakatsanis, N. Towards enhanced PET quantification in clinical oncology. *The British Journal of Radiology*. 2017, **91**(1081), p.20170508.
67. Shur, J.D., Doran, S.J., Kumar, S., Ap Dafydd, D., Downey, K., O'Connor, J.P., Papanikolaou, N., Messiou, C., Koh, D.-M. and Orton, M.R. Radiomics in oncology: a practical guide. *Radiographics*. 2021, **41**(6), pp.1717-1732.
68. Rizzo, S., Botta, F., Raimondi, S., Origgi, D., Fanciullo, C., Morganti, A.G. and Bellomi, M. Radiomics: the facts and the challenges of image analysis. *European radiology experimental*. 2018, **2**(1), pp.1-8.
69. Gillies, R.J., Kinahan, P.E. and Hricak, H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016, **278**(2), pp.563-577.
70. Polidori, T., De Santis, D., Rucci, C., Tremamunno, G., Piccinni, G., Pugliese, L., Zerunian, M., Guido, G., Pucciarelli, F. and Bracci, B. Radiomics applications in cardiac imaging: a comprehensive review. *La radiologia medica*. 2023, pp.1-12.
71. Rainey, C., O'Regan, T., Matthew, J., Skelton, E., Woznitza, N., Chu, K.-Y., Goodman, S., McConnell, J., Hughes, C., Bond, R., McFadden, S. and Malamateniou, C. Beauty is in the AI of the beholder: are we ready for the clinical integration of artificial intelligence in radiography? An exploratory analysis of perceived AI knowledge, skills, confidence, and education perspectives of UK radiographers. *Frontiers in Digital Health*. 2021, **3**, p.739327.
72. Malamateniou, C., Knapp, K., Pergola, M., Woznitza, N. and Hardy, M. Artificial intelligence in radiography: where are we now and what does the future hold? *Radiography*. 2021, **27**, pp.S58-S62.
73. Hatt, M., Le Rest, C.C., Tixier, F., Badic, B., Schick, U. and Visvikis, D. Radiomics: data are also images. *Journal of Nuclear Medicine*. 2019, **60**(Supplement 2), pp.38S-44S.

74. Shiri, I., Rahmim, A., Ghaffarian, P., Geramifar, P., Abdollahi, H. and Bitarafan-Rajabi, A. The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies. *European Radiology*. 2017, **27**, pp.4498-4509.
75. Pfaehler, E., Beukinga, R.J., de Jong, J.R., Slart, R.H., Slump, C.H., Dierckx, R.A. and Boellaard, R. Repeatability of 18F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. *Medical physics*. 2019, **46**(2), pp.665-678.
76. Da-Ano, R., Masson, I., Lucia, F., Doré, M., Robin, P., Alfieri, J., Rousseau, C., Mervoyer, A., Reinhold, C. and Castelli, J. Performance comparison of modified ComBat for harmonization of radiomic features for multicenter studies. *Scientific reports*. 2020, **10**(1), p.10248.
77. Leithner, D., Nevin, R.B., Gibbs, P., Weber, M., Otazo, R., Vargas, H.A. and Mayerhoefer, M.E. ComBat harmonization for MRI radiomics: impact on nonbinary tissue classification by machine learning. *Investigative Radiology*. 2023, p.10.1097.
78. Papadimitroulas, P., Brocki, L., Chung, N.C., Marchadour, W., Vermet, F., Gaubert, L., Eleftheriadis, V., Plachouris, D., Visvikis, D. and Kagadis, G.C. Artificial intelligence: Deep learning in oncological radiomics and challenges of interpretability and data harmonization. *Physica Medica*. 2021, **83**, pp.108-121.
79. Shiri, I., Rahmim, A., Abdollahi, H., Ghafarian, P., Bitarafan, A., Ay, M. and BakhshaieshKaram, M. Radiomics texture features variability and reproducibility in advance image reconstruction setting of oncological PET/CT. In: *European journal of nuclear medicine and molecular imaging*: SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA, 2016, pp.S150-S150.
80. Van Velden, F.H., Kramer, G.M., Frings, V., Nissen, I.A., Mulder, E.R., de Langen, A.J., Hoekstra, O.S., Smit, E.F. and Boellaard, R.

- Repeatability of radiomic features in non-small-cell lung cancer [18 F] FDG-PET/CT studies: impact of reconstruction and delineation. *Molecular imaging and biology*. 2016, **18**, pp.788-795.
81. Kolossváry, M., Szilveszter, B., Karády, J., Drobni, Z.D., Merkely, B. and Maurovich-Horvat, P. Effect of image reconstruction algorithms on volumetric and radiomic parameters of coronary plaques. *Journal of cardiovascular computed tomography*. 2019, **13**(6), pp.325-330.
 82. Dhawan, A.P. *Medical image analysis*. John Wiley & Sons, 2011.
 83. Nieman, K., Gaemperli, O., Lancellotti, P. and Plein, S. Advanced cardiac imaging. In: *Advanced Cardiac Imaging*. Elsevier, 2015, pp.1-11.
 84. Razifar, P. *Novel approaches for application of principal component analysis on dynamic PET images for improvement of image quality and clinical diagnosis*. thesis, Acta Universitatis Upsaliensis, 2005.
 85. Teymurazyan, A., Riauka, T., Jans, H.-S. and Robinson, D. Properties of noise in positron emission tomography images reconstructed with filtered-backprojection and row-action maximum likelihood algorithm. *Journal of digital imaging*. 2013, **26**, pp.447-456.
 86. Sagheer, S.V.M. and George, S.N. A review on medical image denoising algorithms. *Biomedical signal processing and control*. 2020, **61**, p.102036.
 87. Robins, M., Solomon, J., Hoye, J., Abadi, E., Marin, D. and Samei, E. How reliable are texture measurements? In: *Medical Imaging 2018: Physics of Medical Imaging*: SPIE, 2018, pp.991-1003.
 88. Riddell, C., Carson, R.E., Carrasquillo, J.A., Libutti, S.K., Danforth, D.N., Whatley, M. and Bacharach, S.L. Noise reduction in oncology FDG PET images by iterative reconstruction: a quantitative assessment. *Journal of Nuclear Medicine*. 2001, **42**(9), pp.1316-1323.
 89. Gallivanone, F., Interlenghi, M., D'Ambrosio, D., Trifiro, G. and Castiglioni, I. Parameters influencing PET imaging features: a phantom

- study with irregular and heterogeneous synthetic lesions. *Contrast media & molecular imaging*. 2018, **2018**.
90. Altazi, B.A., Zhang, G.G., Fernandez, D.C., Montejo, M.E., Hunt, D., Werner, J., Biagioli, M.C. and Moros, E.G. Reproducibility of F18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms. *Journal of applied clinical medical physics*. 2017, **18**(6), pp.32-48.
 91. Piri, R., Edenbrandt, L., Larsson, M., Enqvist, O., Nøddeskou-Fink, A.H., Gerke, O. and Høilund-Carlsen, P.F. Aortic wall segmentation in 18F-sodium fluoride PET/CT scans: Head-to-head comparison of artificial intelligence-based versus manual segmentation. *Journal of Nuclear Cardiology*. 2022, **29**(4), pp.2001-2010.
 92. Remeseiro, B. and Bolon-Canedo, V. A review of feature selection methods in medical applications. *Computers in biology and medicine*. 2019, **112**, p.103375.
 93. Steyerberg, E.W. Validation in prediction research: the waste by data splitting. *Journal of clinical epidemiology*. 2018, **103**, pp.131-133.
 94. Kaur, P. and Stoltzfus, J. Type I, II, and III statistical errors: A brief overview. *International Journal of Academic Medicine*. 2017, **3**(2), pp.268-270.
 95. Greenacre, M., Groenen, P.J., Hastie, T., d'Enza, A.I., Markos, A. and Tuzhilina, E. Principal component analysis. *Nature Reviews Methods Primers*. 2022, **2**(1), p.100.
 96. Huang, J. and Ling, C.X. Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on knowledge and Data Engineering*. 2005, **17**(3), pp.299-310.

Chapter 3 Exploring the Utility of Radiomic Feature Extraction to Improve the Diagnostic Accuracy of Cardiac Sarcoidosis Using FDG PET

Abstract

Background: This study aimed to explore the radiomic features from PET images to detect active cardiac sarcoidosis (CS).

Methods: Forty sarcoid patients and twenty-nine controls were scanned using FDG PET/CMR. Five feature classes were compared between the groups. From the PET images alone, two different segmentations were drawn. For segmentation A, a region of interest (ROI) was manually delineated for the patients' myocardium hot regions with standardised uptake value (SUV) higher than 2.5 and the controls' normal myocardium region. A second ROI was drawn in the entire left ventricular myocardium for both study groups, segmentation B. The conventional metrics and radiomic features were then extracted for each ROI. Mann-Whitney U test and a logistic regression classifier were used to compare the individual features of the study groups.

Results: For segmentation A, the SUV_{min} had the highest area under the curve (AUC) and greatest accuracy among the conventional metrics. However, for both segmentations, the AUC and accuracy of the TBR_{max} were relatively high, greater than 0.85. Twenty-two (from segmentation A) and thirty-five (from segmentation B) of 75 radiomic features fulfilled the criteria: p -value less than 0.00061 (after Bonferroni correction), AUC greater than 0.5, and accuracy greater than 0.7. Principal Component Analysis (PCA) was conducted, with five components leading to cumulative variance higher than 90%. Ten machine learning classifiers were then tested and trained. Most of them had AUCs and

accuracies ≥ 0.8 . For segmentation A, the AUCs and accuracies of all classifiers are greater than 0.9, but k-neighbors and neural network classifiers were the highest (=1). For segmentation B, there are four classifiers with AUCs and accuracies ≥ 0.8 . However, the gaussian process classifier indicated the highest AUC and accuracy (0.9 and 0.8, respectively).

Conclusion: Radiomic analysis of the specific PET data was not proven to be necessary for the detection of CS. However, building an automated procedure will help to accelerate the analysis and potentially lead to more reproducible findings across different scanners and imaging centres and consequently improve standardisation procedures that are important for clinical trials and development of more robust diagnostic protocols.

3.1 Introduction

Sarcoidosis is a multisystem, granulomatous inflammatory disease of unknown aetiology, characterised by the presence of non-caseating granulomas in the involved organs (1, 2). Sarcoidosis primarily affects the lungs. The development of this disease in the pulmonary system has been identified in more than 90% of reported cases (3, 4). However, it can affect the extrapulmonary organs as well, including the heart (5). Clinically, cardiac involvement is uncommon, manifesting in only approximately 5% of sarcoid patients, but it can occur without apparent symptoms, i.e., a 'clinically silent' disease, which is reflected in the high rate of cardiac involvement in autopsy studies. At least 25% of patients with sarcoidosis are diagnosed with cardiac involvement (6-8).

The challenging in diagnosing cardiac sarcoidosis (CS) is due to the probability of involving any organ, leads to variability in clinical presentation (9). In

addition, a lack of reliable biomarkers or diagnostic tests poses a challenge to diagnosing cardiac sarcoidosis. Furthermore, the role of advanced imaging modalities such as Cardiovascular Magnetic Resonance Imaging (CMR) with Late Gadolinium Enhancement (LGE) and [¹⁸F]Fluorodeoxyglucose Positron Emission Tomography ([¹⁸F]FDG PET) have been demonstrated in the literature to improve the identification and treatment of patients with CS. Currently, these imaging tools are critical for early diagnosis, disease prediction and progression, and therapeutic response monitoring.

To increase the diagnostic performance of [¹⁸F]FDG PET, it is important to suppress the use of glucose by normal cardiomyocytes as this improves its specificity. Several approaches have been proposed, including following a ketogenic diet (high fats and low carbohydrates), prolonged fasting, intravenous heparin, and usually, a combination of these methods (10). However, strategies to improve diagnostic performance do not help in up to 25% of patients, which can result in false-positive findings (11) due to failure to suppress the physiological uptake of the myocardium. A semi-quantitative analysis can be used to diagnose CS. A common tool, a maximum standardised uptake value (SUV_{max}), can identify the highest uptake value within the region of interest (ROI). This can differentiate positive (CS⁺) and negative (CS⁻) results; however, in the presence of high physiological uptake, this metric fails to detect sarcoidosis within this region (12). In addition, the maximum target-to-background ratio (TBR_{max}) is more robust than SUV_{max} due to the effective normalisation for blood uptake (12, 13), which makes it more reliable for comparing data across patients and institutions. Radiomic features, which rely on the spatial correlations of image values or derived image-based metrics, have the potential to elucidate features robust to background physiological

uptake. The purpose of this study is to explore radiomic features from PET images to identify potential candidate radiomic metrics. Specifically, this study will characterise radiomic features that separate active CS from controls.

3.2 Materials and Methods

3.2.1 Ethical Approval

This study was conducted with the approval of the Institutional Review Board at Mount Sinai (GCO # 01-1032), and all subjects gave written informed consent.

3.2.2 Subject Selection

Subjects with clinical suspicion of CS based on demonstrated clinical manifestations of extracardiac lesions and / or disease were recruited at Mount Sinai Hospital in New York, to undertake a PET/CMR examination. All subjects were treatment-naïve and had to avoid carbohydrate diet for 24 hours before the scan and fast during the last 12 hours. The preparation for imaging followed the recent recommendations by Ishida, Y. et al. (14). After the acquisition, the results were assessed by an expert cardiologist for indications of CS and had no indications of failed suppression of FDG uptake. Subjects were divided into patients and controls based on their results. Subjects with patchy FDG uptake were designated as CS+ and were assigned to the patient group for this study (15), and those without either FDG or CMR findings were designated as control subjects for this study. Control population had normal cardiac appearance and regular echocardiography. Forty patients and twenty-nine controls met these criteria for this study. Exclusion criteria include insulin-dependent diabetes

mellitus, pretest blood glucose > 200 mmol/dl, menopausal phobia, pregnancy / lactation, the presence of a cardiac pacemaker or automatic implantable cardioverter-defibrillator, and renal dysfunction.

3.2.3 Imaging Protocol

The simultaneous CMR with LGE and [¹⁸F]FDG PET on an integrated PET/CMR system (Biograph™ mMR, Siemens Healthcare, Erlangen, Germany) was used in this study. 5 MBq/kg of [¹⁸F]FDG was injected into the patients intravenously, who then waited for 10 minutes. Thoracic PET acquisition (one-bed position centred on the heart) took about 90 minutes but for this study only a late time window (last 60 minutes) was selected. PET images were reconstructed using the iterative ordinary Poisson ordered subset expectation maximisation (OP-OSEM) with three iterations and 21 subsets on a 344x344x129 image matrix and an isotropic voxel size of 2 mm, followed by an isotropic 4 mm Gaussian post-filtering. The data obtained with PET were not respiratory-gated or ECG-gated and were not corrected for any potential motion artefacts. A 3D breath-hold Dixon-based MR image was used for attenuation correction. Simultaneously with PET imaging, CMR was performed with electrocardiograph triggered; the scan included short-axis T2 mapping and cine images. Approximately 15 min after 0.2 mmol/kg gadolinium injection, inversion-recovery fast gradient-echo LGE sequences were acquired.

3.2.4 Segmentations

3D slicer software (Version 4.11.2; <https://www.slicer.org>) was used for the segmentation (16, 17). Segmentations were performed by study personnel according to methods used in a previous study (12).

3.2.4.1 Segmentation A

From the PET images (with use of CMR for anatomical localisation, and aiding in focal lesion identification when possible) of the patient group, an ROI was manually drawn in the hot region of the myocardium with an SUV higher than 2.5, which is a cut-off value previously used to differentiate between benign (normal in cases of CS) and malignant (abnormal in cases of CS) lesions (18, 19). For patients with more than one focal lesion, the largest and most active was selected. Due to the focal nature of the disease, applying a threshold helped ensure that the extracted features are only from voxels with abnormal uptakes. For the control group, an ROI was drawn manually in the normal myocardium. Once the SUV_{max} and SUV_{mean} (in the blood pool of the right atrium) were extracted, the TBR_{max} was calculated using the following equation:

$$TBR_{max} = \frac{SUV_{max} (target)}{SUV_{mean} (background)} \quad \text{Equation 3.1}$$

Thirty-five subjects out of forty who had a TBR_{max} within the range of 1 to 3 and patchy uptake were labelled as patients. The remaining five subjects who had $TBR_{max} > 3$ were excluded as failed suppression could not be completely discounted in these cases (12) even though the FDG was patchy and initially included in the study cohort and subsequently in the study cohort for segmentation B.

3.2.4.2 Segmentation B

As the approach A took into account both intensity and pattern, it was useful to investigate a different approach that was independent of these. From the PET images, an ROI was drawn in the entire left ventricular myocardium for forty patients and twenty-nine controls regardless of the TBR_{max} findings and SUV thresholds to compare the reliability of features among segmentation approaches. Radiomic features and conventional metrics were then extracted.

3.2.5 Feature Extraction

PyRadiomics (Version 3.0.1) was used to extract five feature classes (75 features in total) from the PET image ROIs of the patients and controls (20) in addition to the conventional metrics (7 metrics). PyRadiomics adheres to the image biomarker standardisation initiative (IBSI's feature definitions). A bin width of 0.05 was applied. All other parameters were left as default. Harmonisation was not required for these datasets as they originated from a single scanner. A list of all radiomic features and conventional metrics is shown in Supplementary Material 1.

3.2.6 Statistical Analysis

Statistical analyses were undertaken using Scikit-learn software (Version 0.23.2) (21). Mann–Whitney U test was used to compare the radiomic features of the study groups. The p -value was adjusted using a Bonferroni correction approach for multiple tests (p -value [0.05] divided by the number of features

[82]) and the corrected p -value of < 0.00061 was considered to be statistically significant. Logistic regression classifiers were then trained with individual features. Stratified five-fold cross-validation was used to determine the mean area under the curve (AUC), mean accuracy, and 95% confidence intervals (CIs). Features with a p -value < 0.00061 , $AUC > 0.5$, and accuracy > 0.7 were retained. In addition, principal component analysis (PCA) was used to identify highly correlated features and reduce feature redundancy. PCA reduces a large number of features into a small number of principal components. Components that explained 90% of the cumulative variance were retained. Lastly, to find the best machine learning (ML) algorithm, principal components were used as an input to test and train the following ten classifiers: Random Forest, Logistic Regression, Support Vector Machine, Decision Tree, Gaussian Process Classifier, Stochastic Gradient Descent, Perceptron Classifier, Passive Aggressive Classifier, Neural Network Classifier and K-neighbors Classifier with stratified five-fold cross-validation.

3.3 Results

3.3.1 Conventional Metrics Diagnostic Utility

The results are relatively different by applying the Mann–Whitney U tests on the conventional metrics of the different study groups for each segmentation separately. Predictably, for segmentation A, the SUV_{min} had the highest AUC and greatest accuracy due to specifying $SUV > 2.5$ as the minimum value for the patient group, while for segmentation B, the highest performance was for TBR_{max} (see Figure 3.1). However, for both segmentations, the AUC and accuracy of the TBR_{max} were relatively high and had similar results regardless of

the segmentation approach (AUC 0.96; accuracy 0.88–0.89 for segmentation A & B, respectively). This slight difference in TBR_{max} results between both segmentations came from the difference in the number of participants in the patient group who met the criteria for each segmentation.

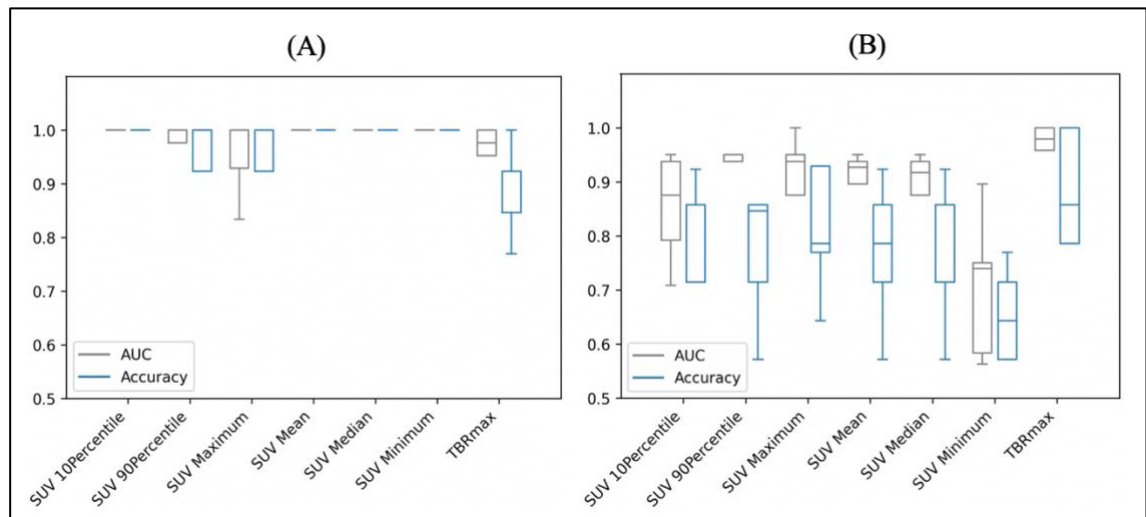


Figure 3.1 Area under the curve (AUC) and accuracy with stratified five-fold cross-validation of the conventional metrics of (A) segmentation A and (B) segmentation B. SUV: standardised uptake value, TBR_{max} : maximum target-to-background ratio.

