THE USE OF AN AUTOMATED SARCOPENIA MEASUREMENT TOOL AS A PREDICTOR OF OUTCOME IN PATIENTS WITH ABDOMINAL AORTIC ANEURYSMS (SPAR)

Thesis submitted to Imperial College London

for the degree of MD (Research)

Mr Abhilash Sudarsanam

Department of Surgery & Cancer

SUPERVISORS

Professor Alun H Davies

Academic Section of Vascular Surgery

Division of Surgery

Department of Surgery & Cancer

Faculty of Medicine

Imperial College London

Professor Andrea Rockall

Academic Section of Radiology

Faculty of Medicine

Imperial College London

17/11/2024

STATEMENT OF ORIGINALITY

The material presented in this report is the original work of the author. Where results or diagrams have been reproduced from the work of others, the source is clearly stated.

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

ACKNOWLEDGEMENTS

I would like to thank Professor Davies, my supervisor, for providing me with this opportunity to undertake a period of formal research and utilising his valuable time to provide excellent supervision and support during this period.

I would like to thank Professor Rockall, my co-supervisor, for introducing me to the concepts of automated segmentation and it's use in clinical research, and providing enthusiastic supervision and support throughout this process.

I would also like to thank the extended clinical research team in the academic section of radiology for helping me process and analyse CT scans of patients for the purposes of this research project, and providing support with segmentation and image analysis.

I am very grateful to the intercalating BSc students, Mr M Yanai and Miss H Sekhon, for their hard work and efforts in the undertaking of the systematic review / meta-analysis and feasibility study of quadriceps muscle as a surrogate measure of sarcopenia, respectively.

I am also incredibly grateful to Ms S Peerbux (clinical trial manager) who was instrumental in the set-up of this study and obtaining ethics approval to commence this study.

I would also like to thank Miss S Salim (post-doctoral research fellow) who helped me undertake the statistical analysis involved in this study.

Finally, I would like to thank all my day-to-day colleagues in the academic department who made the overall process enjoyable, and provided regular feedback and support whenever required.

TABLE OF CONTENTS

STATEMENT OF ORIGINALITY..... 3

COPYRIGHT DECLARATION 4

ACKNOWLEDGEMENTS 5

ABSTRACT 11

ABBREVIATIONS 13

CHAPTER 1 – INTRODUCTION 16

CHAPTER 2 – SYSTEMATIC REVIEW AND META-ANALYSIS 31

CHAPTER 3 – THE USE OF AN AUTOMATED SARCOPENIA MEASUREMENT TOOL AS A PREDICTOR OF OUTCOME IN PATIENTS WITH ABDOMINAL AORTIC ANEURYSMS (SPAR) 72

CHAPTER 4 – USE OF QUADRICEPS MUSCLE AS A SURROGATE MEASURE OF MUSCLE AREA AND DENSITY IN PATIENTS UNDERGOING ABDOMINAL AORTIC ANEURYSM (AAA) REPAIR 125

CHAPTER 5 – CONCLUSIONS AND FUTURE WORK 152

REFERENCES 162

APPENDIX 1 - ETHICS APPROVAL 178

LIST OF FIGURES / TABLES

Figure 1. Two types of surgical intervention for infrarenal AAA repair – (a) Endovascular Aneurysm Repair (EVAR) (b) Open Surgical Repair (OSR) ……... 20

Figure 2. Table summarising the search terms used in Embase and Medline …… 47

Figure 3. Study flow diagram based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) ……..... 52

Figure 4. Summary table listing the 7 studies included in the systematic review and meta-analysis ……..... 53

Figure 5. Comparative data on baseline demographics and clinical characteristics …… 55

Figure 6. Risk of bias graph and summary …….. 57

Figure 7. Quality in Prognostic Studies (QUIPS) detailed checklist summary for individual studies risk of bias ……..... 58

Figure 8. Forest plot of 30-day or in-hospital mortality in patients with Normal vs Low SMI undergoing vascular surgery ……..... 59

Figure 9. Forest plot of post-operative morbidity in patients with Normal vs Low SMI undergoing vascular surgery ……..... 60

Figure 10. Forest plot of all-cause long-term mortality in patients with Normal vs Low SMI undergoing vascular surgery ……..... 60

Figure 11. Sub-group analysis of patients undergoing intervention for AAA vs PAD, comparing 30-day mortality in patients with Normal vs Low SMI …… 62

Figure 12. Sub-group analysis of patients undergoing intervention for AAA vs PAD, comparing 30-day morbidity in patients with Normal vs Low SMI …… 62

Figure 13. Sub-group analysis of patients undergoing intervention for AAA vs PAD, comparing all-cause long-term mortality in patients with Normal vs Low SMI …… 63

Figure 14. Sub-group analysis of Manual vs Semi-automated SMI calculation, comparing 30-day mortality in patients with Normal vs Low SMI undergoing vascular surgery …… 64

Figure 15. Sub-group analysis of Manual vs Semi-automated SMI calculation, comparing all-cause long-term mortality in patients with Normal vs Low SMI …… 64

Figure 16. Table demonstrating Accordion scoring system to grade the severity of post-operative complications of patients undergoing AAA repair …….. 90

Figure 17. Flow diagram demonstrating study patient cohort and reasons for inclusion / exclusion, and final sample size for testing the sarcopenia segmentation tool, and associated outcomes analysis ……..... 95