3.3.2 Individual Radiomic Features Diagnostic Utility

From the Mann–Whitney U tests, for segmentation A: 40 of the 75 radiomic features and for segmentation B: 61 of the 75 showed statistically significant differences between patients and controls, with a p -value < 0.00061 . The five best radiomic features based on p -values for both segmentations are shown in Table 3.1. After applying a logistic regression classifier, only 22 radiomic features for segmentation A and 35 radiomic features for segmentation B fulfilled the following criteria: p -value < 0.00061 , AUC > 0.5 , and accuracy > 0.7 . The AUC and accuracy (95% CI for each criterion) with stratified five-fold cross-validation of the five best-performing radiomic features based on the AUC value are shown in Figure 3.2.

Table 3.1 Conventional metrics and five best performing radiomic features for the different segmentations based on p -values. SUV: Standardised Uptake Value, TBR_{max} : maximum Target-to-Background Ratio, GLDM: Gray Level Dependence Matrix, GLCM: Gray Level Co-occurrence Matrix, GLSZM: Gray Level Size Zone Matrix, NGTDM: Neighbouring Gray Tone Difference Matrix.

	Segmentation A			Segmentation B		
	Feature	p -value	AUC	Feature	p -value	AUC
Conventional	<i>SUV 10 Percentile</i>	1×10^{-11}	0.99	<i>SUV 10 Percentile</i>	6×10^{-7}	0.85
	<i>SUV 90 Percentile</i>	1×10^{-10}	0.96	<i>SUV 90 Percentile</i>	3×10^{-8}	0.90
	<i>SUV maximum</i>	3×10^{-10}	0.95	<i>SUV maximum</i>	8×10^{-9}	0.90
	<i>SUV mean</i>	1×10^{-10}	0.97	<i>SUV mean</i>	6×10^{-8}	0.88
	<i>SUV median</i>	1×10^{-10}	0.97	<i>SUV median</i>	2×10^{-7}	0.88
	<i>SUV minimum</i>	6×10^{-13}	1.00	<i>SUV minimum</i>	9×10^{-3}	0.71
	<i>TBR_{max}</i>	1×10^{-10}	0.96	<i>TBR_{max}</i>	3×10^{-11}	0.96
Radiomics	<i>GLDM_Small Dependence Low Gray Level Emphasis</i>	3×10^{-13}	1.00	<i>GLSZM_Low Gray Level Zone Emphasis</i>	5×10^{-8}	0.85
	<i>GLCM_Inverse Difference Normalised</i>	1×10^{-11}	1.00	<i>GLDM_Dependence Non-Uniformity</i>	1×10^{-7}	0.87
	<i>GLSZM_Small Area Low Gray Level Emphasis</i>	1×10^{-11}	0.99	<i>NGTDM_Complexity</i>	1×10^{-7}	0.85
	<i>GLSZM_Large Area High Gray Level Emphasis</i>	3×10^{-11}	1.00	<i>GLSZM_High Gray Level Zone Emphasis</i>	1×10^{-7}	0.85
	<i>GLCM_Maximal Correlation Coefficient</i>	5×10^{-11}	0.98	<i>GLSZM_Small Area High Gray Level Emphasis</i>	1×10^{-7}	0.85

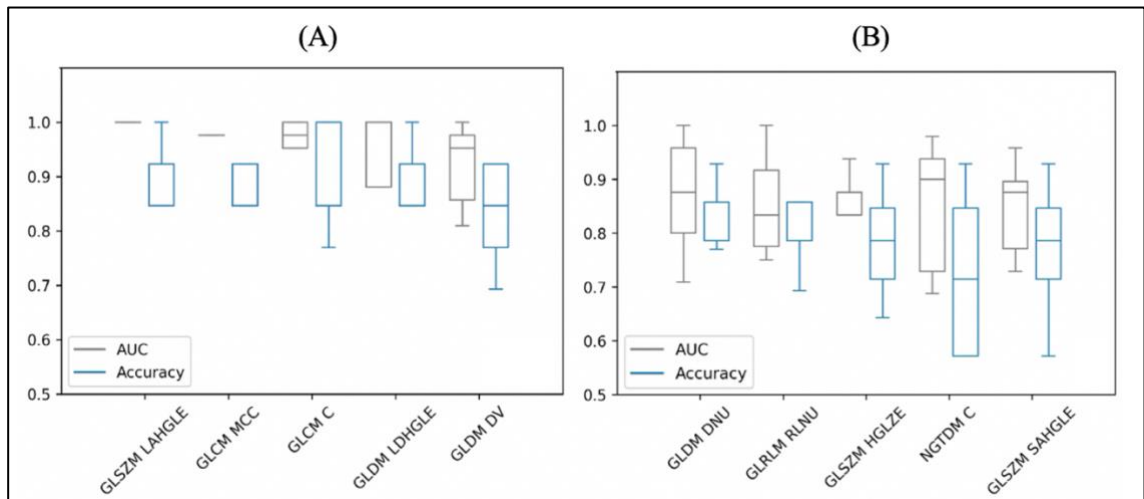


Figure 3.2 Area under the curve (AUC) and accuracy with stratified five-fold cross-validation of the five best-performing radiomic features of (A) segmentation A and (B) segmentation B based on AUC values. GLSZM: Gray Level Size Zone Matrix, LAHGLE: Large Area High Gray Level Emphasis, GLCM: Gray Level Co-occurrence Matrix, MCC: Maximal Correlation Coefficient, GLCM C: Correlation, GLDM: Gray Level Dependence Matrix, LDHGLE: Large Dependence High Gray Level Emphasis, DV: Dependence Variance, DNU: Dependence Non-Uniformity, GLRLM: Gray Level Run Length Matrix, RLNU: Run Length Non-Uniformity, HGLZE: High Gray Level Zone Emphasis, NGTDM: Neighbouring Gray Tone Difference Matrix, NGTDM C: Complexity, SAHGLE: Small Area High Gray Level Emphasis.

3.3.3 Principal Component Analysis and Machine Learning

As the SUV-related metrics tend to overperform, and to study the performance of non-first order features, the SUV-related metrics were excluded from the PCA. By applying PCA, five principal components were retained to explain 90% of the information. These components were used to test and train the ML classifiers. Most of them had AUCs and accuracies ≥ 0.8 . For segmentation A, all classifiers showed high performance in terms of AUC (95% CI 0.88–1.00) and accuracy (95% CI 0.87–1.00), with values > 0.9 . A k-neighbors and neural network classifiers showed the highest AUC and greatest accuracy, with values equal to 1.00, as shown in Figure 3.3. For segmentation B, there are four

classifiers with AUCs and accuracies ≥ 0.8 , Figure 3.3. However, the gaussian process classifier indicated the highest AUC and accuracy (0.9 and 0.8, respectively). The ROC curves of the k-neighbors, neural network, and gaussian process classifiers are shown in Figure 3.4. The actual values of Figures 3.2 and 3.3 are provided in Supplementary Material 2.

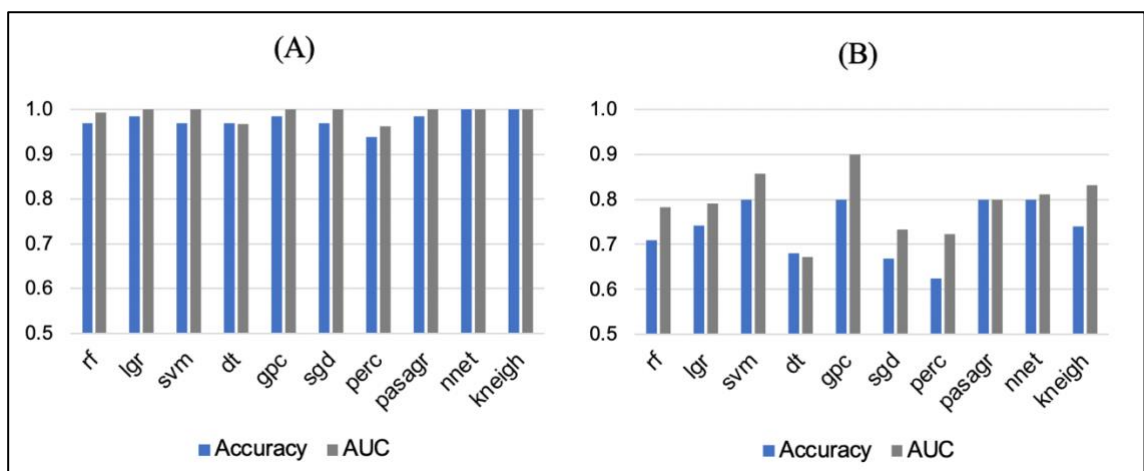


Figure 3.3 Areas under the curve (AUC) and accuracies of machine learning classifiers for (A) Segmentation A and (B) Segmentation B. rf: Random Forest, lgr: Logistic Regression, svm: Support Vector Machine, dt: Decision Tree, gpc: Gaussian Process Classifier, sgd: Stochastic Gradient Descent, perc: Perceptron Classifier, pasagr: Passive Aggressive Classifier, nnet: Neural Network Classifier, kneigh: K-neighbors Classifier.

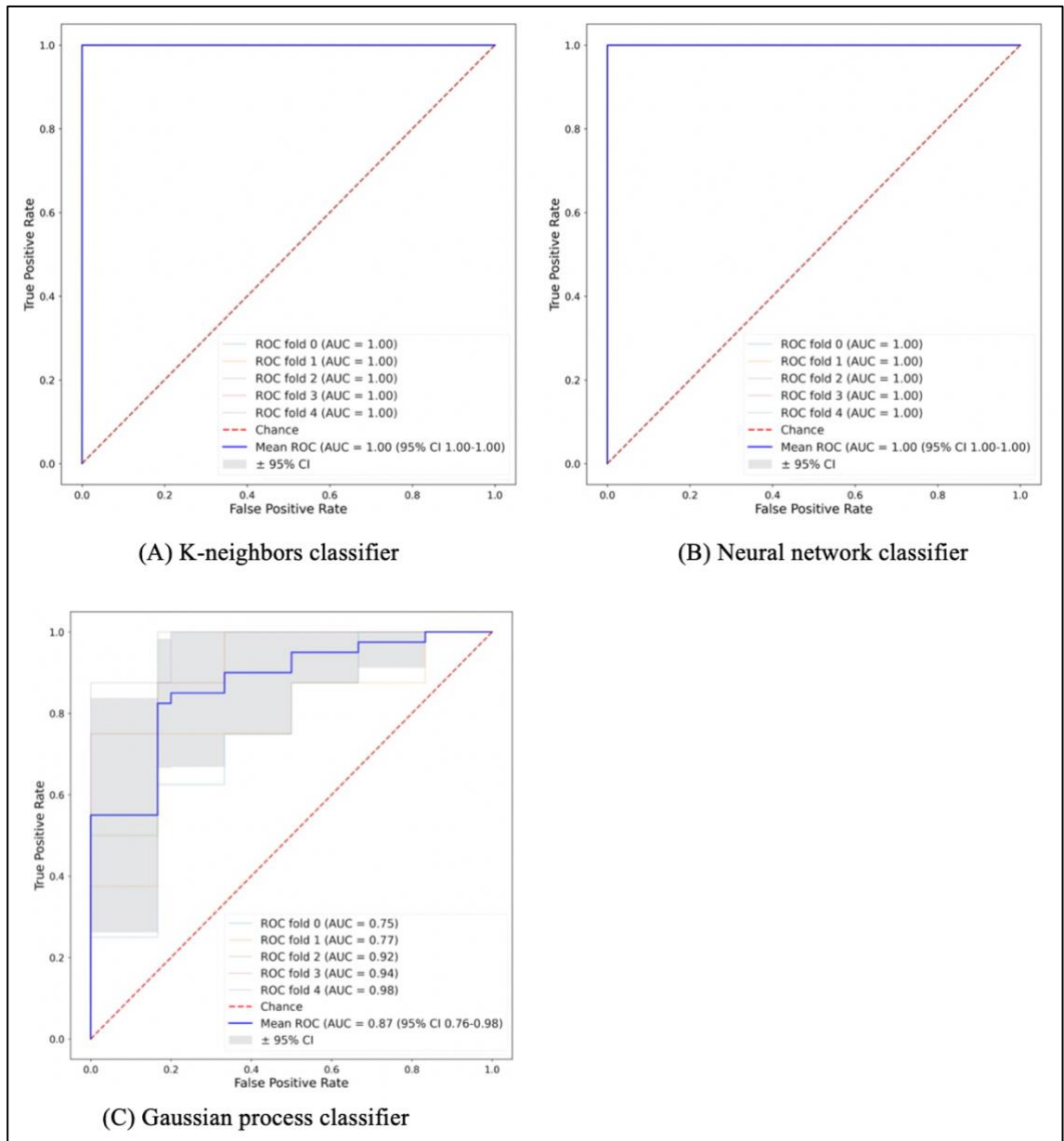


Figure 3.4 The machine learning classifiers with high performance in (A) & (B) Segmentation A and (C) Segmentation B.

3.4 Discussion

This study aimed to explore the diagnostic utility of radiomic features compared to conventional metrics to distinguish between study groups and find the best performance ML classifier to create an automated model. From segmentation A, some conventional metrics like SUV_{min} showed high performance individually. These results were predictable as they are affected by the distribution of voxel intensities within the ROI, one of the criteria for including the patients at the first place. In addition, these features cannot be relied upon because they are greatly affected by the success of glucose suppression in normal cardiomyocytes. TBR_{max} was the most reliable metric over other conventional metrics among both segmentations. Although the TBR_{max} is sensitive to noise and it is not necessarily easy to harmonise across different scanners and imaging centres, types of data, and parameters, this is not the case in this study as datasets originated from a single scanner and institution. Therefore, when comparing TBR_{max} with those of the five-best performance radiomic features, the superiority of TBR_{max} over the rest of the features can be clearly seen. This outcome supports any previous studies that utilised TBR_{max} .

From segmentation A, by comparing the diagnostic utility of individual radiomic features, *GLSZM-Large Area High Gray Level Emphasis* radiomic feature showed the best performance in terms of AUC and accuracy. This feature measures the proportion in the image of the joint distribution of larger size zones with higher gray level values. This means there is a difference in gray level zones between patients and controls. However, it cannot be reliable due to the criteria of this segmentation approach that is based on SUV threshold and TBR_{max} . On the other hand, from segmentation B, the best performing radiomic

feature was *GLDM_Dependence Non-Uniformity* with AUC (0.87) and accuracy (0.83). This feature measures the heterogeneity in the ROIs. The values of this feature are higher in sarcoid patients than controls which illustrates more heterogeneous regions in the group of patients. In addition, many other features measure heterogeneity with high AUCs and accuracies. These features look at the spatial relationships rather than voxels values themselves. However, these features had large error bars, unlike the TBR_{max} which had very small bars regardless of the segmentation approach.

Several studies of different diseases advocated the importance of radiomic analysis to predict outcomes (22, 23). However, the findings across these studies are not replicated; instead, they are conflicted. Technical issues may illustrate this difference in results among studies, such as ROI size, scanner resolution, reconstruction, and segmentation algorithms, or any other unrevealed factors. High scanner resolution and large number of voxels can affect some radiomic features by increasing their values (24). In terms of segmentation algorithms, numerous studies indicated that using different segmentation methods gave close results in survival analyses (23, 25). In addition, Cheng, N.-M. et al. (23) argued that no significant difference exists between radiomic features when using different segmentation methods, unlike SUV_{max} and SUV_{mean} . They reported, in addition, that the effect of utilising different attenuation correction methods on radiomic features was not significant. At the same time Yip, S. et al. (26) had contrasting results, as some of the features were affected by the attenuation correction method. However, in this study, there was a clear difference between radiomic features when using different segmentation approaches. This may be due to the different sizes of

ROIs and the voxel intensities included in each segmentation. Applying the approach of segmentation A, it can provide a good differentiation between study groups based on the conventional metrics such as SUV_{min} and TBR_{max} .

However, this approach can be influenced by observer experience, especially for cases with very small hotspots. Conversely, segmentation B approach is more robust and efficient.

This study is subject to some limitations. First, the sample size is relatively small, and more extensive studies are needed to confirm these results. This is of great significance to prevent overfitting and type I errors. Applying a Bonferroni correction and dimensionality reduction techniques resulted in reducing the effect of this issue. In addition, the lack of an automated segmentation, a segmentation reference to compare with, unavailability of an independent clinical gold standard to validate the performance of the model that was trained on initial input data are other limitations for this study. In addition, the selection of only one focal lesion per patient in segmentation A was considered a limitation of this approach. Furthermore, the models proposed in this study should be validated in normal controls showing non-specific physiological uptake. This study showed uncertainty results of radiomic features and expanding the study to test the reproducibility of the results is required. New knowledge gained from this study is that using radiomic analysis does not provide any additional information related to disease activity in these patients. However, building an automated model regardless of the strategies used for glucose suppression and/or observer experience may prove helpful in further studies. Furthermore, in this study, the MR acquisitions were not utilised, except

for providing anatomical information. In this study the main goal was the radiomic features on PET; the designated tool for CS.

3.5 Conclusion

Radiomic analysis of PET data may not be a useful approach to detect CS.

Several radiomic features that were not related to first-order tracer uptake showed high AUC and accuracy with p -value < 0.00061 . However, by measuring AUCs and accuracies, large error bars can weaken the results.

TBR_{max} showed its superiority over all other conventional and radiomic features in both segmentation approaches. This methodology needs to be validated further in normal control subjects showing non-specific physiological uptake.

References

1. Birnie, D.H., Nery, P.B., Ha, A.C. and Beanlands, R.S. Cardiac sarcoidosis. *Journal of the American College of Cardiology*. 2016, **68**(4), pp.411-421.
2. Hulten, E., Aslam, S., Osborne, M., Abbasi, S., Bittencourt, M.S. and Blankstein, R. Cardiac sarcoidosis—state of the art review. *Cardiovascular diagnosis and therapy*. 2016, **6**(1), p.50.
3. Ginelliová, A., Farkaš, D., Iannaccone, S.F. and Vyhnálková, V. Sudden unexpected death due to severe pulmonary and cardiac sarcoidosis. *Forensic Science, Medicine, and Pathology*. 2016, **12**(3), pp.319-323.
4. Petek, B.J., Rosenthal, D.G., Patton, K.K., Behnia, S., Keller, J.M., Collins, B.F., Cheng, R.K., Ho, L.A., Bravo, P.E. and Mikacenic, C. Cardiac sarcoidosis: diagnosis confirmation by bronchoalveolar lavage and lung biopsy. *Respiratory Medicine*. 2018, **144**, pp.S13-S19.
5. Deng, J.C., Baughman, R.P. and Lynch Iii, J.P. Cardiac involvement in sarcoidosis. In: *Seminars in respiratory and critical care medicine*: Copyright© 2002 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, 2002, pp.513-528.
6. Iwai, K., Tachibana, T., Takemura, T., Matsui, Y., Kitalchi, M. and Kawabata, Y. Pathological studies on sarcoidosis autopsy. I. Epidemiological features of 320 cases in Japan. *Pathology International*. 1993, **43**(7-8), pp.372-376.
7. Perry, A. and Vuitch, F. Causes of death in patients with sarcoidosis. A morphologic study of 38 autopsies with clinicopathologic correlations. *Archives of pathology & laboratory medicine*. 1995, **119**(2), p.167.
8. Kim, J.S., Judson, M.A., Donnino, R., Gold, M., Cooper Jr, L.T., Prystowsky, E.N. and Prystowsky, S. Cardiac sarcoidosis. *American heart journal*. 2009, **157**(1), pp.9-21.

9. Trivieri, M.G., Spagnolo, P., Birnie, D., Liu, P., Drake, W., Kovacic, J.C., Baughman, R., Fayad, Z.A. and Judson, M.A. Challenges in cardiac and pulmonary sarcoidosis: JACC state-of-the-art review. *Journal of the American College of Cardiology*. 2020, **76**(16), pp.1878-1901.
10. Chareonthaitawee, P., Beanlands, R.S., Chen, W., Dorbala, S., Miller, E.J., Murthy, V.L., Birnie, D.H., Chen, E.S., Cooper, L.T. and Tung, R.H. Joint SNMMI–ASNC expert consensus document on the role of 18F-FDG PET/CT in cardiac sarcoid detection and therapy monitoring. *Journal of nuclear medicine*. 2017, **58**(8), pp.1341-1353.
11. Osborne, M.T., Hulten, E.A., Murthy, V.L., Skali, H., Taqueti, V.R., Dorbala, S., DiCarli, M.F. and Blankstein, R. Patient preparation for cardiac fluorine-18 fluorodeoxyglucose positron emission tomography imaging of inflammation. *Journal of Nuclear Cardiology*. 2017, **24**(1), pp.86-99.
12. Dweck, M.R., Abgral, R., Trivieri, M.G., Robson, P.M., Karakatsanis, N., Mani, V., Palmisano, A., Miller, M.A., Lala, A. and Chang, H.L. Hybrid magnetic resonance imaging and positron emission tomography with fluorodeoxyglucose to diagnose active cardiac sarcoidosis. *JACC: Cardiovascular Imaging*. 2018, **11**(1), pp.94-107.
13. Chen, W. and Dilsizian, V. PET assessment of vascular inflammation and atherosclerotic plaques: SUV or TBR? *Journal of Nuclear Medicine*. 2015, **56**(4), pp.503-504.
14. Ishida, Y., Yoshinaga, K., Miyagawa, M., Moroi, M., Kondoh, C., Kiso, K. and Kumita, S. Recommendations for 18 F-fluorodeoxyglucose positron emission tomography imaging for cardiac sarcoidosis: Japanese Society of Nuclear Cardiology recommendations. *Annals of nuclear medicine*. 2014, **28**(4), pp.393-403.
15. Birnie, D.H., Sauer, W.H., Bogun, F., Cooper, J.M., Culver, D.A., Duvernoy, C.S., Judson, M.A., Kron, J., Mehta, D. and Nielsen, J.C. HRS expert consensus statement on the diagnosis and management of

- arrhythmias associated with cardiac sarcoidosis. *Heart rhythm*. 2014, **11**(7), pp.1304-1323.
16. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F. and Sonka, M. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging*. 2012, **30**(9), pp.1323-1341.
 17. Kikinis, R., Pieper, S.D. and Vosburgh, K.G. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In: *Intraoperative imaging and image-guided therapy*. Springer, 2014, pp.277-289.
 18. Kadaria, D., Freire, A.X., SultanAli, I., Zaman, M.K., Archie, D.S. and Weiman, D.S. Dual time point positron emission tomography/computed tomography scan in evaluation of intrathoracic lesions in an area endemic for histoplasmosis and with high prevalence of sarcoidosis. *The American journal of the medical sciences*. 2013, **346**(5), pp.358-362.
 19. Wang, T., Sun, H., Guo, Y. and Zou, L. 18F-FDG PET/CT Quantitative Parameters and Texture Analysis Effectively Differentiate Endometrial Precancerous Lesion and Early-Stage Carcinoma. *Molecular imaging*. 2019, **18**, p.1536012119856965.
 20. Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.-C., Pieper, S. and Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer research*. 2017, **77**(21), pp.e104-e107.
 21. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. and Dubourg, V. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*. 2011, **12**, pp.2825-2830.
 22. Apostolova, I., Ego, K., Steffen, I.G., Buchert, R., Wertz, H., Achenbach, H.J., Riedel, S., Schreiber, J., Schultz, M. and Furth, C. The asphericity of the metabolic tumour volume in NSCLC: correlation with

- histopathology and molecular markers. *European Journal of Nuclear Medicine and Molecular Imaging*. 2016, **43**(13), pp.2360-2373.
23. Cheng, N.-M., Fang, Y.-H.D., Tsan, D.-L., Hsu, C.-H. and Yen, T.-C. Respiration-averaged CT for attenuation correction of PET images– impact on PET texture features in non-small cell lung cancer patients. *PLoS One*. 2016, **11**(3), p.e0150509.
24. Han, S., Woo, S., Suh, C.H., Kim, Y.J., Oh, J.S. and Lee, J.J. A systematic review of the prognostic value of texture analysis in 18 F-FDG PET in lung cancer. *Annals of nuclear medicine*. 2018, **32**(9), pp.602-610.
25. Bashir, U., Azad, G., Siddique, M.M., Dhillon, S., Patel, N., Bassett, P., Landau, D., Goh, V. and Cook, G. The effects of segmentation algorithms on the measurement of 18 F-FDG PET texture parameters in non-small cell lung cancer. *Ejnmri Research*. 2017, **7**(1), p.60.
26. Yip, S., McCall, K., Aristophanous, M., Chen, A.B., Aerts, H.J. and Berbeco, R. Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer. *PLoS One*. 2014, **9**(12), p.e115510.

Supplementary Materials

Supplementary Material 1:

Table1: List of all conventional and radiomic features

Conventional metrics	
TBR _{max}	firstorder_Mean
firstorder_10Percentile	firstorder_Median
firstorder_90Percentile	firstorder_Minimum
firstorder_Maximum	
GLCM features	
glcm_Autocorrelation	glcm_Idn
glcm_ClusterProminence	glcm_Imc1
glcm_ClusterShade	glcm_Imc2
glcm_ClusterTendency	glcm_InverseVariance
glcm_Contrast	glcm_JointAverage
glcm_Correlation	glcm_JointEnergy
glcm_DifferenceAverage	glcm_JointEntropy
glcm_DifferenceEntropy	glcm_MCC
glcm_DifferenceVariance	glcm_MaximumProbability
glcm_Id	glcm_SumAverage
glcm_Idm	glcm_SumEntropy
glcm_Idmn	glcm_SumSquares
GLRLM features	
glrlm_GrayLevelNonUniformity	glrlm_RunEntropy
glrlm_GrayLevelNonUniformityNormalised	glrlm_RunLengthNonUniformity
glrlm_GrayLevelVariance	glrlm_RunLengthNonUniformityNormalised
glrlm_HighGrayLevelRunEmphasis	glrlm_RunPercentage
glrlm_LongRunEmphasis	glrlm_RunVariance
glrlm_LongRunHighGrayLevelEmphasis	glrlm_ShortRunEmphasis
glrlm_LongRunLowGrayLevelEmphasis	glrlm_ShortRunHighGrayLevelEmphasis
glrlm_LowGrayLevelRunEmphasis	glrlm_ShortRunLowGrayLevelEmphasis
GLSZM features	
glszm_GrayLevelNonUniformity	glszm_SizeZoneNonUniformity
glszm_GrayLevelNonUniformityNormalised	glszm_SizeZoneNonUniformityNormalised
glszm_GrayLevelVariance	glszm_SmallAreaEmphasis
glszm_HighGrayLevelZoneEmphasis	glszm_SmallAreaHighGrayLevelEmphasis
glszm_LargeAreaEmphasis	glszm_SmallAreaLowGrayLevelEmphasis

glszm_LargeAreaHighGrayLevelEmphasis	glszm_ZoneEntropy
glszm_LargeAreaLowGrayLevelEmphasis	glszm_ZonePercentage
glszm_LowGrayLevelZoneEmphases	glszm_ZoneVariance

GLDM features

gldm_DependenceEntropy	gldm_LargeDependenceEmphasis
gldm_DependenceNonUniformity	gldm_LargeDependenceHighGrayLevelEmphasis
gldm_DependenceNonUniformityNormalised	gldm_LargeDependenceLowGrayLevelEmphasis
gldm_DependenceVariance	gldm_LowGrayLevelEmphasis
gldm_GrayLevelNonUniformity	gldm_SmallDependenceEmphasis
gldm_GrayLevelVariance	gldm_SmallDependenceHighGrayLevelEmphasis
gldm_HighGrayLevelEmphasis	gldm_SmallDependenceLowGrayLevelEmphasis

NGTDM features

ngtdm_Busyness	ngtdm_Contrast
ngtdm_Coarseness	ngtdm_Strength
ngtdm_Complexity	

Supplementary Material 2:

Table 1: Area under the curve (AUC) and accuracy of the five best-performing radiomic features for the different segmentations based on AUC values. GLSZM: Gray Level Size Zone Matrix, LAHGLE: Large Area High Gray Level Emphasis, GLCM: Gray Level Co-occurrence Matrix, MCC: Maximal Correlation Coefficient, GLCM_C: Correlation, GLDM: Gray Level Dependence Matrix, LDHGLE: Large Dependence High Gray Level Emphasis, DV: Dependence Variance, DNU: Dependence Non-Uniformity, GLRLM: Gray Level Run Length Matrix, RLNU: Run Length Non-Uniformity, HGLZE: High Gray Level Zone Emphasis, NGTDM: Neighbouring Gray Tone Difference Matrix, NGTDM_C: Complexity, SAHGLE: Small Area High Gray Level Emphasis.

Segmentation A			Segmentation B		
Feature	Accuracy	AUC	Feature	Accuracy	AUC
<i>GLSZM_LAHGLE</i>	0.91	1.00	<i>GLDM_DNU</i>	0.83	0.87
<i>GLCM_MCC</i>	0.88	0.98	<i>GLRLM_RLNU</i>	0.81	0.86
<i>GLCM_C</i>	0.89	0.96	<i>GLSZM_HGLZE</i>	0.78	0.85
<i>GLDM_LDHGLE</i>	0.86	0.95	<i>NGTDM_C</i>	0.73	0.85
<i>GLDM_DV</i>	0.83	0.92	<i>GLSZM_SAHGLE</i>	0.77	0.85

Table 2: Area under the curve (AUC) and accuracy of the machine learning classifiers for the different segmentations.

Machine Learning Classifier	Segmentation A		Segmentation B	
	Accuracy	AUC	Accuracy	AUC
Random Forest	0.97	0.99	0.71	0.78
Logistic Regression	0.99	1.00	0.74	0.79
Support Vector Machine	0.97	1.00	0.80	0.86
Decision Tree	0.97	0.97	0.68	0.67
Gaussian Process Classifier	0.99	1.00	0.80	0.90
Stochastic Gradient Descent	0.97	1.00	0.67	0.73
Perceptron Classifier	0.94	0.96	0.63	0.72
Passive Aggressive Classifier	0.99	1.00	0.80	0.80
Neural Network Classifier	1.00	1.00	0.80	0.81
K-neighbors Classifier	1.00	1.00	0.74	0.83

Chapter 4 Exploring the Utility of Cardiovascular Magnetic Resonance Radiomic Feature Extraction for Evaluation of Cardiac Sarcoidosis

Abstract

Background: The aim of this study is to explore the utility of cardiac magnetic resonance (CMR) imaging of radiomic features to distinguish active and inactive cardiac sarcoidosis (CS).

Methods: Subjects were classified into active cardiac sarcoidosis (CS_{active}) and inactive cardiac sarcoidosis ($CS_{inactive}$) based on PET/CMR imaging. CS_{active} was classified as featuring patchy [^{18}F]fluorodeoxyglucose ([^{18}F]FDG) uptake on PET and presence of late gadolinium enhancement (LGE) on CMR, while $CS_{inactive}$ was classified as featuring no [^{18}F]FDG uptake in the presence of LGE on CMR. Among those screened, thirty CS_{active} and thirty-one $CS_{inactive}$ patients met these criteria. A total of 94 radiomic features were subsequently extracted using PyRadiomics. The values of individual features were compared between CS_{active} and $CS_{inactive}$ using the Mann–Whitney U test. Subsequently, machine learning (ML) approaches were tested. ML was applied to two sub-sets of radiomic features (signatures A and B) that were selected by logistic regression and PCA, respectively.

Results: Univariate analysis of individual features showed no significant differences. Of all features, *gray level co-occurrence matrix (GLCM) joint entropy* had a good area under the curve (AUC) and accuracy with the smallest confidence interval, suggesting it may be a good target for further investigation. Some ML classifiers achieved reasonable discrimination between CS_{active} and $CS_{inactive}$ patients. With signature A, support vector machine and k-neighbors showed good performance with AUC (0.77 and 0.73) and accuracy (0.67 and

0.72), respectively. With signature B, decision tree demonstrated AUC and accuracy around 0.7.

Conclusion: CMR radiomic analysis in CS provides promising results to distinguish patients with active and inactive disease.

4.1 Introduction

Sarcoidosis, an inflammatory multisystem disorder of unknown origin, is characterised by the formation of non-caseating granulomas (1). Although the disease affects the lungs in more than 90% of patients, other organs and tissues such as the heart, skin, and lymph nodes may also be affected (2). Despite the low incidence of clinical manifestations of cardiac disease, cardiac sarcoidosis can exist as a potentially fatal disease because of associated ventricular arrhythmias. Therefore, it is crucial to identify individuals with active cardiac sarcoidosis, even if they are subclinical (3). The diagnosis of cardiac sarcoidosis can be established with certainty by means of an endomyocardial biopsy if non-caseating granulomas are identified. However, an invasive approach poses a high degree of risk that does not yield a significant improvement in sensitivity because myocardial involvement is patchy (4).

There has been an increase in the use of non-invasive advanced imaging approaches, including cardiovascular magnetic resonance imaging (CMR) and [¹⁸F]fluorodeoxyglucose ([¹⁸F]FDG) positron emission tomography (PET). [¹⁸F]FDG PET imaging is effective in detecting myocardial inflammation in cardiac sarcoidosis (5). To increase the specificity of PET imaging, it is important to follow a ketogenic diet 24 h before the scan. However, while dietary

restrictions can suppress myocardial physiological glucose uptake, they may not be effective, potentially resulting in false-positive results (6,7). On the other hand, CMR has a significant role diagnosing or screening patients with cardiac sarcoidosis. It can detect signs that might indicate disease, such as myocardial fibrosis, myocardial oedema, and perfusion defects; it is also used to assess the geometry and function of both ventricles. Late gadolinium enhancement (LGE) is a technique used to detect myocardial fibrosis in cardiac sarcoidosis which typically appears in a non-coronary distribution (3). Presence of fibrosis, however, is not capable of determining whether the disease is active or chronic.

Using quantitative measurements can provide complementary information that may overcome limitations of non-invasive approaches (8). Radiomic analysis is an emerging methodology that automatically extracts high dimensional features from imaging data, which can later be mined and analysed for decision support (9). The aim of this study is to explore the utility of radiomic analysis of LGE-CMR to separate those with active cardiac sarcoidosis, based on patchy [^{18}F]FDG uptake, from those with inactive cardiac sarcoidosis, without [^{18}F]FDG uptake (10). Such an outcome may prove useful in detecting active cardiac sarcoidosis even in the presence of inconclusive or false-positive results on [^{18}F]FDG PET.

4.2 Materials and Methods

4.2.1 Ethical Approval

This study is an extension of a previous study (7) and was conducted with the approval of the Institutional Review Board at Mount Sinai (GCO # 01-1032), and all participants provided written informed consent.

4.2.2 Subject selection

PET/CMR imaging was performed at Mount Sinai Hospital in New York, on patients with clinical suspicion of cardiac sarcoidosis based on extracardiac disease and cardiac symptoms. All subjects were treatment-naïve. As recommended by Ishida et al. (11), patient preparation required 24 h of carbohydrate abstinence and a 12 h fast before the scan. Exclusion criteria included renal dysfunction, insulin-dependent diabetes, blood glucose levels greater than 200 mg/dL, pregnancy and lactation, the presence of a cardiac pacemaker or an automatic implantable cardioverter-defibrillator, as well as failed myocardial suppression defined by high maximum target-to-background ratio ($TBR_{max} > 3$) (10) with widespread uptake in a non-specific pattern as determined by an expert reader in the use of PET/MR to diagnose sarcoid cardiomyopathies.

4.2.3 Imaging Protocols

Simultaneous CMR and [^{18}F]FDG PET was performed on an integrated PET/MR system (BiographTM mMR, Siemens Healthcare, Erlangen, Germany). 5 MBq/kg of [^{18}F]FDG was injected into the patients intravenously. After 10 min, thoracic PET acquisition (one-bed position centred on the heart) began and lasted for 90 min; for this study, one time window (40–100 min post injection) was reconstructed. PET images were reconstructed using the iterative ordinary

Poisson ordered subset expectation maximisation (OP-OSEM) with three iterations and 21 subsets on a $344 \times 344 \times 129$ image matrix and an isotropic voxel size of 2 mm, followed by post-filtering with an isotropic 4 mm Gaussian kernel. The data obtained with PET were not respiratory-gated or ECG-gated and were not corrected for any potential motion artifacts. A 3D breath-hold Dixon-based MR image was used for attenuation correction. Simultaneously with PET imaging, CMR was performed with electrocardiograph triggering; the scan included short-axis T2 mapping and cine images covering the whole left ventricle. Approximately 15 min after injection of 0.2 mmol/kg gadolinium-based contrast agent (MultiHance, Bracco, NJ, USA), inversion-recovery fast gradient-echo LGE sequences were acquired with 8 mm slice thickness and 10 mm spacing between short-axis slices across the entire myocardium.

4.2.4 Patient Classification

Following acquisition, a single expert cardiologist evaluated the results for signs of cardiac sarcoidosis. Firstly, only subjects with LGE on CMR in a non-coronary distribution representative of cardiac sarcoidosis were selected.

Subsequently, subjects with patchy [^{18}F]FDG uptake on PET were classified as active cardiac sarcoidosis ($\text{CS}_{\text{active}}$) and those who did not show any [^{18}F]FDG findings were classified as inactive cardiac sarcoidosis ($\text{CS}_{\text{inactive}}$). For this study of 148 patients scanned at the institution, thirty $\text{CS}_{\text{active}}$ and thirty-one $\text{CS}_{\text{inactive}}$ met these criteria.

4.2.5 Segmentation

3D slicer software (Version 4.11.2; <https://www.slicer.org>, accessed on 1 April 2022) was used for segmentation (12,13). Segmentations were performed by study personnel. From the LGE-CMR images, epicardial and endocardial boundaries were drawn to define a region of interest (ROI) encompassing the entire left ventricular myocardium. Radiomic features and conventional metrics were then extracted from the ROI.

4.2.6 Feature Extraction

PyRadiomics (Version 3.0.1; Harvard Medical School, Boston, MA, USA) was used to extract six feature classes (94 features in total) from the LGE-CMR images (14). The first-order statistical features consist of histogram (HISTO)-based properties. The features in this class are used to determine the statistical values of voxel intensities and evaluate the shape of the histogram, regardless of spatial relationships (15). The second-order statistical features comprise features that are utilised to calculate the statistical inter-relationships between adjacent voxels and can be derived from the *gray level cooccurrence matrix (GLCM)* (16,17). The higher-order statistical features can be used to extract areas with increasingly coarse texture patterns (18). They are derived from the *gray level run length matrix (GLRLM)*, *gray level dependence matrix (GLDM)*, *gray level size zone matrix (GLSZM)*, and *neighbouring gray tone difference matrix (NGTDM)* (19). PyRadiomics adheres to most of the image biomarker standardisation initiative (IBSI) feature definitions. All parameters were maintained at their default values. Harmonisation was not required for these datasets as they originated from a single scanner.

4.2.7 Statistical Analysis

Statistical analyses were undertaken using Scikit-learn software (Version 0.23.2) (20). The Mann–Whitney U test was used to compare the individual radiomic features of the study groups. The p -value was adjusted using a Bonferroni correction approach for multiple tests. Given an initial significance level of 0.05, and 94 features, a p -value of < 0.00053 was considered to be statistically significant.

For machine learning (ML), sub-sets of radiomic features (signatures A and B) were selected using two approaches. For signature A, logistic regression classifiers were trained with individual features. Stratified five-fold cross-validation was used to determine the mean area under the curve (AUC), mean accuracy, and 95% confidence intervals (CIs). Features with $AUC > 0.5$ and $accuracy > 0.7$ were retained. Then, Spearman correlation was used to detect the correlated features and the feature with the lower AUC was removed. For signature B, principal component analysis (PCA), which reduces a large number of features into a small number of principal components, was used to identify highly correlated features and reduce feature redundancy. Components that explained 90% of the cumulative variance were retained. Both signatures were used as input to test and train the following ten ML classifiers: random forest, logistic regression, support vector machine, decision tree, Gaussian process classifier, stochastic gradient descent, perceptron classifier, passive aggressive classifier, neural network classifier and k-neighbors classifier with stratified five-fold cross-validation.

4.3 Results

4.3.1 Individual Features—Diagnostic Utility

From the univariate analysis of individual features, none of the radiomic features showed statistically significant differences on Mann–Whitney U tests between CS_{active} and CS_{inactive} , with p -values > 0.00053 . Furthermore, by measuring the effect size, the majority of the radiomic features presented small effect size values ≤ 0.5 , which can be improved by increasing the sample size. The ten best-performing radiomic features based on the p -values were shown in Table 4.1.

Table 4.1 The ten best-performing radiomic features based on the p -values. CS_{active} : active cardiac sarcoidosis, CS_{inactive} : inactive cardiac sarcoidosis, GLCM: gray level co-occurrence matrix, GLRLM: gray level run length matrix, GLDM: gray level dependence matrix, GLSZM: gray level size zone matrix.

Feature	CS_{active} Mean	CS_{inactive} Mean	U Statistic	p-Value	Effect Size
<i>glcm_Cluster Shade</i>	9.27	7.69	701	0.0007	0.09
<i>glcm_Cluster Prominence</i>	79.37	85.63	693	0.0010	0.03
<i>firstorder_Variance</i>	721.11	485.32	689	0.0013	0.50
<i>glrlm_Gray Level Variance</i>	1.36	1.01	689	0.0013	0.38
<i>gldm_Gray Level Variance</i>	1.21	0.83	677	0.0023	0.50
<i>glcm_Maximal Correlation Coefficient</i>	0.43	0.36	676	0.0024	0.60
<i>firstorder_Mean Absolute Deviation</i>	19.05	14.50	674	0.0026	0.75
<i>glcm_Correlation</i>	0.36	0.28	673	0.0028	0.61
<i>glszm_Size Zone Non Uniformity</i>	57.69	38.59	673	0.0028	0.81
<i>glrlm_Run Entropy</i>	2.98	2.81	672	0.0029	0.65

4.3.2 Signature Building and Machine Learning Performance

For signature A, after applying a logistic regression, only nine radiomic features had $AUC > 0.5$ and accuracy > 0.7 . The correlated features were then removed by removing the one with the lower AUC; five were retained (Figure 4.1). One of the retained five features with the smallest confidence interval after correlated features were removed was *GLCM joint entropy*. The performance of *GLCM joint entropy* in distinguishing those with active disease (with [^{18}F]FDG) from those with inactive disease (without [^{18}F]FDG) is shown in Figure 4.2. Following a qualitative assessment, there were 31 CS_{inactive} . Of these, 77.5% were distinguished as CS_{inactive} but with a 22.5% type I error using radiomic analysis (*GLCM joint entropy*). On the other side, there were 30 CS_{active} based on the PET/CMR visual assessment. The *GLCM joint entropy* was able to identify two-thirds of these subjects as CS_{active} . For signature B, by applying PCA, six principal components were retained to explain 90% of the information. These signatures were used to test and train the ML classifiers.