Figure 18. SPAR study demographics table …… 97

Figure 19. Study demographics table of patients undergoing manual segmentation of pseudonymised CT scans for 'ground truth' ……... 99

Figure 20. Sample automated segmentation output produced by the sarcopenia segmentation tool demonstrating patient and scan identifiers, along with automated muscle segmentation (in purple) and relevant outputs ……..... 101

Figure 21. Scatterplot demonstrating excellent IRR between manual segmentations undertaken by 2 independent clinical researchers for skeletal muscle area (SMA) ……. 104

Figure 22. Scatterplot demonstrating excellent IRR between manual segmentations undertaken by 2 independent clinical researchers for skeletal muscle attenuation (in Hounsfield units) ……..... 105

Figure 23. Scatterplots demonstrating excellent IRR between manual and automated segmentations for skeletal muscle area (SMA) ……. 106

Figure 24. Scatterplots demonstrating excellent IRR between manual and automated segmentations for skeletal muscle attenuation (in Hounsfield units) ……. 107

Figure 25. Histogram demonstrating SMI values of the entire SPAR patient cohort …….. 108

Figure 26. Histogram demonstrating SMD values (in Hounsfield units) of the entire SPAR patient cohort ……..... 109

Figure 27. Summary of outcomes analysis for men undergoing open AAA repair …… 112

Figure 28. Summary of outcomes analysis for women undergoing open AAA repair …….. 114

Figure 29. Summary of outcomes analysis for men undergoing EVAR …… 116

Figure 30. Summary of outcomes analysis for women undergoing EVAR …….. 117

Figure 31. Summary of outcomes analysis for patients undergoing TAAA repair …… 119

Figure 32. Examples of segmentation 'bleed out' by the sarcopenia segmentation on the tool - On the left, segmentation bleed out onto haematoma associated with a ruptured abdominal aortic aneurysm and on the right, segmentation bleed out onto the liver …… 120

Figure 33. Examples of L3 skeletal muscle and quadriceps muscle segmentation at anatomical levels defined by the study methodology ……... 132

Figure 34. Study flow diagram demonstrating sample size, and inclusion / exclusion criteria for overall and restricted analysis ……..... 134

Figure 35. Study demographics of the quadriceps segmentation study to assess the utility of quadriceps muscle as a surrogate muscle marker to skeletal muscle at the L3 level ……... 135

Figure 36. Intraclass correlation coefficient (ICC) to compare inter-rater reliability of manual muscle segmentations undertaken by the 2 clinical researchers, in both overall and restricted analyses ……..... 137

Figure 37. Bland-Altman plots demonstrating excellent IRR between manual segmentations undertaken by the 2 clinical researchers, in both overall and restricted analyses ……. 139

Figure 38. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle area against SMA at L3 for observer 1 ……..... 141

Figure 39. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle area against SMA at L3 for observer 2 ……..... 142

Figure 40. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle density against SMD at L3 for observer 1 ……... 144

Figure 41. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle density against SMD at L3 for observer 2 ……... 145

ABSTRACT

Introduction

Sarcopenia is associated with adverse outcomes in a variety of surgical specialties, including vascular surgery. Skeletal muscle index (SMI) and skeletal muscle density (SMD) have been shown to be a surrogate marker of sarcopenia and frailty. We explore the feasibility of using a fully automated deep-learning section-based muscle segmentation tool in calculating SMI and SMD, and predicting outcomes in patients undergoing abdominal aortic aneurysm (AAA) repair.

Methods

Pre-operative CT scans of patients undergoing AAA repair between 2016 and 2020 were run through and analysed by the automated sarcopenia segmentation tool, with 20% of scans undergoing manual segmentation to serve as 'ground truth'. The primary outcome measure was the correlation between SMI at L3 (level of the 3rd lumbar vertebra) and 30-day composite outcome, which was a combination of death and defined post-operative complications.

Results

280 CT scans underwent successful automated segmentation, with excellent IRR noted between the manual and automated segmentations with an ICC 0.96 (95% CI 0.95 – 0.97, $p < 0.05$) for skeletal muscle area (SMA) and ICC 0.98 (95% CI 0.98 – 0.99, $p < 0.05$) for skeletal muscle density (SMD).

On regression modelling of SMI and 30-day composite outcome, good negative correlation was noted amongst women undergoing open AAA repair ($r = -0.86$ / $R^2 = 0.75$, $p<0.05$).

Men undergoing open AAA repair in the lowest SMD quartile had poorer 30-day composite outcomes as compared to the highest quartile ($t(34) = 2.741$, $p = 0.01$).

Conclusion

This automated sarcopenia segmentation tool provides a rapid and accurate measure of SMI and SMD, and may help produce a sarcopenia score to guide clinicians on optimal treatment strategies and improve post-operative outcomes by identifying high-risk patients.

Future research should look to validate the findings of this study using prospective, multicentre trials.