Most of the ML classifiers showed poor performance in terms of AUC (95% CI 0.09–0.95 and 0.39–0.94) and accuracy (95% CI 0.35–0.82 and 0.36–0.82), for signature A and signature B respectively, as shown in Figure 4.3. For signature A, the support vector machine and k-neighbors ML classifiers had good performance with AUC (0.77 and 0.73) and accuracy (0.67 and 0.72), respectively. For signature B, the decision tree ML classifier was the only one that showed good performance with AUC and accuracy ≈ 0.7 , while the random forest, Gaussian process, and passive aggressive ML classifiers had high AUCs > 0.7 but poor accuracies ≤ 0.6 . The performance of the best ML classifiers of

both signatures is shown in Figure 4.4. The values for accuracy and AUC presented in Figures 4.1, 4.2, and 4.4 are provided in Supplementary Material 1.

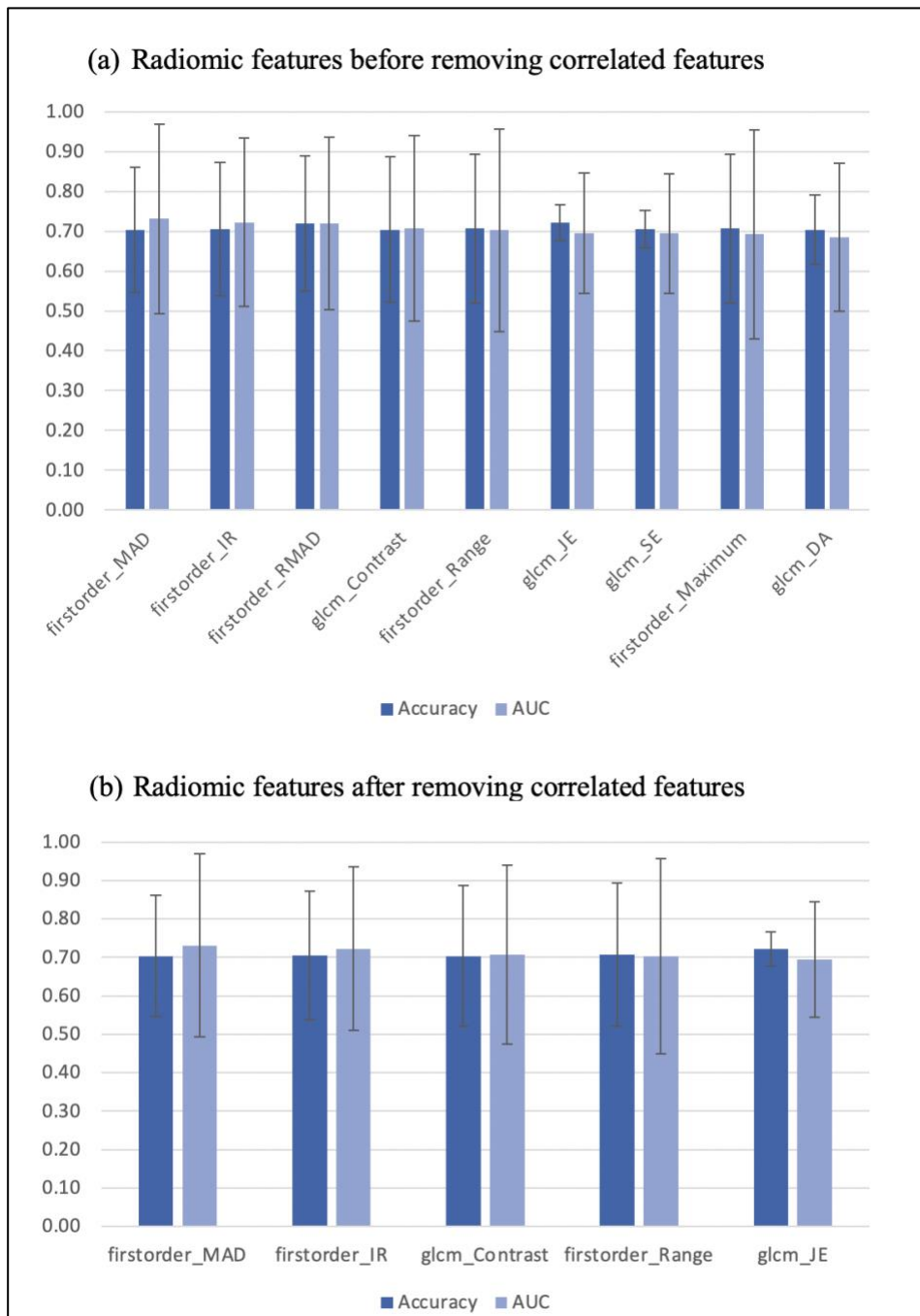


Figure 4.1 Radiomic features before and after removing correlated features with 95% confidence intervals. MAD: mean absolute deviation, IR: interquartile range, RMAD: robust mean absolute deviation, GLCM: gray level co-occurrence matrix, JE: joint entropy, SE: sum entropy, DA: difference average.

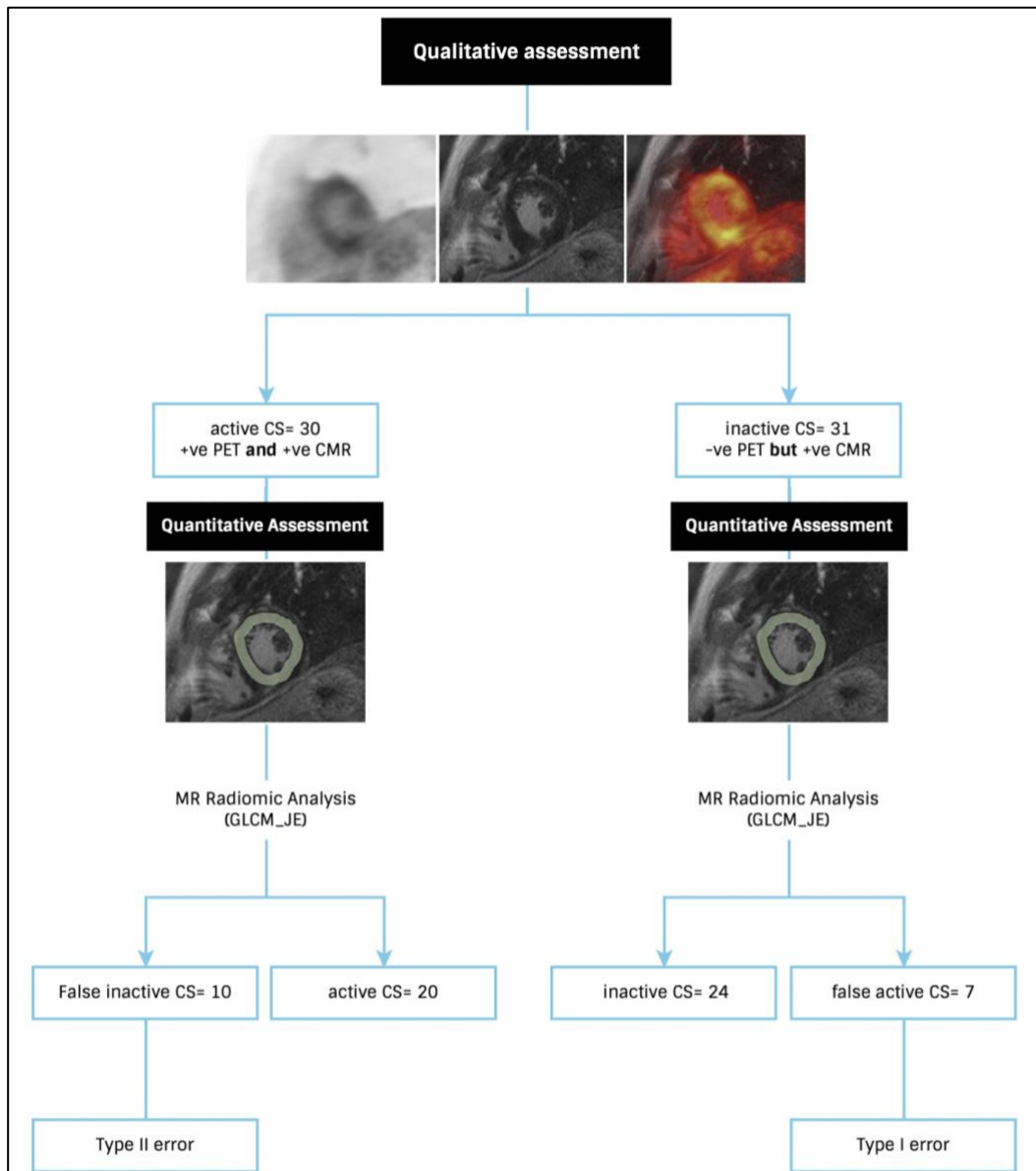


Figure 4.2 Flowchart showing the performance of the LGE-CMR based radiomic feature gray level co-occurrence matrix joint entropy (GLCM_JE) in discriminating active from inactive disease in cardiac sarcoidosis (CS).

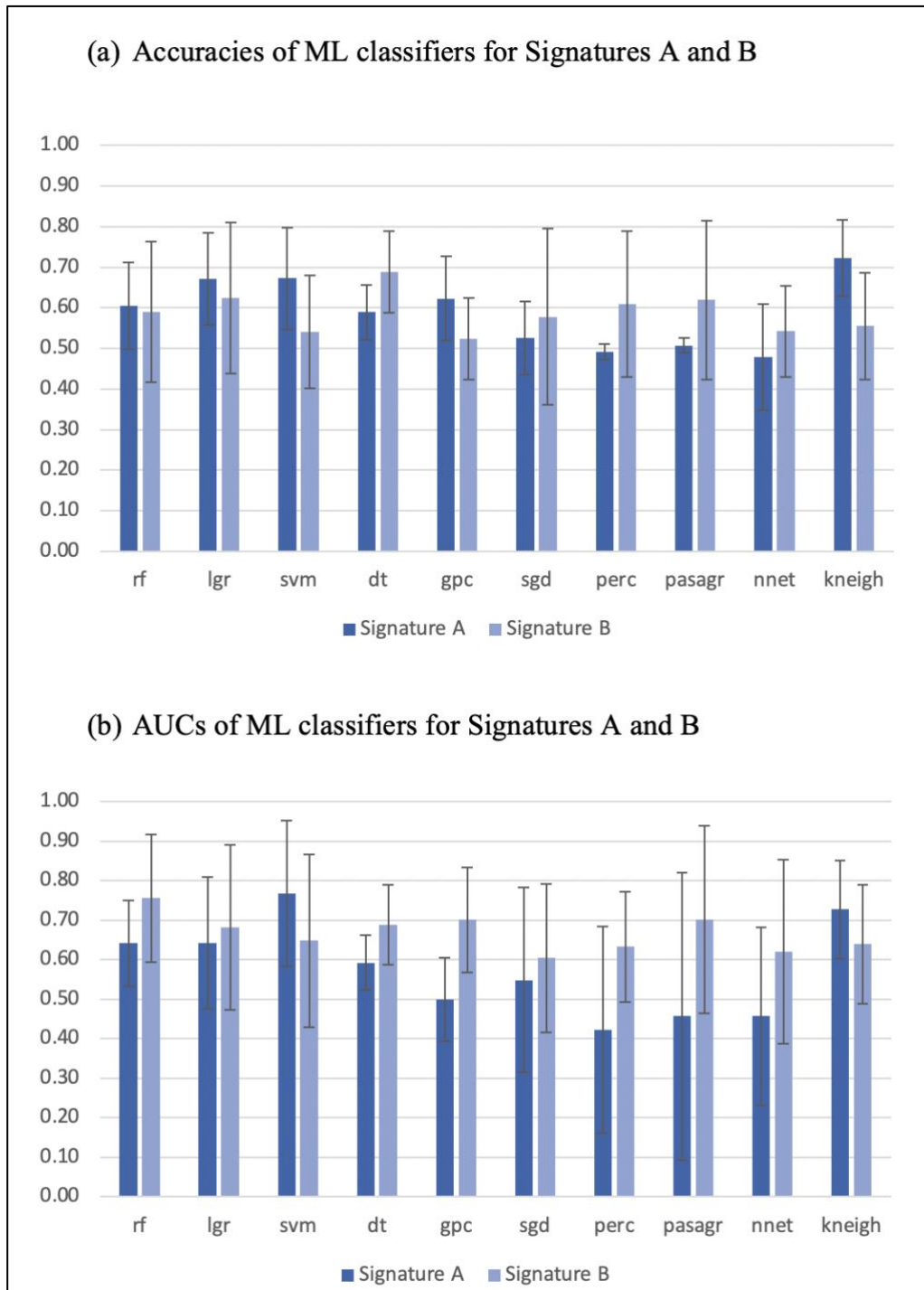


Figure 4.3 Areas under the curve (AUC) and accuracies of machine learning (ML) classifiers of both signatures with 95% confidence intervals. rf: random forest, lgr: logistic regression, svm: support vector machine, dt: decision tree, gpc: Gaussian process classifier, sgd: stochastic gradient descent, perc: perceptron classifier, pasagr: passive aggressive classifier, nnet: neural network classifier, kneigh: k-neighbors classifier.

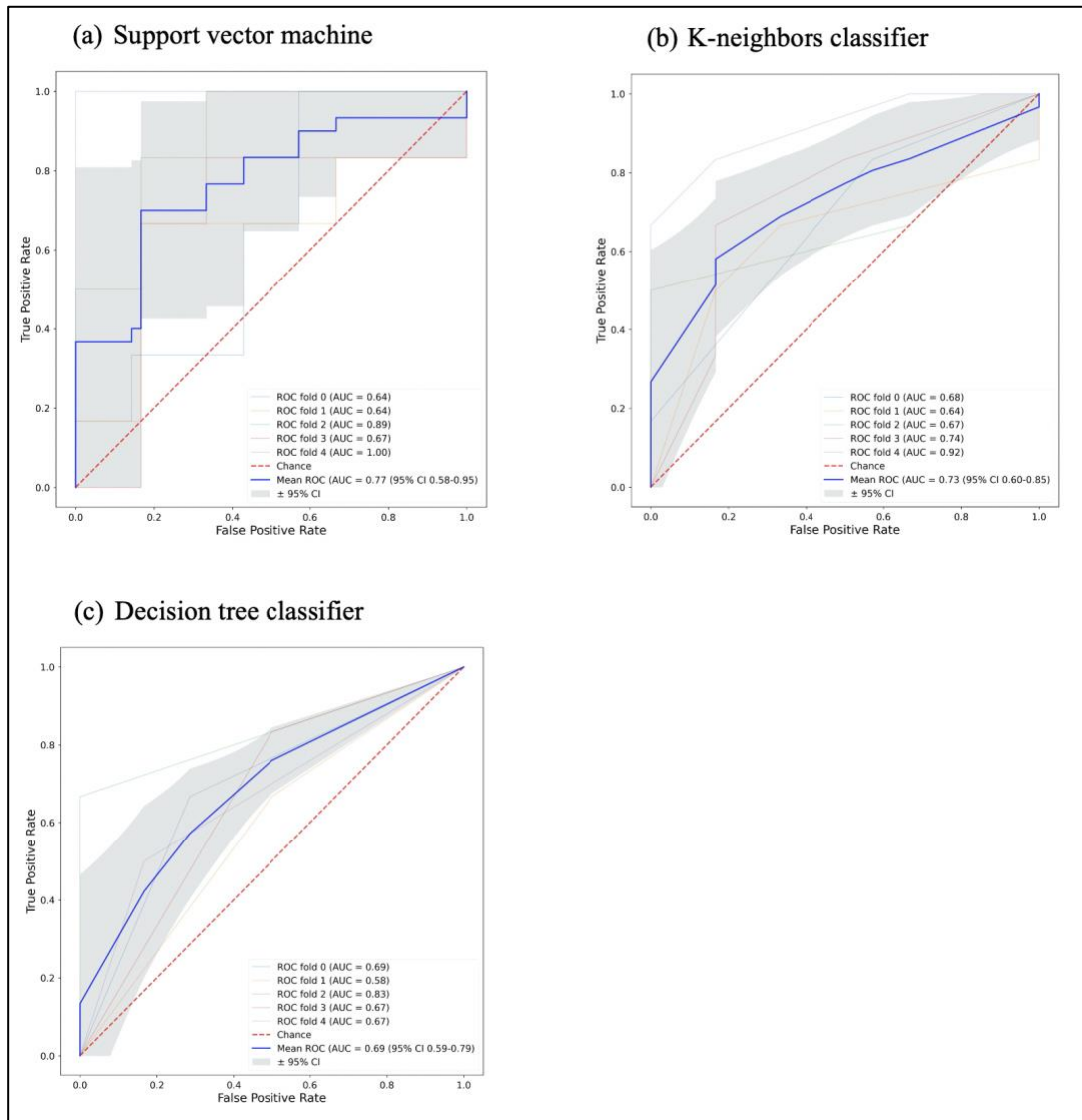


Figure 4.4 The performance of the best machine learning classifiers for both signatures, (a,b) for signature A, and (c) for signature B.

4.4 Discussion

This study aimed to investigate the utility of radiomic features derived from LGE-CMR images to distinguish active from inactive cardiac sarcoidosis. The univariate analysis of individual radiomic features showed that the individual

radiomic features were not adequate to differentiate between patients in the CS_{active} and $CS_{inactive}$ groups with statistical significance. However, many features in the *GLCM* class were among the top ten individual features, although they did not reach significance. These second-order features calculate the statistical inter-relationships between adjacent voxels (16,17). These features inform on the spatial distribution of voxel intensities and therefore are a measure of signal heterogeneity (21,22). It is noteworthy that these second-order features outperformed many first-order features. Given that the presence of LGE (a first-order criterion) was a pre-requisite for both CS_{active} and $CS_{inactive}$ groups, it could be expected that first-order features would not be effective discriminators; however, the overall burden of fibrosis (and therefore LGE) would also be a likely indicator of active disease. In addition, it is interesting to note similar findings in the ML approach using signature A. About half of the features were from the *GLCM* class. From the uncorrelated features, *GLCM joint entropy* represented the best performing feature with the highest accuracy (0.72) and smallest confidence interval. The propensity for *GLCM joint entropy* to separate the study groups can be explained by the fact that *joint entropy* is a measure of the degree of randomness or variance in the pixel intensities in a given pixel's neighborhood (23,24). Slightly higher values related to this feature were observed in the CS_{active} group than in the $CS_{inactive}$ group, pointing to the presence of heterogeneity, one of the characteristics of this disease (25).

A number of limitations apply to this study. It should be noted, first, that the sample size is relatively small, and further studies would be necessary to verify these findings. The small sample size may lead to overfitting and type I errors. In order to reduce the impact of this issue, Bonferroni correction and

dimensionality reduction techniques were applied. Caution should be taken in interpreting our findings; of note, while the radiomic feature *GLCM joint entropy* showed the highest performance using logistic regression during feature selection for signature A, it was not among the top-ten features under univariate analysis, indicating that many features exhibit similar performance. Further limitations of this study include the lack of an automated segmentation, a segmentation reference, and the absence of an independent clinical gold standard based on biopsy to validate the model's performance. Further studies are required to test the reproducibility of our findings based on radiomic features.

The new knowledge gained from this study is that radiomic analysis of LGE-CMR has the potential to identify active disease based on CMR alone. CMR-based analysis of disease activity could improve management for patients who undergo [¹⁸F]FDG PET/CMR imaging and exhibit non-specific findings related to failed suppression of physiological uptake of [¹⁸F]FDG in the myocardium. Such non-specific findings may be present in up to 25% of PET scans (6). However, given the limitations of the study, the results should be interpreted with caution. Moreover, the trend towards successful discrimination of CS_{active} from CS_{inactive} using second-order radiomic features such as *GLCM joint entropy* is encouraging and may indicate promising candidates for further evaluation. In future studies, the combined radiomic analysis of [¹⁸F]FDG PET (7) and LGE-CMR data may further increase the accuracy of discriminating active from inactive disease, may be useful in determining prognosis, and may aid in clinical decision-making (26).

4.5 Conclusion

This study explored the use of radiomic analysis of LGE-CMR images in patients with cardiac sarcoidosis to distinguish between active cardiac sarcoidosis (CS_{active}) and inactive cardiac sarcoidosis (CS_{inactive}) based on LGE-CMR data alone. Both individual radiomic features and ML classifiers based on groups of radiomic features showed a modest ability to separate CS_{active} from CS_{inactive} . *GLCM joint entropy* (95% CI accuracy 0.68 to 0.77; AUC 0.54 to 0.85) emerged as the individual radiomic feature with the greatest accuracy in separating CS_{active} from CS_{inactive} .

References

1. Birnie, D.H.; Nery, P.B.; Ha, A.C.; Beanlands, R.S. Cardiac sarcoidosis. *J. Am. Coll. Cardiol.* **2016**, *68*, 411–421.
2. Ginelliová, A.; Farkaš, D.; Iannaccone, S.F.; Vyhnálková, V. Sudden unexpected death due to severe pulmonary and cardiac sarcoidosis. *Forensic Sci. Med. Pathol.* **2016**, *12*, 319–323.
3. Nappi, C.; Ponsiglione, A.; Imbriaco, M.; Cuocolo, A. *¹⁸F-FDG PET/CMR in Cardiac Sarcoidosis: A Wild Card in the Deck?* Springer: Berlin/Heidelberg, Germany, 2020; pp. 1–3.
4. Bennett, M.K.; Gilotra, N.A.; Harrington, C.; Rao, S.; Dunn, J.M.; Freitag, T.B.; Halushka, M.K.; Russell, S.D. Evaluation of the role of endomyocardial biopsy in 851 patients with unexplained heart failure from 2000–2009. *Circ. Heart Fail.* **2013**, *6*, 676–684.
5. Chareonthitawee, P.; Beanlands, R.S.; Chen, W.; Dorbala, S.; Miller, E.J.; Murthy, V.L.; Birnie, D.H.; Chen, E.S.; Cooper, L.T.; Tung, R.H. *Joint SNMMI–ASNC Expert Consensus Document on the Role of ¹⁸F-FDG PET/CT in Cardiac Sarcoid Detection and Therapy Monitoring*; Springer: Berlin/Heidelberg, Germany, 2017.
6. Osborne, M.T.; Hulten, E.A.; Murthy, V.L.; Skali, H.; Taqueti, V.R.; Dorbala, S.; DiCarli, M.F.; Blankstein, R. Patient preparation for cardiac fluorine-18 fluorodeoxyglucose positron emission tomography imaging of inflammation. *J. Nucl. Cardiol.* **2017**, *24*, 86–99.
7. Mushari, N.A.; Soutanidis, G.; Duff, L.; Trivieri, M.G.; Fayad, Z.A.; Robson, P.; Tsoumpas, C. Exploring the utility of radiomic feature extraction to

- improve the diagnostic accuracy of cardiac sarcoidosis using FDG PET. *Front. Med.* **2022**, *9*, 840261.
8. Rainey, C.; O'Regan, T.; Matthew, J.; Skelton, E.; Woznitza, N.; Chu, K.-Y.; Goodman, S.; McConnell, J.; Hughes, C.; Bond, R. and Malamateniou, C. An insight into the current perceptions of UK radiographers on the future impact of AI on the profession: A cross-sectional survey. *J. Med. Imaging Radiat. Sci.* **2022**, *53*, 347–361.
 9. Duff, L.; Scarsbrook, A.F.; Mackie, S.L.; Froom, R.; Bailey, M.; Morgan, A.W.; Tsoumpas, C. A methodological framework for AI-assisted diagnosis of active aortitis using radiomic analysis of FDG PET–CT images: Initial analysis. *J. Nucl. Cardiol.* **2022**, *29*, 3315–3331.
 10. Dweck, M.R.; Abgral, R.; Trivieri, M.G.; Robson, P.M.; Karakatsanis, N.; Mani, V.; Palmisano, A.; Miller, M.A.; Lala, A.; Chang, H.L. Hybrid magnetic resonance imaging and positron emission tomography with fluorodeoxyglucose to diagnose active cardiac sarcoidosis. *JACC Cardiovasc. Imaging* **2018**, *11*, 94–107.
 11. Ishida, Y.; Yoshinaga, K.; Miyagawa, M.; Moroi, M.; Kondoh, C.; Kiso, K.; Kumita, S. Recommendations for 18F-fluorodeoxyglucose positron emission tomography imaging for cardiac sarcoidosis: Japanese Society of Nuclear Cardiology recommendations. *Ann. Nucl. Med.* **2014**, *28*, 393–403.
 12. Fedorov, A.; Beichel, R.; Kalpathy-Cramer, J.; Finet, J.; Fillion-Robin, J.-C.; Pujol, S.; Bauer, C.; Jennings, D.; Fennessy, F.; Sonka, M. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **2012**, *30*, 1323–1341.

13. Kikinis, R.; Pieper, S.D.; Vosburgh, K.G. 3D Slicer: A platform for subject-specific image analysis, visualization, and clinical support. In *Intraoperative Imaging and Image-Guided Therapy*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 277–289.
14. Van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.-C.; Pieper, S.; Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107.
15. Rizzo, S.; Botta, F.; Raimondi, S.; Origgi, D.; Fanciullo, C.; Morganti, A.G.; Bellomi, M. Radiomics: The facts and the challenges of image analysis. *Eur. Radiol. Exp.* **2018**, *2*, 36.
16. Balagurunathan, Y.; Kumar, V.; Gu, Y.; Kim, J.; Wang, H.; Liu, Y.; Goldgof, D.B.; Hall, L.O.; Korn, R.; Zhao, B. Test–retest reproducibility analysis of lung CT image features. *J. Digit. Imaging* **2014**, *27*, 805–823.
17. Ergen, B.; Baykara, M. Texture based feature extraction methods for content based medical image retrieval systems. *Bio-Med. Mater. Eng.* **2014**, *24*, 3055–3062.
18. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563.
19. Lovinfosse, P.; Hatt, M.; Visvikis, D.; Hustinx, R. Heterogeneity analysis of ¹⁸F-FDG PET imaging in oncology: Clinical indications and perspectives. *Clin. Transl. Imaging* **2018**, *6*, 393–410.
20. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

21. Xu, X.; Zhang, X.; Tian, Q.; Zhang, G.; Liu, Y.; Cui, G.; Meng, J.; Wu, Y.; Liu, T.; Yang, Z. Three-dimensional texture features from intensity and high-order derivative maps for the discrimination between bladder tumors and wall tissues via MRI. *Int. J. Comput. Assist. Radiol. Surg.* **2017**, *12*, 645–656.
22. Duff, L.; Tsoumpas, C. Automated diagnosis and prediction in cardiovascular diseases using tomographic imaging. *Big Data Multimodal Med. Imaging* **2019**, 71–102. DOI:10.1201/b22410-4.
23. Zwanenburg, A.; Leger, S.; Vallières, M.; Löck, S. Image biomarker standardisation initiative. *arXiv* **2016**, arXiv:1612.07003.
24. Avard, E.; Shiri, I.; Hajianfar, G.; Abdollahi, H.; Kalantari, K.R.; Houshmand, G.; Kasani, K.; Bitarafan-Rajabi, A.; Deevband, M.R.; Oveisi, M. Non-contrast Cine Cardiac Magnetic Resonance image radiomics features and machine learning algorithms for myocardial infarction detection. *Comput. Biol. Med.* **2022**, *141*, 105145.
25. Tana, C.; Mantini, C.; Donatiello, I.; Mucci, L.; Tana, M.; Ricci, F.; Cipollone, F.; Giamberardino, M.A. Clinical Features and Diagnosis of Cardiac Sarcoidosis. *J. Clin. Med.* **2021**, *10*, 1941.
26. Vallières, M.; Freeman, C.R.; Skamene, S.R.; El Naqa, I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys. Med. Biol.* **2015**, *60*, 5471.

Supplementary Materials

Supplementary Material 1:

Table 1: Area under the curve (AUC) and accuracy of the five best-performing radiomic features based on AUC values before and after correlated features removed. GLCM: Gray Level Co-occurrence Matrix.

	Feature	Accuracy	AUC	β	p-value
Correlation retained	<i>Firstorder_Mean Absolute Deviation</i>	0.70	0.73	0.139	1.25×10^{-08}
	<i>Firstorder_Interquartile Range</i>	0.71	0.72	0.018	1.41×10^{-05}
	<i>Firstorder_Robust Mean Absolute Deviation</i>	0.72	0.72	-	-
	<i>GLCM_Contrast</i>	0.70	0.71	-0.001	1.97×10^{-13}
	<i>Firstorder_Range</i>	0.71	0.70	-0.185	1.24×10^{-02}
Correlation removed	<i>Firstorder_Mean Absolute Deviation</i>	0.70	0.73	0.139	1.25×10^{-08}
	<i>Firstorder_Interquartile Range</i>	0.71	0.72	0.018	1.41×10^{-05}
	<i>GLCM_Contrast</i>	0.70	0.71	-0.001	1.97×10^{-13}
	<i>Firstorder_Range</i>	0.71	0.70	-0.185	1.24×10^{-02}
	<i>GLCM_Joint Entropy</i>	0.72	0.70	-0.389	1.85×10^{-02}

Table 2: Area under the curve (AUC) and accuracy of the machine learning classifiers for the different signatures.

Machine Learning Classifier	Signature A		Signature B	
	Accuracy	AUC	Accuracy	AUC
Random Forest	0.61	0.64	0.59	0.76
Logistic Regression	0.67	0.64	0.62	0.68
Support Vector Machine	0.67	0.77	0.54	0.65
Decision Tree	0.59	0.59	0.69	0.69
Gaussian Process Classifier	0.62	0.50	0.52	0.70
Stochastic Gradient Descent	0.53	0.55	0.58	0.61
Perceptron Classifier	0.49	0.42	0.61	0.63
Passive Aggressive Classifier	0.51	0.46	0.62	0.70
Neural Network Classifier	0.48	0.46	0.54	0.62
K-neighbors Classifier	0.72	0.73	0.56	0.64

Chapter 5 An Assessment of PET and CMR Radiomic Features for Detection of Cardiac Sarcoidosis

Abstract

Background: Visual interpretation of PET and CMR may fail to identify cardiac sarcoidosis (CS) with high specificity. This study aimed to evaluate the role of [¹⁸F]FDG PET and late gadolinium enhancement (LGE)-CMR radiomic features in differentiating CS from another cause of myocardial inflammation, in this case patients with cardiac-related clinical symptoms following COVID-19.

Methods: [¹⁸F]FDG PET and LGE-CMR were treated separately in this work. There were thirty-five post-COVID-19 (PC) and forty CS datasets. Regions of interest were delineated manually around the entire left ventricle for PET and LGE-CMR datasets. Radiomic features were then extracted. The ability of individual features to correctly identify image data as CS or PC was tested to predict clinical classification of CS versus PC using Mann–Whitney U tests and logistic regression. Features were retained if p -value < 0.00053 , AUC > 0.5 and accuracy > 0.7 . After applying correlation test, uncorrelated features were used as a signature (joint features) to train machine learning classifiers. For LGE-CMR analysis, to further improve the results, different classifiers were used for individual features besides logistic regression and the results of individual features of each classifier were screened to create a signature that include all features that followed the previously mentioned criteria and use them as input for machine learning classifiers.

Results: The Mann–Whitney U tests and logistic regression were trained on individual features to build a collection of features. For [¹⁸F]FDG PET analysis, the maximum target-to-background ratio (TBR_{max}) showed high area under the curve (AUC) and accuracy with small p -values (< 0.00053) but the signature

performed better (AUC 0.98 and accuracy 0.91). For LGE-CMR analysis, *Gray Level Dependence Matrix (gldm)-Dependence Non-Uniformity* showed good results with small error bars (accuracy 0.75 and AUC 0.87). However, by applying a Support Vector Machine classifier on individual LGE-CMR features and creating a signature, a Random Forest classifier displayed better AUC and accuracy (0.91 and 0.84, respectively).

Conclusion: Using radiomic features may prove useful in identifying individuals with CS. Some features showed promising results to differentiate between PC and CS. By automating the analysis, the patient management process can be accelerated and improved.

5.1 Introduction

Cardiac sarcoidosis (CS) is a granulomatous inflammatory disease that can be diagnosed with [¹⁸F]-fluorodeoxyglucose positron emission tomography ([¹⁸F]FDG PET). [¹⁸F]FDG PET is performed in suspected CS due to the avid uptake of glucose by the active inflammation cells in sarcoid granulomas. It is recommended that a low-carbohydrate, high-fat diet followed by fasting be used to inhibit the physiologic glucose metabolism of the heart to enable diagnostic imaging. Moreover, a cardiac PET with abnormal [¹⁸F]FDG uptake on suppressed myocardial uptake is crucial to CS diagnosis (1). A PET image can also be used to quantify inflammation in addition to a visual review. Several metrics exist to describe the intensity and heterogeneity of [¹⁸F]FDG uptake. PET is less specific for CS when there is no extracardiac uptake (2). In addition, it is critical to note that approximately 25% of cardiac PET studies fail due to the inadequate suppression of physiologic glucose uptake (3).

Conversely, cardiovascular magnetic resonance (CMR) is a non-invasive imaging technique that plays a significant role in diagnosing or screening patients with CS. It can detect scar tissue that may indicate inactive CS (4). Myocardial scarring can be evaluated using late gadolinium enhancement (LGE) imaging. Gadolinium is an extracellular contrast agent that demonstrates a slow washout in fibrotic regions compared to the normal myocardium. Although LGE is helpful for identifying CS, based on the LGE distribution and pattern (5, 6), it is a non-specific tool. In addition, LGE-CMR has limited sensitivity prior to myocardial scar development (7).

Moreover, [¹⁸F]FDG PET can detect the inflammation related to CS, which theoretically leads to its early diagnosis (8). On the other hand, CMR with LGE is capable of identifying myocardial scarring even in small areas, owing to its high spatial resolution. The specificity of CMR in diagnosing CS might be higher than [¹⁸F]FDG PET, however, both have high sensitivity (9). There is controversy among studies regarding the identification of the appropriate technique for diagnosing CS (10-13). Similarly, the feasibility of combining the findings of both [¹⁸F]FDG PET and LGE-CMR has not been adequately explored and could enhance the accuracy of the assessment owing to the identification of different pathologic features.

Additionally, it may be possible to gain additional information by employing quantitative measurements that may supply complementary information greater than that provided by non-invasive methods (14). A method of analysing

imaging data uses radiomics to automatically extract high-dimensional features. Subsequently, researchers can mine and analyse these features to support decision making (15, 16). First-order statistical features comprise properties based on histograms (*HISTO*). Regardless of the spatial relationship between the voxels, these features are based on the shape of the histogram and discern statistical values of the voxel intensities (17, 18). Statistical inter-relationships between neighbouring voxels are calculated using second-order statistical features, which can be derived from the *gray-level cooccurrence matrix (GLCM)* (17). In addition, areas with coarser textures can be extracted using higher-order statistical features (19). They are derived from the *gray level run length matrix (GLRLM)*, *gray level dependence matrix (GLDM)*, *gray level size zone matrix (GLSZM)*, and *neighbouring gray tone difference matrix (NGTDM)*.

Correspondingly, this work investigates the precision of PET and CMR radiomic features in differentiating CS from another cause of myocardial inflammation, in this case, patients with cardiac-related symptoms following COVID-19, or post-COVID-19 (PC) patients. Myocardial inflammation can be a symptom observed in some PC patients. The severity and prevalence of myocardial inflammation may vary among individuals, and it is one of the potential complications associated with COVID-19. It is important to note that not all PC patients will experience myocardial inflammation, and the manifestation of symptoms can vary widely (20).

5.2 Materials and Methods

5.2.1 Ethical Approval

This study was conducted with the approval of the Institutional Review Board at Mount Sinai Hospital (GCO # 01-1032). All the subjects supplied their written informed consent.

5.2.2 Subject Selection

Both PET and CMR imaging were performed at Mount Sinai Hospital in New York on two types of patients: patients suspected of having cardiac sarcoidosis based on extracardiac disease and cardiac symptoms, as well as PC patients. The majority of the CS cohort predates the COVID-19 era, ensuring that these patients did not exhibit post-COVID-19 symptoms. CS diagnosis is consistent with the Heart Rhythm Society (HRS) expert consensus statement (10). PC patients had either chest pain, palpitations, or shortness of breath following COVID-19 that could not be attributed to another cause. This retrospective study encompassed CS and PC patients exhibiting abnormal FDG uptake in the myocardium that were evaluated by a cardiologist who is an expert in the use of PET/MR to diagnose cardiomyopathies. Exclusions were made for individuals with renal dysfunction, insulin-dependent diabetes, blood glucose levels exceeding 200 mg/dL, pregnant or lactating individuals, and those with cardiac pacemakers or automatic implantable cardioverter-defibrillators. In preparation for the scan, the patient was required to abstain from carbohydrate consumption for 24 hours and fast for 12 hours. Initially, there were 90 suspected PC patients and 69 patients with CS. However, for the purpose of this study, only cases with myocarditis were included. Therefore, the study included thirty-five datasets from PC patients and forty datasets from patients with CS, as summarised in Figure 5.1. The demographic information of the patients is provided in Table 5.1.

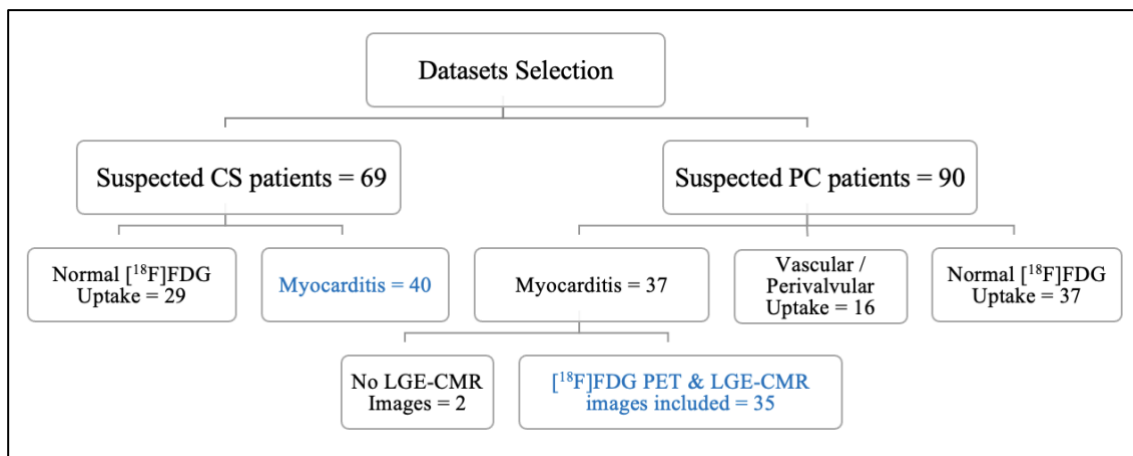


Figure 5.1 A flowchart of datasets selection

Table 5.1 Demographic information of the study population. PC: Post-COVID-19 Patients; CS: Cardiac Sarcoidosis Patients

Group	Sex	Mean Age	Standard Deviation
PC	F=19	44.2	12.27
	M=16		
CS	F=16	61.35	9.41
	M=24		

5.2.3 Imaging Protocol

An integrated PET/MR system was used to perform simultaneous CMR and $[^{18}\text{F}]\text{FDG}$ PET (Biograph™ mMR, Siemens Healthcare, Erlangen, Germany). An intravenous injection of $[^{18}\text{F}]\text{FDG}$ containing 5 MBq/kg was given to the patients. Acquisition of thoracic PET (one-bed position centred on the heart) takes approximately 90 minutes to scan the patients in two phases (blood and tissue phases). However, for the purpose of this study, only the last 60 minutes of the time window were chosen because the focus of this study specifically centres on the tissue phase. Iterative ordinary Poisson ordered subset expectation

maximisation (OP-OSEM) was used to reconstruct PET images over a 344×344×129 image matrix with 3 iterations, 21 subsets and 2 mm isotropic voxels, followed by post-filtering using a Gaussian kernel of 4 mm. A PET study was neither respiratory-gated nor electrocardiogram (ECG)-gated, and no motion correction was carried out. Attenuation correction was performed using a 3D breath-hold Dixon-based MR image. Parallel to the PET scan, CMR was performed with ECG triggering covering the whole left ventricle. Inversion-recovery gradient-echo LGE sequences were acquired across the entire myocardium approximately 15 minutes after injection of 0.2 mmol/kg gadolinium-based contrast agent (MultiHance, Bracco, NJ) with 8 mm slice thickness and 10 mm spacing between slices. Bias correction was not performed in CMR images.

5.2.4 Segmentation

3D slicer software (Version 4.11.2; <https://www.slicer.org>) was used for the segmentation (21, 22). Regions of interest (ROI) were drawn manually in the entire left ventricular myocardium for both [¹⁸F]FDG PET and LGE-CMR images by a junior radiographer and reviewed by a Biomedical Engineering expert with 10 years of experience in Medical Imaging. This approach is less likely to be influenced by the intensity and experience of observers compared to the hot regions-only segmentation. The hot regions-only segmentation may exhibit bias and result in unreliable outcomes during testing in our prior study (23). Figure 5.2 provides an illustrative example of the segmentation on PET/CMR image. Subsequently, radiomic features were extracted.

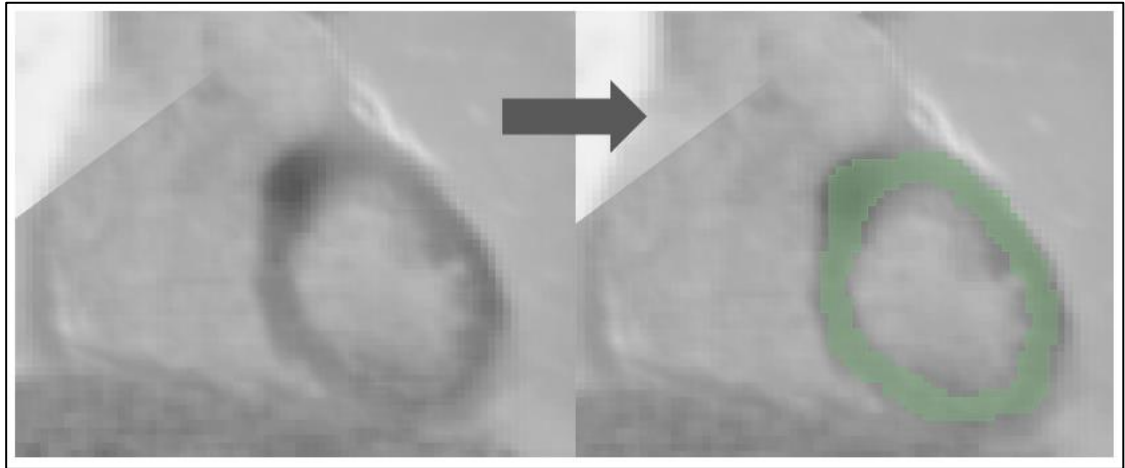


Figure 5.2 An example of the segmentation on PET/CMR image.