ABBREVIATIONS

AAA	Abdominal Aortic Aneurysm
BEVAR	Branched Endovascular Aneurysm Repair
BIA	Bioelectrical Impedance Analysis
BMD	Bone Mineral Density
BMI	Body Mass Index
BQ	Bilateral Quadriceps
CAD	Coronary Artery Disease
CI	Confidence Interval
CKD	Chronic Kidney Disease
CLI	Critical Limb Ischaemia
CLTI	Chronic Limb Threatening Ischaemia
Cm	Centimetre
COPD	Chronic Obstructive Pulmonary Disease
CRP	C-Reactive Protein
CSMA	Cross Sectional Muscle Area
CT	Computed Tomography
CTD	Connective Tissue Disease
CVD	Chronic Obstructive Pulmonary Disease
DEXA	Dual X-ray Absorptiometry
DXA	Dual energy X-ray Absorptiometry
ECG	Electrocardiogram
ESC	European Society of Cardiology
ESVS	European Society for Vascular Surgery
EVAR	Endovascular Aneurysm Repair
EWGSOP	European Working Group on Sarcopenia in Older People
FCNN	Fully Convolutional Neural Networks
FEVAR	Fenestrated Endovascular Aneurysm Repair
HbA1c	Glycated Haemoglobin
HCRW	Health and Care Research Wales (HCRW)
HDL	High Density Lipoprotein

HDU	High Dependency Unit
HR	Hazard Ratio
HRA	Health Regulatory Authority
HU	Hounsfield Units
IC	Intermittent Claudication
ICC	Intraclass Correlation Coefficient
ICD	International Classification of Disease
ICHT	Imperial College Healthcare NHS Trust
ID	Identity Document
IEP	Image Exchange Portal
IFU	Instructions For Use
IRAS	Integrated Research Application System
IRR	Interrater Reliability
IT	Information Technology
ITU	Intensive Therapy Unit
IQR	Interquartile Range
JRCO	Joint Research Compliance Office
Lbs	Pounds
L3	3rd Lumbar Spine Vertebra
LOS	Length of Stay
LQ	Left Quadriceps
m/s	metres / second
MHU	Mean Hounsfield Units
mm	Millimetre
MRI	Magnetic Resonance Imaging
NHS	National Health Service
NVR	National Vascular Registry
OR	Odds Ratio
OSR	Open Surgical Repair
PACS	Picture Archiving and Communication System
PAD	Peripheral Arterial Disease
PMA	Psoas Muscle Area

PMD	Psoas Muscle Density
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International Prospective Register of Systematic Reviews
PVD	Peripheral Vascular Disease
QMA	Quadriceps Muscle Area
QMD	Quadriceps Muscle Density
QUIPS	Quality in Prognostic Studies
RCT	Randomised Controlled Trial
REDCap	Research Electronic Data Capture
ROC	Receiver Operating Characteristic
RQ	Right Quadriceps
SD	Standard Deviation
SMA	Skeletal Muscle Area
SMD	Standardised Mean Difference
SMI	Skeletal Muscle Index
SOP	Standard Operating Procedure
SPAR	The Use of An Automated Sarcopenia Measurement Tool as a Predictor of Outcome in Patients with Abdominal Aortic AneuRysms
SPPB	Short Physical Performance Battery
TAAA	Thoracoabdominal Aortic Aneurysm
TBAD	Type B Aortic Dissection
TEVAR	Thoracic Endovascular Aortic Repair
TUG	Timed-Up-and-Go Test
UK	United Kingdom
US	Ultrasound
USA	United States of America
VSGBI	Vascular Society of Great Britain & Ireland

1. INTRODUCTION

Introduction

The aorta is the main blood vessel originating from the heart, receiving the output of the left ventricle during cardiac systole and responsible for supplying oxygenated blood to the majority of the body through it's various branches. The abdominal aorta extends from the diaphragmatic crus down to the aortic bifurcation into the common iliac arteries, which usually occurs in front of the 4th lumbar vertebral body (L4). An abdominal aortic aneurysm (AAA) is defined as an abnormal, permanent dilatation of > 50% (as compared to it's 'normal diameter') or > 3 cm of any segment of the abdominal aorta. The pathophysiology of AAAs is complex and includes a combination of upregulation of proteolytic pathways, apoptosis, oxidative stress, inflammation, and loss of arterial wall matrix, with advanced atherosclerosis thought to be main underlying mechanism. However, whilst most AAAs are confined to the segment of aorta below the renal arteries (infrarenal aorta), aortic aneurysmal disease is considered a system wide disorder with generalised arteriopathy and ectasia commonly seen in the rest of the arterial tree (1).

Ultrasound screening studies have shown AAAs to be prevalent in about 4 – 8 % of men over the age of 65 years in 'Western' populations, though the overall incidence of AAAs is decreasing on a year-on-year basis, possibly due to reduction in smoking prevalence. The prevalence in women is lower at 1 – 2 % though outcomes are poorer in women as compared to men. The exact reasons for this remain unknown (1,2). The main risk factors for developing an abdominal aortic aneurysm include smoking, male gender, family history, age, hypertension, and low HDL cholesterol levels. Diabetes mellitus is thought to offer a 'protective' benefit in preventing the development of AAAs or slowing down the rate of expansion and disease progression, though the exact pathophysiology remains unclear (1–3).

Glycated haemoglobin (HbA1c) is used to monitor the level of blood sugar control and disease severity in diabetic patients, which reflects the average blood sugar levels over the preceding 3 months. It has been noted that patients with higher HbA1c levels tend to have lower rates of aneurysm growth (3).

The mainstay of AAA management has been surveillance and intervention beyond a certain threshold diameter to prevent rupture. This is because ruptured AAAs have a mortality rate approaching 80%, while the 30-day mortality rate of elective AAA repair is about 5% (1,3). However, most AAAs are asymptomatic at presentation, with initial symptoms only occurring at the time of rupture. This has led to the development of population wide screening programs to diagnose AAAs whilst still asymptomatic for men aged 65 and over, but has not been adopted in women yet as it was not found to be cost effective (1). The increasing 'routine' use of imaging modalities such as ultrasound (US), CT and MRI in clinical practice for a variety of clinical presentations and disease management has also led to the incidental diagnosis of asymptomatic aortic aneurysmal disease. When symptomatic before rupture, the usual presentations tend to be that of abdominal, flank, loin or back pain usually due to local compressive effects or due to lower limb or pelvic / buttock ischaemia due to distal embolic 'trashing' of thrombus from within the aneurysm sac.

Cardiovascular disease is the leading cause of death in England and Wales, and peripheral vascular disease (PVD) is a common occurrence in this patient cohort and in our ageing population. The incidence of PVD varies from 3 – 10 % in people aged less than 70 years, to 15 – 20% in people aged over 70 years. PVD can be caused by luminal stenosis or occlusion of any artery supplying the limb, and may be symptomatic or asymptomatic. PVD is most commonly caused by atherosclerosis, the risk of which increases with age (amongst other risk

factors such as smoking, sex, ethnicity and diet), but may have alternative causes such as inflammatory arteriopathies (4). The range of symptoms associated with PVD is variable, ranging from asymptomatic to intermittent claudication (IC) to disabling rest pain and tissue loss / ulceration. Classification systems like the Rutherford and Fontaine classifications have been used to describe the severity of symptoms. Patients with ischaemic rest pain, with or without tissue loss are classified as having critical limb ischaemia (CLI) or chronic limb threatening ischaemia (CLTI) and are at significantly increased risk of major limb amputation, as well as at least a 3-fold increased risk of cardiac or cerebrovascular events and death, as compared to patients with IC (5).

Patients with AAAs have a poorer prognosis as compared to their age and sex matched 'healthy' controls, likely due to their increased cardiovascular risk. To date, no pharmacological therapy exists that limits or reverses aneurysm growth, and surgical intervention remains the treatment of choice once the threshold diameter for intervention has been reached (1,6). There are 2 primary types of surgical intervention (see Figure 1) available for infrarenal AAA repair – Open Surgical Repair (OSR) or Endovascular Aneurysm Repair (EVAR). OSR involves clamping a 'healthy' segment of aorta proximally and iliac arteries distally, opening the aneurysm sac thereafter and sewing a synthetic 'tube' or 'bifurcated' / 'trouser' graft (usually made of Dacron) to healthy inflow and outflow segments of artery to restore flow through this segment. EVAR involves the endoluminal placement of a stent graft within the aneurysm sac, to divert flow through this stent graft effectively excluding flow through the aneurysm sac and is usually placed in the patient through endovascular access of the common femoral arteries in the groin, under fluoroscopy (x-ray) guidance. These stent grafts usually have a nitinol 'skeleton' with fabric attached to associated struts, and have a set of anatomical considerations to be met to ensure effective device placement in the patient. Each

individual device manufacturer provide their own device specific instructions for use (IFU), the adherence to which improves technical success and overall patient outcomes (1,6). EVAR is generally considered to be the 'less' invasive or 'minimally' invasive approach to AAA repair, but requires lifelong surveillance due to the risk of development of 'endoleaks' where blood flow persists within the aneurysm sac resulting in a risk of aneurysm rupture and associated complications, including death.