To calculate the maximum target-to-background ratio (TBR_{max}) in PET images, the standardised uptake value (SUV_{max}) was extracted, and another ROI was drawn in the blood pool to extract the (SUV_{mean}) of the background and then follow the following equation:

$$TBR_{max} = \frac{SUV_{max} (target)}{SUV_{mean} (background)} \quad \text{Equation 5.1}$$

5.2.5 Feature Extraction

PyRadiomics (Version 3.0.1) was used to extract six feature classes (94 features in total) from the PET/CMR images (24). A list of all radiomic features is shown in Supplementary Material 1. PyRadiomics adheres to most of the image biomarker standardisation initiative (IBSI) feature definitions. In the case

of PET images, a fixed bin size of 0.075 was utilised, this gave a good number of bins and a good representation of the data. However, in LGE-CMR images, the default fixed bin size of 25 was used. The impact of gray-level discretisation on extracted feature values from PET images has been well-documented (25). Nevertheless, there is limited research exploring the effect of gray-level discretisation on clinical MR images. According to Duron, L. et al. (26), an experimental study aimed at examining the impact of gray-level discretisation on the reproducibility of texture features from MR images, it was found that utilising different fixed bin sizes had a minimal effect on the variability of these features. The PET images were subjected to SUV normalisation. Since the datasets were obtained from a single scanner, pre-processing (except post-filtering using a Gaussian kernel in reconstruction process in PET images) and harmonisation were not performed. The feature extraction was conducted in 3D, as it provides more informative results compared to 2D analysis. In addition, to mitigate the risk of overfitting caused by limited data, the models were not optimised.

5.2.6 Statistical Analysis

Statistical analyses were undertaken using the Scikit-learn software (Version 0.23.2) (27). The individual radiomic features of the study groups were compared using Mann-Whitney U test to assess their ability to separate CS from PC. In addition, Bonferroni correction was used to adjust the p -value for multiple tests. According to the significance level of 0.05, and with 94 features, the corrected p -value was < 0.00053 . The radiomic features were then trained and tested using logistic regression classifiers. This analysis used stratified five-fold cross-validation to obtain the mean area under the curve (AUC), mean

accuracy, and 95% confidence intervals (CIs). An AUC > 0.5 and an accuracy > 0.7 were considered acceptable for the retention of features with a p -value of less than 0.00053. When $0.5 < \text{AUC} < 1$, there is a high chance that the feature will be able to distinguish the positive class values from the negative ones. In addition, accuracy > 0.7 can be considered as a decent score. Subsequently, Spearman correlation was used to detect the correlated features with 0.70 correlation coefficient. This threshold was selected because higher thresholds indicate a strong resemblance between the two features, with at least half of their variance being shared. Of these correlated features, the feature with the highest AUC was retained. Following that, the uncorrelated features were then used as input for machine learning classifiers to create a signature (joint features). In LGE-CMR features, to find a classifier that can provide high values of AUC and accuracy, other classifiers besides logistic regression were explored. The retained features were then used as input for machine learning classifiers. The selection of the top-performing machine learning classifier was based on the highest mean AUC and mean accuracy values from stratified five-fold cross-validation. Due to the small sample size in this study, only the training cross-validation outcomes were documented. This approach has been recommended in situations where the sample size is insufficient to support an independent validation set (28). By using cross-validation, the potential overestimation of the model's performance was reduced. The workflow of the statistical analysis is illustrated in Figure 5.3. In this study, the PET and CMR datasets were analysed separately, enabling a more focused investigation of the specific features and characteristics inherent to each modality. This approach yields valuable insights into the individual contributions of PET and CMR, enhancing the understanding of the subject under investigation.

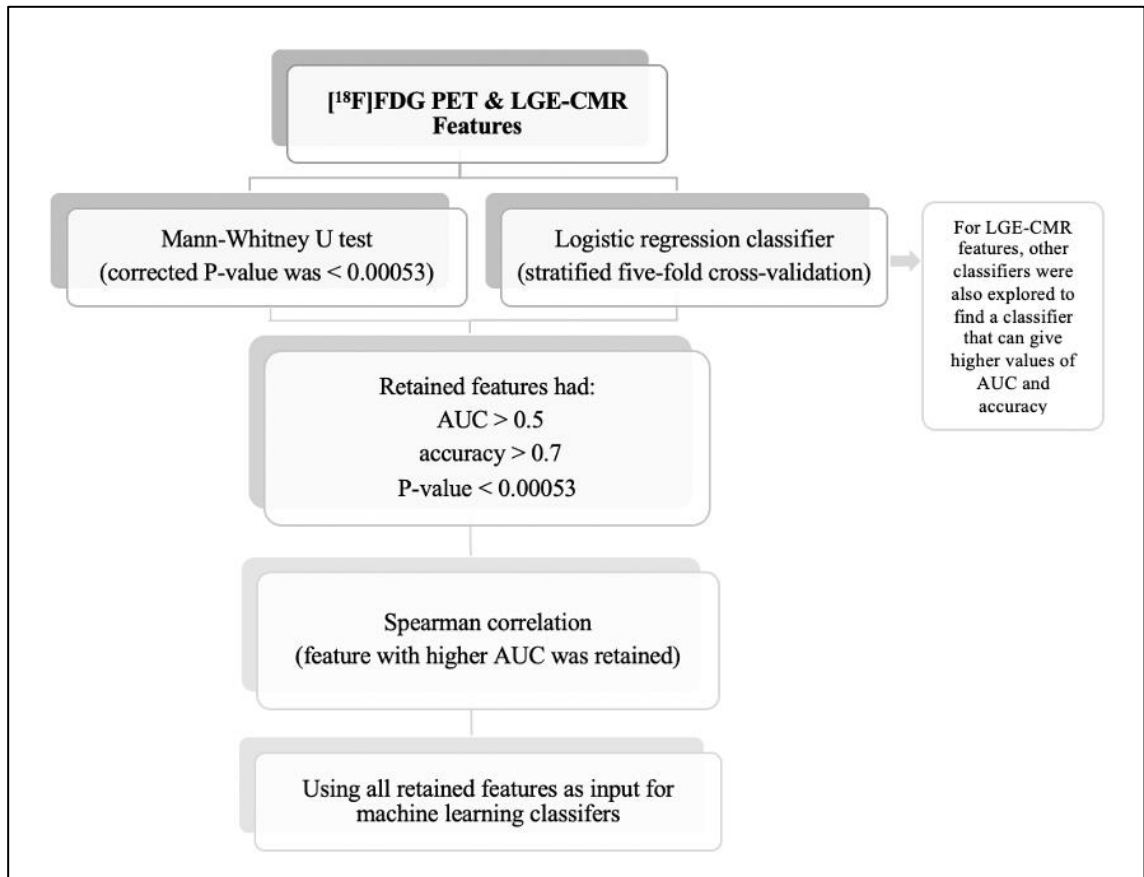


Figure 5.3 Statistical analysis workflow

5.3 Results

5.3.1 Individual Features Assessment

The univariate analysis of individual features in each dataset revealed that [18F]FDG PET and LGE-CMR data had 5 features and 11 features, respectively, with p -values < 0.00053. For all datasets, Table 5.2 demonstrates the five best radiomic features based on the p -values.

Table 5.2 Five best radiomic features based on the p -values. GLSZM: gray level size zone matrix; GLRLM: gray level run length matrix; GLDM: gray level dependence matrix.

PET Features	p-value	LGE-CMR Features	p-value
<i>TBR_{max}</i>	1.5×10^{-11}	<i>glszm_Small Area Low Gray Level Emphasis</i>	7.2×10^{-7}
<i>glszm_Large Area High Gray Level Emphasis</i>	2×10^{-5}	<i>gldm_Dependence Non-Uniformity</i>	7.3×10^{-7}
<i>glrlm_Gray Level Non-Uniformity</i>	1×10^{-4}	<i>gldm_Small Dependence Low Gray Level Emphasis</i>	8.8×10^{-7}
<i>gldm_Gray Level Non-Uniformity</i>	1.2×10^{-4}	<i>glszm_Low Gray Level Zone Emphasis</i>	1.3×10^{-6}
<i>glszm_Zone Variance</i>	1.8×10^{-4}	<i>glrlm_Run Length Non-Uniformity</i>	8.8×10^{-6}

5.3.2 Feature Selection

The logistic regression was applied on individual features. [¹⁸F]FDG PET and LGE-CMR datasets had only three and five features, respectively, conforming to the inclusion criteria (p -value < 0.00053, AUC > 0.5 and accuracy > 0.7). Those features that met the inclusion criteria were again screened based on correlation. To detect the correlated features, a correlation test was conducted. Features with a higher AUC were retained. The number of selected features in the PET and LGE-CMR features decreased to two uncorrelated features for each one. Table 5.3 presents AUC and accuracy values, along with their corresponding 95% CI, for the uncorrelated features in each dataset.

Scrutinizing PET features in greater detail, *TBR_{max}* conveyed high AUC and

accuracy with relatively small confidence intervals while *glszm_Zone Variance* had acceptable values but with large confidence intervals. Creating a signature using these uncorrelated features as input for machine learning classifiers improved the performance. Random Forest was the best one (95% CI AUC 0.95 – 1.00: accuracy 0.83 – 0.99). The performance of all machine learning classifiers is displayed in Table 5.4.

Table 5.3 Areas under the curve (AUCs) and accuracies (ACC) of uncorrelated features. CI: confidence interval; GLSZM: gray level size zone matrix; GLDM: gray level dependence matrix; GLRLM: gray level run length matrix.

	Feature	ACC	ACC CI	AUC	AUC CI	Sensitivity	Specificity
PET	<i>TBR_{max}</i>	0.89	0.07	0.95	0.09	0.91	0.88
	<i>glszm_Zone Variance</i>	0.71	0.15	0.69	0.22	0.49	0.90
LGE-CMR	<i>gldm_Dependence</i>	0.75	0.06	0.87	0.05	0.69	0.80
	<i>Non-Uniformity</i>						
	<i>glrlm_Long Run High Gray Level Emphasis</i>	0.71	0.15	0.78	0.21	0.57	0.83

Table 5.4 The machine learning classifiers performance of PET joint features with 95% confidence intervals (CI). ACC: Accuracy; AUC: area under the curve.

Machine Learning Classifier	ACC	ACC CI	AUC	AUC CI	Sensitivity	Specificity
Random Forest	0.91	0.08	0.98	0.03	0.94	0.90
Logistic Regression	0.87	0.09	0.96	0.07	0.83	0.90
Support Vector Machine	0.63	0.08	0.56	0.31	0.26	0.95
Decision Tree	0.88	0.03	0.88	0.04	0.86	0.93
Gaussian Process	0.61	0.10	0.70	0.16	0.40	0.80
Stochastic Gradient Descent	0.48	0.03	0.71	0.19	0.80	0.20
Perceptron	0.44	0.07	0.69	0.22	0.94	0.00
Passive Aggressive	0.63	0.12	0.69	0.31	0.40	1.00
Neural Network	0.53	0.17	0.64	0.12	0.97	0.20
K-neighbors	0.69	0.04	0.73	0.11	0.60	0.78

For LGE-CMR features, *gldm_Dependence Non-Uniformity* presented good AUC and accuracy (95% CI AUC 0.82– 0.92: accuracy 0.69– 0.81). Examples of PET and CMR images for CS and PC with the related features are shown in Figure 5.4. However, to further improve the AUC and accuracy findings of LGE-CMR, additional measures were taken. After applying many other classifiers besides logistic regression, Support Vector Machine showed six features following the criteria, Table 5.5. The total number of features were not normalised before applying Support Vector Machine because the focus was on the relative relationships between the data points rather than their absolute

values or scales. Additionally, normalisation had the potential to modify the original distances and influence the underlying relationships within the data, which are crucial for distance-based algorithms. The retained features were used as input for machine learning classifiers, and Random Forest proved to have the greatest AUC and accuracy values (95% CI AUC 0.82 – 1.00: accuracy 0.73 – 0.95). The machine learning classifiers performance for joint features are shown in Table 5.6.




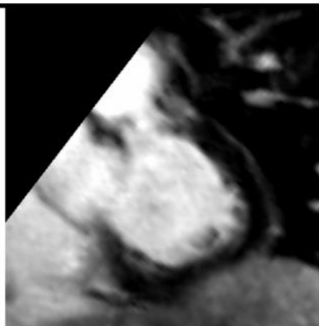
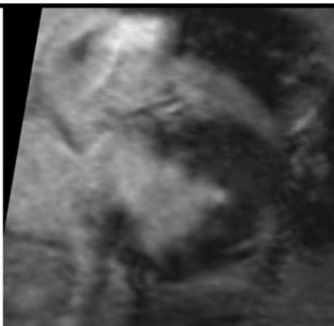
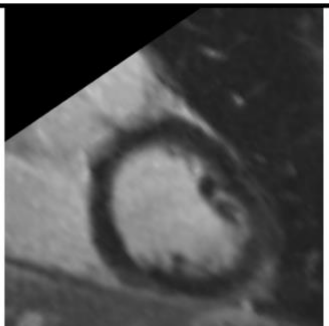
	CS (case 1)	CS (case 2)	PC
PET	 TBR _{max} = 2.44 ↓	 TBR _{max} = 5.76 ↑	 TBR _{max} = 4.25 ↑
CMR	 gldm_DNU= 134.20 ↓	 gldm_DNU= 361.70 ↑	 gldm_DNU= 277.24 ↑

Figure 5.4 Two cases of PET/CMR cardiac sarcoidosis (CS) and one case of post-COVID-19 (PC) associated with the best-performance features. PET/CMR CS (case 1) has significantly lower values than PC values, whereas (case 2) has values within the range of PC patients' values, potentially leading to a misdiagnosis. Display intensity of PET images is from 0 to 6.

Table 5.5 Areas under the curve (AUCs) and accuracies (ACC) of uncorrelated features using the Support Vector Machine (SVM) that used to create a CMR signature. CI: confidence interval; GLSZM: gray level size zone matrix; GLRLM: gray level run length matrix; GLDM: gray level dependence matrix.

Feature	ACC	ACC CI	AUC	AUC CI	Sensitivity	Specificity
<i>glszm_Low Gray Level Zone Emphasis</i>	0.72	0.18	0.83	0.17	0.66	0.78
<i>glrlm_Run Entropy</i>	0.72	0.14	0.80	0.14	0.77	0.68
<i>glszm_Small Area Low Gray Level Emphasis</i>	0.81	0.12	0.79	0.17	0.63	0.98
<i>gldm_Dependence Non-Uniformity</i>	0.73	0.09	0.78	0.11	0.69	0.78
<i>gldm_Small Dependence Low Gray Level Emphasis</i>	0.76	0.12	0.77	0.17	0.60	0.90
<i>glrlm_Gray Level Non-Uniformity</i>	0.71	0.11	0.66	0.16	0.46	0.93

SVM Classifier

Table 5.6 The machine learning classifiers performance of LGE-CMR joint features with 95% confidence intervals (CI). ACC: Accuracy; AUC: area under the curve.

Machine Learning Classifier	ACC	ACC CI	AUC	AUC CI	Sensitivity	Specificity
Random Forest	0.84	0.11	0.91	0.09	0.77	0.93
Logistic Regression	0.77	0.07	0.88	0.07	0.74	0.80
Support Vector Machine	0.72	0.15	0.79	0.15	0.51	0.90
Decision Tree	0.75	0.11	0.75	0.11	0.74	0.75
Gaussian Process	0.53	0.09	0.52	0.04	0.20	0.83
Stochastic Gradient Descent	0.48	0.03	0.62	0.09	0.49	0.53
Perceptron	0.55	0.03	0.24	0.15	0.03	1.00
Passive Aggressive	0.49	0.04	0.58	0.34	0.54	0.45
Neural Network	0.68	0.19	0.78	0.26	0.71	0.33
K-neighbors	0.68	0.08	0.69	0.09	0.60	0.75

5.4 Discussion

CS is an inflammatory disease with an unknown cause. To aid in the diagnostic process, advanced imaging techniques like [^{18}F]FDG PET and LGE-CMR are recommended. [^{18}F]FDG PET is utilised in suspected CS cases due to its ability to detect glucose uptake by active inflammatory cells in sarcoid granulomas while LGE-CMR can identify scar tissue that may indicate inactive CS. However, both techniques have limitations that contribute to their lack of specificity for CS. This study focused on investigating the potential of [^{18}F]FDG PET and LGE-CMR radiomic features in differentiating CS from other causes of myocardial inflammation, specifically in patients with post-COVID-19 symptoms related to the heart.

After applying several steps to filter the radiomic features of PET images, TBR_{max} succeeded in being the best performed feature. TBR_{max} was able to discriminate about 90% of CS cases from PC cases. Most of the CS cases had a TBR_{max} range between 1 to 3 while PC cases had higher values. This result is supported by other studies that revealed similar range values of TBR_{max} in CS patients, which were between 1 to 3 (12, 23). To some extent, TBR_{max} can make fair comparisons between institutions by looking at the equation of extracting their values, which essentially means a blood uptake correction (29). Although TBR_{max} has successfully discriminated about 90% of cases, there is still about 10% of cases that have been misdiagnosed such as the TBR_{max} value of case 2 in Figure 5.4 which provided values that were approximately similar to those of PC patients. $glszm_Zone\ Variance$ was the second-best performing feature but had significant error bars that made it unreliable.

However, the PET features performance improved significantly after using joint features as input for machine learning classifiers, especially Random Forest classifier.

For the radiomic features of CMR images, *gldm_Dependence Non-Uniformity* presented acceptable results, but with some errors. About 67.5% of CS patients had values less than 88.5% of PC patients (average value in CS = 205.8, average value in PC = 323.4). This measure determines the degree of similarity of dependence within an ROI (24). Therefore, the higher values of *gldm_Dependence Non-Uniformity* indicate a greater level of heterogeneity. As an interpretation of the values of each group, CS appears to have a lower variance than PC. However, it is recommended that this feature should be interpreted cautiously because it may contain errors especially for CS group, as third of CS patients had values similar to the majority of PC patients. In Figure 5.4, there was a big difference between the feature values in the first and second cases, and the second case even gave higher values than the PC patient value.

One of the approaches followed to augment the performance of the LGE-CMR features is creating a signature (joint features) that includes all the uncorrelated features with the best AUCs and good accuracies. This step was applied to the output of the logistic regression as well as other classifiers. The signature from the Support Vector Machine illustrated great results and ameliorated the findings compared to the individual features. The individual feature of LGE-CMR dataset from logistic regression, *gldm_Dependence Non-Uniformity*, had lower AUC but smaller confidence intervals than when using the signature. Employing

the joint features gives only a little advantage in making it the superior choice for LGE-CMR dataset classification.

In our previous study (23), it was found that *gldm_Dependence Non-Uniformity* emerged as one of the top features in PET images for distinguishing between CS patients and controls. However, it exhibited larger error bars compared to TBR_{max} indicating greater variability in its measurements. In contrast, the evaluation of LGE-CMR radiomic features to differentiate between active CS and inactive CS (4) yielded different top features compared to the current study. This discrepancy can be attributed to the distinct types of comparisons conducted in each study, as well as slight variations in the methodology employed. These alterations resulted in improved outcomes of the analysis of LGE-CMR images. Radiomic analysis is affected by several factors that make the comparisons among studies are difficult. The findings across studies are not consistently replicated; instead, they often exhibit conflicting results (30, 31). This divergence in outcomes could potentially be attributed to technical factors. Efforts should be made to minimise variation up to the reconstruction step to ensure consistency. It is crucial to avoid introducing variation in factors that occur after reconstruction whenever possible. This entails making consistent choices, such as employing the same image segmentation algorithm and utilising a uniform discretisation scheme for all the data (32). In addition, the higher the resolution and number of voxels can impact certain radiomic features by inflating their values (33). By mitigating variability at these stages, the reliability and comparability of the results can be enhanced.

Several limitations are associated with this study. First, all studies evaluating patients with suspected CS have well-known limitations due to the absence of a standard for diagnosing the condition. However, it is possible to detect CS more effectively by combining data from both CMR and PET. This patient cohort routinely did not undergo endomyocardial biopsies. It is, however, difficult to rule out CS with an endomyocardial biopsy due to its low sensitivity and high sampling error rate because of its focal distribution (34). In addition, considering the size of the sample, further studies are needed to verify this conclusion to avoid overfitting and type I errors. This issue was reduced by applying the Bonferroni correction. Furthermore, validating the AI approach on a larger and diverse patient population as well as normal controls would indeed enhance the robustness and applicability of the results. Moreover, no automated segmentation was performed, and reference segmentation was not provided in this study.