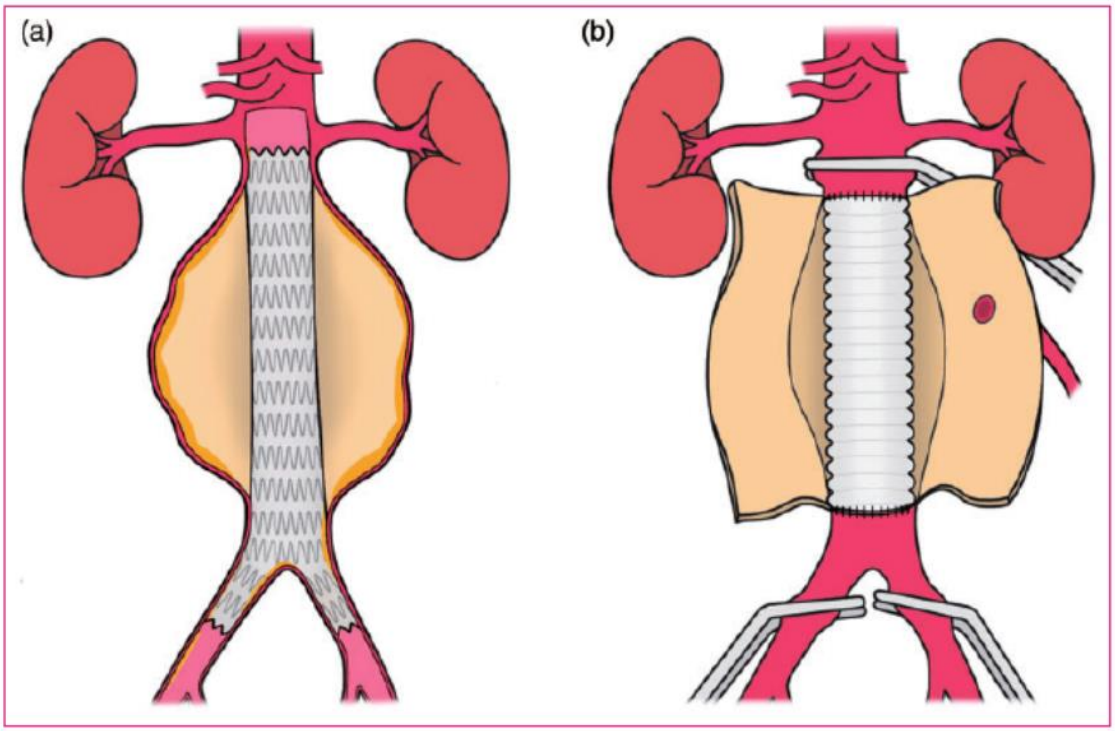


Figure 1. Two types of surgical intervention for infrarenal AAA repair – (a) Endovascular Aneurysm Repair (EVAR) (b) Open Surgical Repair (OSR) (7)

Multiple previous randomised control trials (RCT) have been undertaken to identify the optimal threshold diameter for intervention on abdominal aortic aneurysms. The UK Small Aneurysm Trial (6) and a similar RCT in the USA (8) did not identify any survival benefit in treating patients with AAAs less than 5.5 cm in diameter with OSR. No significant cost benefits

were noted either. Whilst a significant proportion of patients with small aneurysms will require surgical intervention in the future with progressive aneurysm size increase, no effects on peri-operative mortality or long-term survival were noted with later intervention. Whilst endovascular aneurysm repair (EVAR) has improved 30-day post-operative mortality as compared to open surgical repair (OSR), EVAR in patients with AAAs less than 5.5 cm did not demonstrate any survival benefit either (9). This had led to 5.5 cm being globally accepted as the diameter threshold for intervention on AAAs, though these studies are now at least 20 years old and significant improvements have been made in peri-operative care and device technology since.

The optimal choice of intervention between OSR and EVAR also remains unclear. Whilst minimally invasive or 'keyhole' surgery has now become the 'gold standard' for common surgical procedures such as cholecystectomy, appendicectomy and prostatectomy, the same cannot be said for EVAR. This is largely due to the long-term results of EVAR and increased need for future re-intervention, which has questioned it's durability and cost effectiveness compared to OSR. The 15 year follow up results from the EVAR-1 trial (10) which compared patients undergoing OSR and EVAR demonstrated that the early survival benefit from EVAR was lost within 2 years, and beyond 8 years OSR had a lower aneurysm-related and all-cause mortality, due to the significant number of late aneurysm sac ruptures in the EVAR group (due to endoleaks), along with an increased cancer risk. It has been hypothesised that the increased exposure to ionising radiation during the EVAR and subsequent surveillance from CT scans may confer this increased cancer risk. On the contrary, OSR does not involve any radiation exposure during the procedure and long-term CT surveillance is usually not undertaken, unless clinical concerns, as endoleaks are not seen with OSR.

Whilst the debate of OSR versus EVAR as the optimal treatment of choice in the management of aortic aneurysmal disease is likely to continue for the foreseeable future, patient selection and optimisation may play a more important role in improving overall outcomes. The European Society of Cardiology (ESC) guidelines consider open AAA repair to be a 'high' risk surgical intervention with a 5% risk of cardiovascular morbidity or mortality in the 30-day post-operative period, whereas EVAR is considered 'intermediate' risk with a 1 – 5 % risk of 30-day cardiovascular morbidity or mortality (5).

The European Society for Vascular Surgery (ESVS) guidelines recommend that all patients undergoing AAA repair should be subject to a detailed clinical history and physical examination, specified blood tests, an electrocardiogram (ECG) and a 'functional assessment' (11). However, the definition of this functional assessment and what it entails remains unclear in these guidelines.

Furthermore, specific recommendations for cardiac, respiratory, renal and nutritional work-up and referral are provided, with specific objective tests for each. For example, pre-operative cardiac investigations may be undertaken using standard or stress echocardiography, and respiratory assessments may be undertaken using pulmonary function tests with spirometry. The role of pre-operative optimisation in these domains is also clearer, for instance, with pre-operative chest physiotherapy or supervised exercise therapy (12).

However, no such diagnostic tests exist for assessing a patient's functional status, and this may therefore need to be measured indirectly. Furthermore, surgical risk prediction is central to both the decision making and patient counselling on the ideal approach to AAA repair. However, current surgical risk prediction models under-perform in the elderly as they only capture a snapshot of the patient's condition and do not reflect the patient's physiologic resilience to the complications of surgery. The inclusion of a frailty assessment has therefore been advocated to improve the prediction of surgical risk. Risk stratification and frailty assessment may help

in predicting which patients would benefit from surgery and also the choice of surgical intervention (conservative vs endovascular vs open surgery) (13).

Is measurement of frailty the answer?

The degree of frailty may be a useful measure of a patient's functional status, and this may therefore be used as a surrogate marker of function in a patient's pre-operative assessment, with a view to providing risk stratification and help identify those patients in need of pre-operative functional optimisation. However, the complexities of frailty definitions and assessments have made this a challenging process.

Some definitions of frailty have considered it to be synonymous with disability, co-morbidity or advanced aged and have led to the use of these terms interchangeably. This may often be incorrect as not all 'old' patients can be considered frail, and not all 'young' patients can be considered non-frail. Moreover, disability may be unrelated to frailty and may be secondary to trauma or congenital abnormalities (14).

More medical definitions of frailty have considered it as a syndrome of gradual loss of inherent physical homeostasis, reducing the body's ability to cope with a variety of physiological stressors due to decreased reserve, and generalised decline across organ systems, leading to adverse outcomes and poor physical health. Several surrogate markers of frailty have been identified that include, but are not limited to, lean body mass, physical strength, physical endurance, balance, gait, walking performance and activity levels (14).

A cross sectional, population-based study of over 5000 patients in the USA concluded that 'frailty' may be defined as a physical state of high vulnerability for adverse health outcomes, such as falls, hospitalisation, physical disability and mortality (14,15). It was identified as a

distinct clinical syndrome, and may be diagnosed by the presence of 3 or more of the following criteria:

1) Unintentional weight loss (10 lbs in the preceding 1 year)

2) Self-reported exhaustion

3) Weakness (measured using grip strength)

4) Slow walking speed

5) Low physical activity

Considering frailty's definition as a distinct clinical syndrome, certain risk factors for frailty were also identified such as ethnicity (increased risk in African Americans), lower socio-economic status and co-morbidities (14). Furthermore, it was established that in patients undergoing surgery, frailty confers poorer outcomes, both in terms of morbidity and mortality, as these patients have an increased risk of post-operative complications, longer hospital length of stay (LOS) and slower return to baseline functional status following an intervention. Patients on the spectrum of frailty, who fulfil only 1 or 2 of the criteria listed above have an intermediate risk as compared to non-frail patients and are more likely to develop 'complete' frailty over time, thereby highlighting the importance of early recognition of these patients, and the importance of attempting to delay the onset of frailty and need for optimisation in this patient cohort. The transformation of a 'non-frail' to a 'frail' patient is a linear process and identification of patients on the path to frailty is key to improving overall outcomes and quality of life (14).

But therein lies the problem. Whilst we may have a definition of frailty and understand it's detrimental effects, the identification of frail patients is not as straightforward. Frailty, which is typically measured by combining information from the patient's medical history and

assessments of physical and functional status, is reliant on multiple, time consuming subjective assessments, which may be 'operator-dependent' (16). For example, the Frailty Index (FI) or Clinical Frailty Scale (CFS) developed by the Canadian Study of Health and Aging (CSHA) includes 70 items derived from a patient's history and physical examination to reach a score (17). Whilst this is a well validated tool to measure frailty, the time consuming nature of this assessment has meant that we are generally unable to utilize it for every patient that we see in the out-patient or in-patient setting due to our already stretched services and limited resources. This has, therefore, limited the routine utility of frailty assessments in day-to-day clinical workflow. As a result, the true degree of frailty may not be clear to clinicians in the consulting room or in-patient hospital admission bed prior to embarking on major surgical intervention.

How can we measure frailty more easily?

An objective or quantitative measure of frailty, which can be obtained rapidly, would allow better appreciation of appropriateness of intervention and associated risks, especially in a disease process like aortic aneurysmal disease that is largely asymptomatic until the terminal event (18). This would also provide clinicians with an easier and more standardised outcome measure to help guide treatment decisions thereafter.

Muscle bulk has long been felt to correlate well with frailty and patient physical capacity and lean body mass is considered to be one of the surrogate markers of frailty. 'Sarcopenia' has been defined as a progressive, generalised and / or accelerated loss of muscle mass and function that is associated with increased adverse outcomes including falls, functional decline, frailty, other morbidity, and mortality. It may be an age-related process but may also be influenced by genetic and lifestyle factors. Whilst sarcopenia is more commonly seen in the elderly, it is by

no means seen only in this cohort, and can also be seen in younger patients, particularly associated with certain medical conditions. There has been an increasing attempt over the last few years to discover biomarkers to diagnose early sarcopenia, and also develop therapeutic strategies aimed at reversing or slowing down the advent of this process, due to the recognition of it's long-term adverse outcomes (19).

The psoas muscle is one such muscle group that has been studied extensively as a marker of sarcopenia to measure and compare patient outcomes. The frequency of imaging of the psoas muscle, particularly in cross sectional imaging like CT or MRI and it's ease of identification and assessment on these imaging modalities has led to many researchers analysing it's effects on outcomes, and whether it could be used to diagnose sarcopenia.

Psoas muscle area (PMA) is a simple measure of cross-sectional muscle area measured at a pre-determined axial level (e.g. middle of 3rd lumbar spine – L3) on cross sectional imaging, which can be readily obtained from CT or MRI scans. CT imaging forms a standard part of the pre-operative evaluation of patients undergoing AAA repair, and is also an integral part of post-intervention surveillance, particularly following endovascular intervention. Low PMA has been associated with post-operative morbidity and mortality in a variety of surgical specialities, including general and orthopaedic surgery (20). However, previous work on the effect of PMA in vascular surgery has shown differing results with some showing no correlation (21), whilst others have shown that sarcopenia may identify patients who will not benefit from intervention (13,22–24).

One such study (13) identified a 14% reduction in all-cause mortality for every 1 cm^2 decrease in psoas muscle area. After adjusting for various patient specific variables such as age, sex and

type of surgery, a significant association was noted between PMA and all-cause mortality with a hazard ratio (HR) of 0.86 per cm^2 (95% CI 0.79 – 0.93).