The new knowledge gained from this study is that radiomic analysis can enhance the objectivity and complementarity of PET and CMR in identifying CS from PC. PET-based analyses could effectively differentiate CS from PC. The PET joint features demonstrated high performance that can be used alone without resorting to CMR. However, CMR-based analysis is helpful when PET images suffer from failed suppression of the physiological uptake of [^{18}F]FDG in the myocardium (3). The results may vary from one institution to another due to different scanning procedures and protocols, and characteristics of each scanner. However, the methodology is straightforward and transferable to PET/CT-only and MR-only studies.

5.5 Conclusion

This work adds to the growing evidence that radiomic analysis may assist [¹⁸F]FDG PET and LGE-CMR to precisely discern cardiac sarcoidosis with a specific focus on TBR_{max} . These features hold promise for heightening the accuracy of diagnoses. Nonetheless, more research is warranted to validate and refine these results and guarantee their wider clinical applicability.

References

1. Lehtonen, J., Uusitalo, V., Pöyhönen, P., Mäyränpää, M.I. and Kupari, M. Cardiac sarcoidosis: phenotypes, diagnosis, treatment, and prognosis. *European Heart Journal*. 2023, **44**(17), pp.1495-1510.
2. Divakaran, S., Stewart, G.C., Lakdawala, N.K., Padera, R.F., Zhou, W., Desai, A.S., Givertz, M.M., Mehra, M.R., Kwong, R.Y. and Hedgire, S.S. Diagnostic accuracy of advanced imaging in cardiac sarcoidosis: an imaging-histologic correlation study in patients undergoing cardiac transplantation. *Circulation: Cardiovascular Imaging*. 2019, **12**(6), p.e008975.
3. Osborne, M.T., Hulten, E.A., Murthy, V.L., Skali, H., Taqueti, V.R., Dorbala, S., DiCarli, M.F. and Blankstein, R. Patient preparation for cardiac fluorine-18 fluorodeoxyglucose positron emission tomography imaging of inflammation. *Journal of Nuclear Cardiology*. 2017, **24**(1), pp.86-99.
4. Mushari, N.A., Soultanidis, G., Duff, L., Trivieri, M.G., Fayad, Z.A., Robson, P.M. and Tsoumpas, C. Exploring the Utility of Cardiovascular Magnetic Resonance Radiomic Feature Extraction for Evaluation of Cardiac Sarcoidosis. *Diagnostics*. 2023, **13**(11), p.1865.
5. Smedema, J.-P., Snoep, G., van Kroonenburgh, M.P., van Geuns, R.-J., Dassen, W.R., Gorgels, A.P. and Crijns, H.J. Evaluation of the accuracy of gadolinium-enhanced cardiovascular magnetic resonance in the diagnosis of cardiac sarcoidosis. *Journal of the American College of Cardiology*. 2005, **45**(10), pp.1683-1690.
6. Ichinose, A., Otani, H., Oikawa, M., Takase, K., Saito, H., Shimokawa, H. and Takahashi, S. MRI of cardiac sarcoidosis: basal and subepicardial localization of myocardial lesions and their effect on left ventricular function. *American Journal of Roentgenology*. 2008, **191**(3), pp.862-869.

7. Nakamori, S., Neisius, U., Nezafat, M., Jang, J., Ngo, L.H., Rodriguez, J., Manning, W.J. and Nezafat, R. Multiparametric Mapping Approach for Detection of Cardiac Involvement in Systemic Sarcoidosis. *Cardiovascular Imaging*. 2020, **13**(9), pp.2058-2060.
8. Skali, H., Schulman, A.R. and Dorbala, S. 18 F-FDG PET/CT for the assessment of myocardial sarcoidosis. *Current cardiology reports*. 2013, **15**(4), p.370.
9. Aggarwal, N.R., Snipelisky, D., Young, P.M., Gersh, B.J., Cooper, L.T. and Chareonthaitawee, P. Advances in imaging for diagnosis and management of cardiac sarcoidosis. *European Heart Journal-Cardiovascular Imaging*. 2015, **16**(9), pp.949-958.
10. Birnie, D.H., Sauer, W.H., Bogun, F., Cooper, J.M., Culver, D.A., Duvernoy, C.S., Judson, M.A., Kron, J., Mehta, D. and Nielsen, J.C. HRS expert consensus statement on the diagnosis and management of arrhythmias associated with cardiac sarcoidosis. *Heart rhythm*. 2014, **11**(7), pp.1304-1323.
11. Terasaki, F. and Yoshinaga, K. New guidelines for diagnosis of cardiac sarcoidosis in Japan. *Annals of Nuclear Cardiology*. 2017, **3**(1), pp.42-45.
12. Dweck, M.R., Abgral, R., Trivieri, M.G., Robson, P.M., Karakatsanis, N., Mani, V., Palmisano, A., Miller, M.A., Lala, A. and Chang, H.L. Hybrid magnetic resonance imaging and positron emission tomography with fluorodeoxyglucose to diagnose active cardiac sarcoidosis. *JACC: Cardiovascular Imaging*. 2018, **11**(1), pp.94-107.
13. Vita, T., Okada, D.R., Veillet-Chowdhury, M., Bravo, P.E., Mullins, E., Hulten, E., Agrawal, M., Madan, R., Taqueti, V.R. and Steigner, M. Complementary value of cardiac magnetic resonance imaging and positron emission tomography/computed tomography in the assessment of cardiac sarcoidosis. *Circulation: Cardiovascular Imaging*. 2018, **11**(1), p.e007030.

14. Rainey, C., O'Regan, T., Matthew, J., Skelton, E., Woznitza, N., Chu, K.-Y., Goodman, S., McConnell, J., Hughes, C., Bond, R. and Malamateniou, C. An insight into the current perceptions of UK radiographers on the future impact of AI on the profession: A cross-sectional survey. *Journal of Medical Imaging and Radiation Sciences*. 2022, **53**(3), pp.347-361.
15. Duff, L., Scarsbrook, A.F., Mackie, S.L., Froom, R., Bailey, M., Morgan, A.W. and Tsoumpas, C. A methodological framework for AI-assisted diagnosis of active aortitis using radiomic analysis of FDG PET–CT images: Initial analysis. *Journal of Nuclear Cardiology*. 2022, pp.1-17.
16. Tingen, H.S.A., van Praagh, G.D., Nienhuis, P.H., Tubben, A., van Rijsewijk, N.D., Ten Hove, D., Mushari, N.A., Martinez-Lucio, T.S., Mendoza-Ibañez, O.I. and van Sluis, J. The clinical value of quantitative cardiovascular molecular imaging: a step towards precision medicine. *The British Journal of Radiology*. 2023, p.20230704.
17. Rizzo, S., Botta, F., Raimondi, S., Origgi, D., Fanciullo, C., Morganti, A.G. and Bellomi, M. Radiomics: the facts and the challenges of image analysis. *European radiology experimental*. 2018, **2**(1), pp.1-8.
18. Basran, P.S. and Porter, I. Radiomics in veterinary medicine: Overview, methods, and applications. *Veterinary Radiology & Ultrasound*. 2022, **63**, pp.828-839.
19. Gillies, R.J., Kinahan, P.E. and Hricak, H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016, **278**(2), p.563.
20. Srinivasan, A., Wong, F., Couch, L.S. and Wang, B.X. Cardiac complications of COVID-19 in low-risk patients. *Viruses*. 2022, **14**(6), p.1322.
21. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F. and Sonka, M. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging*. 2012, **30**(9), pp.1323-1341.

22. Kikinis, R., Pieper, S.D. and Vosburgh, K.G. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In: *Intraoperative imaging and image-guided therapy*. Springer, 2014, pp.277-289.
23. Mushari, N.A., Soultanidis, G., Duff, L., Trivieri, M.G., Fayad, Z.A., Robson, P. and Tsoumpas, C. Exploring the utility of radiomic feature extraction to improve the diagnostic accuracy of cardiac sarcoidosis using FDG PET. *Frontiers in medicine*. 2022, **9**.
24. Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.-C., Pieper, S. and Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer research*. 2017, **77**(21), pp.e104-e107.
25. Leijenaar, R.T., Nalbantov, G., Carvalho, S., Van Elmpt, W.J., Troost, E.G., Boellaard, R., Aerts, H.J., Gillies, R.J. and Lambin, P. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Scientific reports*. 2015, **5**(1), p.11075.
26. Duron, L., Balvay, D., Vande Perre, S., Bouchouicha, A., Savatovsky, J., Sadik, J.-C., Thomassin-Naggara, I., Fournier, L. and Lecler, A. Gray-level discretization impacts reproducible MRI radiomics texture features. *PLoS One*. 2019, **14**(3), p.e0213459.
27. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. and Dubourg, V. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*. 2011, **12**, pp.2825-2830.
28. Steyerberg, E.W. Validation in prediction research: the waste by data splitting. *Journal of clinical epidemiology*. 2018, **103**, pp.131-133.
29. Akerele, M.I., Mushari, N.A., Forsythe, R.O., Syed, M., Karakatsanis, N.A., Newby, D.E., Dweck, M.R. and Tsoumpas, C. Assessment of different quantification metrics of [18 F]-NaF PET/CT images of patients

- with abdominal aortic aneurysm. *Journal of Nuclear Cardiology*. 2020, pp.1-11.
30. Apostolova, I., Ego, K., Steffen, I.G., Buchert, R., Wertz, H., Achenbach, H.J., Riedel, S., Schreiber, J., Schultz, M. and Furth, C. The asphericity of the metabolic tumour volume in NSCLC: correlation with histopathology and molecular markers. *European journal of nuclear medicine and molecular imaging*. 2016, **43**, pp.2360-2373.
 31. Cheng, N.-M., Fang, Y.-H.D., Tsan, D.-L., Hsu, C.-H. and Yen, T.-C. Respiration-averaged CT for attenuation correction of PET images—impact on PET texture features in non-small cell lung cancer patients. *PLoS One*. 2016, **11**(3), p.e0150509.
 32. Zwanenburg, A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *European journal of nuclear medicine and molecular imaging*. 2019, **46**(13), pp.2638-2655.
 33. Han, S., Woo, S., Suh, C.H., Kim, Y.J., Oh, J.S. and Lee, J.J. A systematic review of the prognostic value of texture analysis in 18 F-FDG PET in lung cancer. *Annals of nuclear medicine*. 2018, **32**, pp.602-610.
 34. Kato, S., Sakai, Y., Okabe, A., Kawashima, Y., Kuwahara, K., Shiogama, K., Abe, M., Ito, H. and Morimoto, S.i. Histology of cardiac sarcoidosis with novel considerations arranged upon a pathologic basis. *Journal of Clinical Medicine*. 2022, **11**(1), p.251.

Supplementary Materials

Supplementary Material 1:

Table 1: List of all radiomic features

First order features	
firstorder_10Percentile	firstorder_Minimum
firstorder_90Percentile	firstorder_Range
firstorder_Energy	firstorder_RobustMeanAbsoluteDeviation
firstorder_Entropy	firstorder_RootMeanSquared
firstorder_InterquartileRange	firstorder_Skewness
firstorder_Kurtosis	firstorder_TotalEnergy
firstorder_Maximum	firstorder_Uniformity
firstorder_MeanAbsoluteDeviation	firstorder_Variance
firstorder_Mean	TBR _{max} (for PET only)
firstorder_Median	
GLCM features	
glcm_Autocorrelation	glcm_Idn
glcm_ClusterProminence	glcm_Imc1
glcm_ClusterShade	glcm_Imc2
glcm_ClusterTendency	glcm_InverseVariance
glcm_Contrast	glcm_JointAverage
glcm_Correlation	glcm_JointEnergy
glcm_DifferenceAverage	glcm_JointEntropy
glcm_DifferenceEntropy	glcm_MCC
glcm_DifferenceVariance	glcm_MaximumProbability
glcm_Id	glcm_SumAverage
glcm_Idm	glcm_SumEntropy

glcm_Idmn	glcm_SumSquares
GLRLM features	
glrlm_GrayLevelNonUniformity	glrlm_RunEntropy
glrlm_GrayLevelNonUniformityNormalised	glrlm_RunLengthNonUniformity
glrlm_GrayLevelVariance	glrlm_RunLengthNonUniformityNormalised
glrlm_HighGrayLevelRunEmphasis	glrlm_RunPercentage
glrlm_LongRunEmphasis	glrlm_RunVariance
glrlm_LongRunHighGrayLevelEmphasis	glrlm_ShortRunEmphasis
glrlm_LongRunLowGrayLevelEmphasis	glrlm_ShortRunHighGrayLevelEmphasis
glrlm_LowGrayLevelRunEmphasis	glrlm_ShortRunLowGrayLevelEmphasis
GLSZM features	
glszm_GrayLevelNonUniformity	glszm_SizeZoneNonUniformity
glszm_GrayLevelNonUniformityNormalised	glszm_SizeZoneNonUniformityNormalised
glszm_GrayLevelVariance	glszm_SmallAreaEmphasis
glszm_HighGrayLevelZoneEmphasis	glszm_SmallAreaHighGrayLevelEmphasis
glszm_LargeAreaEmphasis	glszm_SmallAreaLowGrayLevelEmphasis
glszm_LargeAreaHighGrayLevelEmphasis	glszm_ZoneEntropy
glszm_LargeAreaLowGrayLevelEmphasis	glszm_ZonePercentage
glszm_LowGrayLevelZoneEmphasis	glszm_ZoneVariance
GLDM features	
gldm_DependenceEntropy	gldm_LargeDependenceEmphasis

gldm_DependenceNonUniformity	gldm_LargeDependenceHighGrayLevel Emphasis
gldm_DependenceNonUniformityNormalised	gldm_LargeDependenceLowGrayLevel Emphasis
gldm_DependenceVariance	gldm_LowGrayLevelEmphasis
gldm_GrayLevelNonUniformity	gldm_SmallDependenceEmphasis
gldm_GrayLevelVariance	gldm_SmallDependenceHighGrayLevel Emphasis
gldm_HighGrayLevelEmphasis	gldm_SmallDependenceLowGrayLevel Emphasis
NGTDM features	
ngtdm_Busyness	ngtdm_Contrast
ngtdm_Coarseness	ngtdm_Strength
ngtdm_Complexity	

Chapter 6 Discussion

6.1 Overview of Results

The objective of the study in Chapter 3 was to investigate the diagnostic effectiveness of [^{18}F]FDG PET radiomic features in comparison to conventional metrics. The primary goal was to differentiate between CS patients and a control group who exhibited suspicion of CS but tested negative in PET imaging, and identify the most optimal machine learning classifier to develop an automated model. The study evaluated two manual segmentation approaches: the hot regions-only segmentation (segmentation A) and the whole left ventricular myocardium (segmentation B). Significant disparities in radiomic features were observed when employing distinct segmentation approaches. This discrepancy can likely be attributed to variations in the sizes of ROIs and the inclusion of voxel intensities within each segmentation. In addition, it should be noted that segmentation A may be influenced by observer experience, particularly in instances involving small hotspots. Moreover, the outcomes of the top-performing features were influenced by the criteria of segmentation A, introducing a level of bias that should not be overlooked. Conversely, segmentation B demonstrated greater resilience and efficacy in these regards.

Out of all the features considered, the dominance of TBR_{max} stands out prominently. Despite TBR_{max} 's sensitivity to noise and potential challenges in harmonisation across diverse scanners, imaging centres, data types, and parameters, such issues are not applicable in this study. The datasets were exclusively derived from a single scanner and institution. By applying PCA and using of five principal components to test and train the ML classifiers, A k-

neighbors and neural network classifiers showed the highest AUC and greatest accuracy for segmentation A approach. For segmentation B approach, the gaussian process classifier indicated the highest AUC and accuracy.

The primary objective of the study in chapter 4 was to explore the effectiveness of radiomic features extracted from LGE-CMR images in differentiating between active and inactive cardiac sarcoidosis. This study represented an advancement of the methodology employed in the chapter 3. To enhance reliability, the segmentation approach was chosen, favouring segmentation B over segmentation A. Furthermore, the ML classifiers were tested on individual features (referred to as signature A) as well as principal components (referred to as signature B), but it was not necessary to do so in chapter 3 since the ML classifiers performed well on principal components.

It is notable that numerous second-order features demonstrated superior performance compared to many first-order features. Given that the presence of LGE (a first-order criterion) was a prerequisite for both CS_{active} and CS_{inactive} groups, it was anticipated that first-order features might not serve as effective discriminators. However, the overall fibrosis burden, and thus LGE, could still be a meaningful indicator of active disease. Similarly, noteworthy findings emerged in the machine learning approach using signature A, where approximately half of the features belonged to the *GLCM* class. *GLCM joint entropy* emerged as the top-performing feature, exhibiting the highest accuracy and the narrowest confidence interval. The effectiveness of *GLCM joint entropy* in distinguishing between study groups can be attributed to its ability to measure the level of randomness or variance in pixel intensities within a given pixel's neighborhood

(1, 2). Slightly elevated values of this feature were observed in the CS_{active} group compared to the $CS_{inactive}$ group, indicating the presence of heterogeneity—a characteristic of this disease (3).

The study in chapter 5 examined the role of PET and LGE CMR radiomic features in distinguishing CS from other conditions marked by myocardial inflammation. Specifically, the focus was on patients exhibiting cardiac-related symptoms after COVID-19 (PC patients). The methodology in this chapter was further enhanced to include LGE CMR radiomic features. In the quest for a classifier yielding the highest values for AUC and accuracy, various classifiers were investigated beyond logistic regression.

Following a series of steps to refine the radiomic features of PET images, TBR_{max} emerged as the top-performing feature. TBR_{max} demonstrated the ability to distinguish approximately 90% of CS cases from PC cases. Notably, CS cases predominantly exhibited a TBR_{max} range between 1 to 3, while PC cases displayed higher values. Moreover, the performance of PET features witnessed a good improvement when employing joint features as input for machine learning classifiers, particularly the Random Forest classifier.

Concerning the radiomic features extracted from CMR images, *gldm_Dependence Non-Uniformity* yielded acceptable results but with some discrepancies. Approximately 67.5% of CS patients exhibited values less than 88.5% of PC patients. This metric assesses the degree of dependence similarity within an ROI, where higher values signify increased heterogeneity. In

interpreting the values for each group, it appears that CS has lower heterogeneity compared to PC. Nevertheless, caution is advised in interpreting this feature due to its errors that should not be disregarded. Furthermore, by creating joint features, the signature derived from the Support Vector Machine showcased excellent outcomes and improved the findings in comparison to individual features. Utilising joint features provides a slight advantage, making it the preferred option for LGE-CMR dataset classification. The enhanced outcomes of LGE-CMR radiomic features in this study, in contrast to the results in chapter 4, may be attributed to both the different type of comparison and the slight modifications made in the methodology.

6.2 Limitations

This project is accompanied by several limitations. Firstly, insufficient data with failed suppression of the physiological PET uptake, hindering the showcase of the pivotal advantage and importance of radiomic analysis. To address this, the same methodology can be applied to data with failed suppression, aiming to explore the role of radiomic analysis in detecting CS patients even under suboptimal preparation. In addition, as is common in investigations involving patients with suspected CS, there are inherent challenges due to the absence of a standardised diagnostic criterion for the condition. Nonetheless, the combination of data from both CMR and PET has demonstrated enhanced efficacy in detecting CS. It is noteworthy that the patient cohort in this project typically did not undergo endomyocardial biopsies. The utility of such biopsies in ruling out CS is hindered by their low sensitivity and the high sampling error rate attributed to the condition's focal distribution (4). Furthermore, given the sample size, additional studies may be imperative to validate these conclusions and

mitigate the risks of overfitting and type I errors, a concern addressed by the application of the Bonferroni correction, dimensionality reduction as well as avoiding hyperparameters tuning. Although the risk of overfitting is still possible. In addition, by mitigating the risk of type I error, there could be an increase in the risk of type II errors (5). However, due to the rarity of the disease, having more data is difficult compared to oncological studies. Therefore, this project developed training models using cross-validation as an alternative method for model evaluation. This approach allows for assessing the statistical optimism in average predictive performance, providing a more realistic estimate of the model's performance. The creation of a completely independent validation set would have significantly reduced the amount of data available for model development, potentially compromising the robustness and generalisability of the findings (6). Additionally, the absence of automated segmentation and the lack of a reference segmentation introduce further limitations. The inclusion of these components would enhance the robustness and accuracy of the findings. However, the developed models in this project can be used as a first step towards fully automated models. Future studies are imperative to assess the reproducibility of the findings across different sites, and different newer scanners with better sensitivity and resolution.

The used scanner and the implemented imaging protocol introduced certain limitations that could impact the radiomic analysis conducted in this project. The datasets were acquired using Biograph mMR PET/MR scanner, the first fully integrated PET/MR model available commercially (7). While this scanner presents advantages in integrating functional images with high soft tissue resolution, it also possesses several limitations. Standalone PET scanners

typically utilise photomultiplier tubes (PMTs) as their photodetectors. However, PMTs are susceptible to magnetic fields, posing challenges for integration into PET/MR scanners. In contrast, the Biograph mMR employs avalanche photodiodes (APDs) as an alternative photodetector due to their insensitivity to magnetic fields, robustness, and compact design (8, 9). Nevertheless, APDs are constrained by their limited gain and excess noise (10), directly influencing the accuracy of radiomic analysis. Exploring advanced scanners, such as those equipped with silicon photomultipliers (11), may enhance research findings. Additionally, the use of APDs, characterised by poor timing capabilities (timing resolution of 2.9 ns), precludes the provision of TOF information (12, 13). TOF information is vital for precise radiotracer localisation, leading to improved image quality and reduced image noise.