Another cohort study in vascular surgery (23) retrospectively looked at 135 patients who had undergone EVAR between 2010 and 2016 and measured total psoas muscle area on an axial CT slice just inferior to the 4th lumbar (L4) superior end plate. Whilst there were no effects on hospital length of stay, a significant difference in long term mortality was noted. Patients in the lowest tertile had an increased risk of 5-year mortality (odds ratio – OR – 3.9, 95% CI 1.2 – 12.9), with only 42% survival noted in that group of patients at 5 years, as compared to 93% survival in the remaining cohort ($2/3^{rd}$ of patients). Whilst this may help identify a group of patients that may benefit from no surgical intervention, it may also help to increase intervention thresholds for the higher risk group, so as to improve their overall risk / benefit profile, thus warranting intervention at a higher threshold diameter when rupture risks also increase.

Another similar study in vascular surgery (24) looked at 191 patients who had undergone EVAR and measured total psoas muscle area at the most caudal axial slice of the L3 vertebra, and normalised this for height, with sarcopenia being defined as a total psoas area of < 500 mm^2/m^2. Sarcopenic patients were noted to have poorer overall survival (HR 2.37, p = 0.011) and longer hospital length of stay compared to non-sarcopenic patients (4.0 days vs 3.0 days, p = 0.008), but no significant differences were noted in terms of post-operative complications.

Another retrospective single-centre study from the USA (22) looked at 272 patients who underwent EVAR between 2002 and 2014 and measured total psoas muscle area at the L3 vertebral level, and then normalised this for height. Sarcopenia thresholds were set using a receiver operating characteristic (ROC) curve, and a threshold of 500 mm^2/m^2 identified. It was

noted that patients in the lowest quartile of total psoas muscle area had a significantly higher overall mortality as compared to patients in the highest quartile (23.53% vs 5.88%, p = 0.01).

However, not all published studies have demonstrated similar findings when investigating the role of psoas muscle area as a prognostic marker. 382 patients who underwent EVAR or OSR of AAAs at the Leeds Vascular Institute between 2008 and 2016 (21) demonstrated no association between total psoas muscle area and mortality. However, total psoas muscle area was considered and analysed as a continuous variable in this cohort, rather than distributing patients into 'sarcopenic' and 'non-sarcopenic' groups or determining cut-off values / thresholds for sarcopenia.

As previously alluded to, aortic aneurysmal disease is a system wide disorder with generalized arteriopathy seen in the rest of the arterial tree as well. Peripheral vascular disease, specifically lower limb occlusive disease may manifest as critical limb threatening ischaemia (CLTI) in advanced cases, and similar findings have also been noted in this group of patients. CLTI already carries a very poor prognosis with at least 1 in 4 patients (25%) dead within one year of diagnosis (4). Appropriate patient selection to identify those patients who are likely to benefit the most from revascularization strategies is, therefore, even more important.

A retrospective analysis of 64 patients with CLTI (25) measured the cross-sectional skeletal muscle area at the L3 vertebral level and utilised cut-off values of < 114 cm^2 for men and < 89.8 cm^2 for women. The 5-year survival rate was significantly lower for sarcopenic patients, as compared to non-sarcopenic patients (23.5% vs 77.5%, p = 0.001), and suggested that pre-operative optimization via nutritional optimisation and exercise therapy may help improve overall outcomes and survival.

A similar retrospective analysis on 117 patients undergoing infra-inguinal revascularization between 2006 and 2015 (26) measured thigh muscle area and bone area on an axial slice which was at the mid-point of a line drawn from the greater trochanter of the femur to the superior border of the patella. Thigh sarcopenia was defined as a ratio of thigh muscle area to bone area of < 9 (thigh muscle area / bone area < 9). Sarcopenic patients had a significantly worse survival at 2 years as compared to non-sarcopenic patients (HR 2.64, 95% CI 1.11 – 6.70, p = 0.03), with overall survival rates of 55.4% in sarcopenic patients and 86.5% in non-sarcopenic patients at 2 years.

This initial work in patients with critical limb ischaemia has suggested that sarcopenia has a reasonable prognostic prediction ability in this cohort of patients as well, and has opened the door to future research in this group of patients (25,26). However, one must remember that psoas muscle area may not be available in all patients with lower limb vascular pathology, as they may only have duplex ultrasound imaging or contrast / subtraction angiography as their pre-operative investigations or they may have cross sectional imaging (CT / MRA) that does not extend to the level of the psoas muscle. Furthermore, lower limb peripheral vascular disease may lead to unilateral or bilateral lower limb muscle atrophy due to ischaemia, and this needs consideration when determining threshold values for sarcopenia, particularly when measuring muscle area in the lower limb. Reduced mobility due to claudication type symptoms further exacerbates this issue by causing disuse atrophy.

Other cohorts of vascular surgery patients such as those with carotid or cerebrovascular disease, upper limb vascular disease or thoracic aortic disease may encounter similar problems with the lack of availability of psoas muscle or abdominal skeletal muscle in pre-operative imaging, and alternate methods of diagnosing and assessing sarcopenia may be required.