The PET datasets underwent reconstruction using the OP-OSEM algorithm, which, through iterative updates based on acquired projection data, aids in recovering well-defined boundaries and enhancing spatial details. This iterative approach holds promise in mitigating blurring induced by PVE. However, to specifically address PVE, additional techniques such as partial volume correction algorithms may be necessary. The OSEM algorithm, while effective, has a limitation of amplifying noise with each iteration. As a precaution, it is recommended to limit iterations to three or employ Gaussian post-filtering (14, 15), as done in this project. In addition, motion is anticipated to impact the PET images, therefore motion correction may help enhance the use of radiomic features.

6.3 Future Directions and Considerations

Future studies could address the limitations discussed above and advance the understanding by including failed suppression data, validating and automating the models as well as assessing the reproducibility of the results. Moreover, the enhancement of analysis could be achieved by employing more advanced scanners and refining imaging protocols.

For additional future exploration, the presence of radiotracers that are not affected by the physiological activity of the myocardium may provide important diagnostic value (16). ^{68}Ga -labeled somatostatin based PET imaging (e.g. [^{68}Ga]DOTATATE), which is already used in oncological applications, can be promising radiotracers to diagnose CS due to the overexpression of the subtype 2 of somatostatin receptor in sarcoid granulomas (17, 18). However, these radiotracers need more investigations, especially to evaluate in which phase of sarcoidosis (chronic or acute) they can offer greater value than standard [^{18}F]FDG. In addition, increased cell proliferation is another feature that can be seen in sarcoid granulomas as opposed to the normal myocardium which lacks proliferative activity. This can be detected using fluorothymidine ([^{18}F]FLT) (19, 20). However, there are few studies examining the efficiency of this radiotracer compared to [^{18}F]FDG.

Patlak analysis (K_i) in the context of [^{18}F]FDG PET may have been less sensitive to failed myocardial suppression when compared to other quantitative methods like SUV metrics. It is particularly advantageous in situations where there is incomplete or failed suppression of physiological [^{18}F]FDG uptake in the myocardium, which is a common challenge in cardiac imaging. Patlak analysis

is a kinetic modelling approach that focuses on the late phase of [^{18}F]FDG uptake, typically when the tracer has equilibrated between the blood and the tissue. In the myocardium, [^{18}F]FDG uptake is influenced by both perfusion and metabolism. The late phase is characterised by stable kinetics, and Patlak analysis exploits this to estimate the metabolic rate of [^{18}F]FDG (21-23). However, while Patlak analysis is more demanding than calculating SUV, it can pose inconvenience for patients. Nakajo, M. et al. (22) found that the analysis added value to the results of SUV, especially with regard to identifying clinical events of cardiac sarcoidosis such as arrhythmia. As Patlak analysis takes into account the blood input function, which helps in normalising the radiotracer uptake for variations in blood activity, applying radiomic analysis on *Ki* images instead of static images may provide additional information related to tissue characteristics and reduce potential confounding factors that may affect SUV measurements, such as differences in blood flow or injection dose. In addition, by applying a population-averaged input function, the duration of the dynamic imaging can be significantly reduced, which can increase patient throughput (24).

T1 and T2 mapping are advanced CMR techniques that may play a significant role in the assessment of cardiac sarcoidosis, especially when combined with LGE imaging. These mapping techniques provide quantitative information about tissue characteristics and can offer valuable insights into the extent and nature of myocardial involvement in cardiac sarcoidosis (25). T1 mapping measures the longitudinal relaxation time of protons in the myocardium. In the context of cardiac sarcoidosis, it provides quantitative information about tissue composition, helping to identify areas of fibrosis (26). This distinction is crucial

for understanding the heterogeneity of myocardial involvement in cardiac sarcoidosis. T1 mapping has the potential to detect pathological changes in the myocardium even before the appearance of visible structural alterations on conventional imaging. This early detection can contribute to timely intervention and management. T2 mapping is sensitive to changes in tissue water content, making it particularly useful for detecting oedema associated with active inflammation. In cardiac sarcoidosis, areas of oedema may indicate ongoing inflammatory processes. Combining T2 mapping with LGE imaging provides a comprehensive assessment of cardiac sarcoidosis. While LGE highlights areas of irreversible fibrosis, T2 mapping identifies regions with reversible inflammation, offering a more complete understanding of the disease state. In addition, T2 mapping, in conjunction with LGE, aids in differentiating between active and chronic lesions. Active lesions with inflammation may exhibit increased T2 values, while chronic lesions with fibrosis may show LGE without increased T2 (27, 28). However, it exhibits constrained sensitivity in identifying myocardial inflammation in individuals with CS and displays a restricted correlation with [¹⁸F]FDG PET imaging. Extracellular volume mapping, which quantifies the fraction of extracellular volume (potentially elevated due to scar or oedema), requires further validation for its utility in detecting CS, particularly in patients without LGE (29). The integration of radiomic analysis using T1/T2 mapping and LGE CMR images has the potential to offer a comprehensive perspective on CS assessment. This integration may even present an opportunity to reduce reliance on PET imaging, thereby mitigating the associated risks of radiation exposure. This is particularly beneficial for individuals requiring repeat examinations.

6.4 Conclusion

In conclusion, the thesis advanced the diagnostic methodologies for CS and related myocardial inflammation. The incorporation of radiomic features from both PET and CMR, along with the exploration of various classifiers, provides a comprehensive perspective not presented previously for CS patients.

References

1. Zwanenburg, A., Vallières, M., Abdalah, M.A., Aerts, H.J., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R.J. and Boellaard, R. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020, **295**(2), pp.328-338.
2. Avard, E., Shiri, I., Hajianfar, G., Abdollahi, H., Kalantari, K.R., Houshmand, G., Kasani, K., Bitarafan-Rajabi, A., Deevband, M.R. and Oveisi, M. Non-contrast Cine Cardiac Magnetic Resonance image radiomics features and machine learning algorithms for myocardial infarction detection. *Computers in biology and medicine*. 2022, **141**, p.105145.
3. Tana, C., Mantini, C., Donatiello, I., Mucci, L., Tana, M., Ricci, F., Cipollone, F. and Giamberardino, M.A. Clinical Features and Diagnosis of Cardiac Sarcoidosis. *Journal of Clinical Medicine*. 2021, **10**(9), p.1941.
4. Kato, S., Sakai, Y., Okabe, A., Kawashima, Y., Kuwahara, K., Shiogama, K., Abe, M., Ito, H. and Morimoto, S.i. Histology of cardiac sarcoidosis with novel considerations arranged upon a pathologic basis. *Journal of Clinical Medicine*. 2022, **11**(1), p.251.
5. Kaur, P. and Stoltzfus, J. Type I, II, and III statistical errors: A brief overview. *International Journal of Academic Medicine*. 2017, **3**(2), pp.268-270.
6. Steyerberg, E.W. Validation in prediction research: the waste by data splitting. *Journal of clinical epidemiology*. 2018, **103**, pp.131-133.
7. Healthineers, S. *Biograph mMR Simultaneous MR and PET Imaging*. [Online]. 2024. [Accessed 1 Jan]. Available from: <https://www.siemens-healthineers.com/en-us/magnetic-resonance-imaging/mr-pet-scanner/biograph-mmr>

8. Wu, Y., Catana, C., Farrell, R., Dokhale, P.A., Shah, K.S., Qi, J. and Cherry, S.R. PET performance evaluation of an MR-compatible PET insert. *IEEE Transactions on Nuclear Science*. 2009, **56**(3), pp.574-580.
9. Maramraju, S.H., Smith, S.D., Junnarkar, S.S., Schulz, D., Stoll, S., Ravindranath, B., Purschke, M.L., Rescia, S., Southeikal, S. and Pratte, J.-F. Small animal simultaneous PET/MRI: initial experiences in a 9.4 T microMRI. *Physics in Medicine & Biology*. 2011, **56**(8), p.2459.
10. Jung, J.H., Choi, Y. and Im, K.C. PET/MRI: technical challenges and recent advances. *Nuclear medicine and molecular imaging*. 2016, **50**, pp.3-12.
11. Schaart, D.R., Seifert, S., Vinke, R., van Dam, H.T., Dendooven, P., Löhner, H. and Beekman, F.J. LaBr₃: Ce and SiPMs for time-of-flight PET: achieving 100 ps coincidence resolving time. *Physics in Medicine & Biology*. 2010, **55**(7), p.N179.
12. Delso, G., Fürst, S., Jakoby, B., Ladebeck, R., Ganter, C., Nekolla, S.G., Schwaiger, M. and Ziegler, S.I. Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner. *Journal of Nuclear Medicine*. 2011, **52**(12), pp.1914-1922.
13. Vandenberghe, S. and Marsden, P.K. PET-MRI: a review of challenges and solutions in the development of integrated multimodality imaging. *Physics in Medicine & Biology*. 2015, **60**(4), p.R115.
14. Hamill, J. and Bruckbauer, T. Iterative reconstruction methods for high-throughput PET tomographs. *Physics in Medicine & Biology*. 2002, **47**(15), p.2627.
15. Ahn, S., Ross, S.G., Asma, E., Miao, J., Jin, X., Cheng, L., Wollenweber, S.D. and Manjeshwar, R.M. Quantitative comparison of OSEM and penalized likelihood image reconstruction using relative difference penalties for clinical PET. *Physics in Medicine & Biology*. 2015, **60**(15), p.5733.

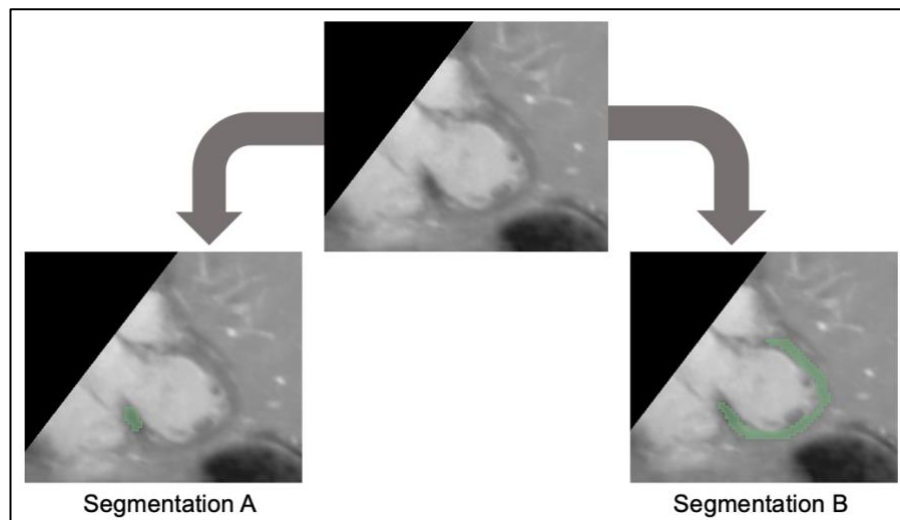
16. Chareonthaitawee, P., Beanlands, R.S., Chen, W., Dorbala, S., Miller, E.J., Murthy, V.L., Birnie, D.H., Chen, E.S., Cooper, L.T. and Tung, R.H. *Joint SNMMI–ASNC expert consensus document on the role of 18 F-FDG PET/CT in cardiac sarcoid detection and therapy monitoring*. Springer. 2017.
17. Slart, R.H., Koopmans, K.-P., van Geel, P.P., Kramer, H., Groen, H.J., Gan, C.T.-J., Prakken, N.H., Glaudemans, A.W. and Nossent, G.D. Somatostatin receptor based hybrid imaging in sarcoidosis. *European journal of hybrid imaging*. 2017, **1**(1), p.7.
18. Lee, H., Schubert, E.K., Vidula, M.K., Pryma, D.A., Marchlinski, F.E., Goldberg, L.R., Clancy, C.B., Rossman, M.D., DiCarli, M.F. and Bravo, P.E. Potential clinical utility of 68Ga-DOTATATE PET/CT for detection and response assessment in cardiac sarcoidosis. *Journal of Nuclear Cardiology*. 2023, **30**(3), pp.1075-1087.
19. Martineau, P., Pelletier-Galarneau, M., Juneau, D., Leung, E., Birnie, D. and Beanlands, R. Molecular imaging of cardiac sarcoidosis. *Current Cardiovascular Imaging Reports*. 2018, **11**(3), p.6.
20. Saric, P., Young, K.A., Rodriguez-Porcel, M. and Chareonthaitawee, P. PET imaging in cardiac sarcoidosis: a narrative review with focus on novel PET tracers. *Pharmaceuticals*. 2021, **14**(12), p.1286.
21. Dweck, M.R., Abgral, R., Trivieri, M.G., Robson, P.M., Karakatsanis, N., Mani, V., Palmisano, A., Miller, M.A., Lala, A. and Chang, H.L. Hybrid magnetic resonance imaging and positron emission tomography with fluorodeoxyglucose to diagnose active cardiac sarcoidosis. *JACC: Cardiovascular Imaging*. 2018, **11**(1), pp.94-107.
22. Nakajo, M., Ojima, S., Kawakami, H., Tani, A., Hirayama, A., Jinguji, M., Kubozono, T., Ohishi, M. and Yoshiura, T. Value of Patlak Ki images from 18F-FDG-PET/CT for evaluation of the relationships between disease activity and clinical events in cardiac sarcoidosis. *Scientific reports*. 2021, **11**(1), p.2729.

23. Régis, C., Benali, K. and Rouzet, F. FDG PET/CT imaging of sarcoidosis. In: *Seminars in Nuclear Medicine*: Elsevier, 2023, pp.258-272.
24. van Sluis, J., van Snick, J.H., Brouwers, A.H., Noordzij, W., Dierckx, R.A., Borra, R.J., Lammertsma, A.A., Glaudemans, A.W., Slart, R.H. and Yaqub, M. Shortened duration whole body 18F-FDG PET Patlak imaging on the Biograph Vision Quadra PET/CT using a population-averaged input function. *EJNMMI physics*. 2022, **9**(1), pp.1-14.
25. Munoz, C., Schneider, A., Botnar, R.M. and Prieto, C. Recent advances in PET-MRI for cardiac sarcoidosis. *Frontiers in Nuclear Medicine*. 2022, **2**, p.1032444.
26. Burt, J.R., Zimmerman, S.L., Kamel, I.R., Halushka, M. and Bluemke, D.A. Myocardial T1 mapping: techniques and potential applications. *Radiographics*. 2014, **34**(2), pp.377-395.
27. Crouser, E.D., Ono, C., Tran, T., He, X. and Raman, S.V. Improved detection of cardiac sarcoidosis using magnetic resonance with myocardial T2 mapping. *American journal of respiratory and critical care medicine*. 2014, **189**(1), pp.109-112.
28. O'Brien, A.T., Gil, K.E., Varghese, J., Simonetti, O.P. and Zareba, K.M. T2 mapping in myocardial disease: a comprehensive review. *Journal of Cardiovascular Magnetic Resonance*. 2022, **24**(1), pp.1-25.
29. Gilotra, N.A., Griffin, J.M., Pavlovic, N., Houston, B.A., Chasler, J., Goetz, C., Chrispin, J., Sharp, M., Kasper, E.K. and Chen, E.S. Sarcoidosis-related cardiomyopathy: current knowledge, challenges, and future perspectives state-of-the-art review. *Journal of cardiac failure*. 2022, **28**(1), pp.113-132.

Thesis Addendum

This addendum provides supplementary information concerning the published chapters (3, 4, and 5):

- The ethical approval for all studies in this thesis (**Chapter 3, 4 and 5**) has been granted by the School of Medicine Research Ethics Committee (Reference number: MREC 20-010).
- An explanation of segmentation A and B in **Chapter 3** can be shown below:



- In all studies (**Chapter 3, 4 and 5**), the machine learning classifiers utilised default parameters as described in scikit-learn, version 0.23.2 (1).

ML Classifier	Description	Parameters
Logistic Regression	Models the binary response through representing log odds as a linear combination of one or multiple independent variables (2)	<i>(penalty='l2', *, dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='lbfgs', max_iter=100, multi_class='auto', verbose=0, warm_start=False, n_jobs=None, l1_ratio=None)</i>
Support Vector Machine	Finding the optimal hyperplane that maximises the margin between different classes in the transformed feature space (3)	<i>(*, C=1.0, kernel='rbf', degree=3, gamma='scale', coef0=0.0, shrinking=True, probability=True, tol=0.001, cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', break_ties=False, random_state=None)</i>
Decision Tree	Splitting the data into subsets based on the value of input features, creating a tree-like model of decisions (4)	<i>(*, criterion='gini', splitter='best', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features=None, random_state=None, max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, class_weight=None, presort='deprecated', ccp_alpha=0.0)</i>
Random Forest	Constructing multiple decision trees during training and outputting the class that is the mode of the classes (for classification) of the individual trees (5)	<i>(n_estimators=100, *, criterion='gini', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=None, random_state=None, verbose=0, warm_start=False, class_weight=None, ccp_alpha=0.0, max_samples=None)</i>

Gaussian Process Classifier	Uses a kernel function to define the similarity between points and provides probabilistic predictions, offering both mean values and uncertainties (6)	<i>(kernel=kernel, *, optimizer='fmin_l_bfgs_b', n_restarts_optimizer=0, max_iter_predict=100, warm_start=False, copy_X_train=True, random_state=None, multi_class='one_vs_rest', n_jobs=None)</i>
Stochastic Gradient Descent	Using the stochastic gradient descent optimisation method, this approach updates the model parameters incrementally for each training example (7)	<i>(loss='hinge', *, penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=1000, tol=0.001, shuffle=True, verbose=0, epsilon=0.1, n_jobs=None, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5, early_stopping=False, validation_fraction=0.1, n_iter_no_change=5, class_weight=None, warm_start=False, average=False)</i>
Perceptron Classifier	Models the relationship between input features and a binary outcome using a single-layer neural network by adjusting weights iteratively based on misclassified examples to find a hyperplane that separates the two classes (8)	<i>(* , penalty=None, alpha=0.0001, fit_intercept=True, max_iter=1000, tol=0.001, shuffle=True, verbose=0, eta0=1.0, n_jobs=None, random_state=0, early_stopping=False, validation_fraction=0.1, n_iter_no_change=5, class_weight=None, warm_start=False)</i>
Passive Aggressive Classifier	Making updates only when a prediction error occurs (aggressive) and otherwise leaves the model unchanged (passive) (9)	<i>(* , C=1.0, fit_intercept=True, max_iter=1000, tol=0.001, early_stopping=False, validation_fraction=0.1, n_iter_no_change=5, shuffle=True, verbose=0, loss='hinge', n_jobs=None, random_state=None, warm_start=False, class_weight=None, average=False)</i>
Neural Network Classifier	Inspired by the structure and function of the human brain. It consists of interconnected layers of nodes (neurons) that work together to classify input data. Each neuron receives input, processes it using an activation function, and passes the result to the next layer (10)	<i>(hidden_layer_sizes=(100,), activation='relu', *, solver='adam', alpha=0.0001, batch_size='auto', learning_rate='constant', learning_rate_init=0.001, power_t=0.5, max_iter=200, shuffle=True, random_state=None, tol=0.0001, verbose=False, warm_start=False, momentum=0.9, nesterovs_momentum=True, early_stopping=False, validation_fraction=0.1, beta_1=0.9, beta_2=0.999,</i>

		<i>epsilon=1e-08, n_iter_no_change=10, max_fun=15000)</i>
K-neighbours Classifier	Classifying a data point based on the majority class among its 'K' nearest neighbours. The value of 'K' is a critical parameter that determines the number of neighbours to consider (10)	<i>(n_neighbors=5, *, weights='uniform', algorithm='auto', leaf_size=30, p=2, metric='minkowski', metric_params=None, n_jobs=None, **kwargs)</i>

References

1. Scikit-learn. *scikit-learn user guide Release 0.23.2*. [Online]. 2020. [Accessed 3 March 2024]. Available from: <https://scikit-learn.org/0.23/downloads/scikit-learn-docs.pdf>
2. Verzani, J. *Using R for introductory statistics*. Chapman and Hall/CRC, 2004.
3. Vapnik, V. *The nature of statistical learning theory*. Springer science & business media, 2013.
4. Quinlan, J.R. Learning decision tree classifiers. *ACM Computing Surveys (CSUR)*. 1996, **28**(1), pp.71-72.
5. Breiman, L. Random forests. *Machine learning*. 2001, **45**, pp.5-32.
6. Seeger, M. Gaussian processes for machine learning. *International journal of neural systems*. 2004, **14**(02), pp.69-106.
7. Ketkar, N. and Ketkar, N. Stochastic gradient descent. *Deep learning with Python: A hands-on introduction*. 2017, pp.113-132.
8. Freund, Y. and Schapire, R.E. Large margin classification using the perceptron algorithm. In: *Proceedings of the eleventh annual conference on Computational learning theory*, 1998, pp.209-217.
9. Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S., Singer, Y. and Warmuth, M.K. Online passive-aggressive algorithms. *Journal of Machine Learning Research*. 2006, **7**(3).
10. Bishop, C.M. Pattern recognition and machine learning. *Springer google schola*. 2006, **2**, pp.1122-1128.