Furthermore, global improvements in healthcare have led to increased life expectancy resulting in more and more elderly patients potentially requiring intervention. Improvements in medical management with the advent of best medical therapy (combination of an antiplatelet and a high dose statin) and an increasing uptake of smoking cessation globally has also resulted in vascular disease presenting at a later stage of life due to a combination of risk factor optimisation and improved life expectancy. However, overall healthcare resources remain limited and stretched to cater for an ever expanding population with complex healthcare needs. These factors further highlight the importance of careful patient selection to identify those that will benefit the most from intervention, and also being able to risk stratify patients to identify the optimal treatment strategy for them.

Diagnosing sarcopenia, and therefore frailty, is certainly not going to be the definitive answer to this complex question, but it can definitely serve as an important tool in the complex armamentarium of surgical decision making.

2. SYSTEMATIC REVIEW & META-ANALYSIS

Skeletal Muscle Index (SMI) as a Predictor of Outcome in Vascular Surgery: A Systematic Review and Meta-Analysis

Abstract

Introduction:

Sarcopenia as a marker of frailty, has been shown to be associated with a higher risk of mortality and morbidity after surgery. Various methods to measure sarcopenia have been explored and some studies have shown that skeletal muscle index (SMI) is an effective objective measure of sarcopenia.

Methods:

A systematic review of the literature was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and registered with the International Prospective Register of Systematic Reviews (PROSPERO registration number: CRD42022314191). Meta-analysis for all-cause mortality using time-to-event data was reported as summary hazard ratio (HR), and 30-day mortality and morbidity were reported using odds ratio (OR).

Results:

Seven (7) prognostic studies were included in the systematic review and meta-analysis, which included a total of 1991 patients. Patients with low SMI had a significantly higher risk of mortality in the long term compared to patients with normal SMI (HR 2.80, 95% CI 2.25 – 3.49; $p<0.001$), as well as a significantly higher risk of 30-day mortality (OR 12.74, 95% CI 3.29 – 49.39; $p<0.001$) and morbidity (OR 6.21, 95% CI 1.31 – 29.94; $p<0.001$). Sub-group

analysis showed similar findings between different types of vascular pathology and muscle segmentation methods (manual vs semi-automated muscle segmentation).

Conclusion:

Low skeletal muscle index (SMI) is associated with an increased risk of mortality and morbidity after vascular surgery, and may serve as a useful risk stratification tool to identify those patients at greater risk. This may also guide the clinician on the optimal treatment strategy (open vs endovascular vs conservative) for each individual patient. Standardised definitions of SMI cut-off values and tools for rapid, automated assessment of muscle mass will further improve it's utility and usability in routine clinical practice.

As discussed in Chapter 1 (Introduction), current surgical risk prediction models are mostly based on individual, subjective clinician assessments of patients' co-morbidities, physical status, and use of incompletely validated risk stratification tools (27). These potentially subjective assessments often do not reflect the true physiological reserve of the patient's body, and therefore may undercompensate or overcompensate for patient frailty. Furthermore, while frailty assessments and scores may be key to the prediction of surgical peri-operative risk and have some utility in the decision-making process, they remain complex and time consuming, and cannot be routinely integrated into day-to-day clinical workflow or electronic records as often as required (28). Hence, we, as clinicians may not be able to effectively predict post-operative outcomes or provide adequate pre-operative risk stratification to keep our patients well informed and aid this sometimes complex decision making process.

This has led to the need for measuring frailty and risk stratifying patients in a more objective / quantitative and less time-consuming manner. Reviews from various fields, such as abdominal surgery (29), adult spinal surgery (30), and oncological surgery (31), have shown that sarcopenia, which is defined as a generalised loss of skeletal muscle and function, may have an association with mortality and major complications after surgery.

To recap, sarcopenia is the progressive loss of skeletal muscle mass and function, which has demonstrable links to mortality, quality of life, post-operative complications and hospital length of stay. Early recognition of sarcopenia may help in optimising or reversing this process through exercise therapy and / or nutritional optimisation. Whilst I have briefly touched on the use of cross-sectional psoas muscle area as a surrogate marker of sarcopenia, there exist various ways of measuring sarcopenia. Other methods include the measurement of total body skeletal muscle mass, appendicular skeletal muscle mass (muscle mass associated with the pelvic or pectoral girdle, or the upper or lower limbs) or cross-sectional muscle area (CSMA) of certain muscle groups (32).

However, despite all this, there seems to be a lack of consensus about sarcopenia guidelines, diagnosis, optimal investigations, 'cut-off values' for diagnosis and overall management of sarcopenia, which has also resulted in it's limited uptake in community practice. The European Working Group on Sarcopenia in Older People (EWGSOP) published an initial document in 2010 to try and answer some of these questions and lay down some initial guidelines. This was followed up in 2018 where this working group (EWGSOP2) met again to further refine these guidelines and provide an update based on the extensive research that had taken place in this relatively new field of clinical interest over the preceding decade (33).

Sarcopenia is now officially recognised as a muscle disease with an ICD-10 diagnosis code. This has also meant that muscle strength is now seen as the principal determinant of sarcopenia, and not muscle quantity or quality, as these parameters are harder to measure when attempting to diagnose a patient with sarcopenia in the out-patient setting. Muscle 'quantity' is a measure of how much muscle an individual has, and this can be a gross, total value in that individual, or can be normalised for height to provide a more standardized, comparative measure. Muscle 'quality' is determined by the microscopic and macroscopic structure of muscle tissue, and it's associated architecture and composition, and can be measured using imaging or muscle tissue analysis (more invasive). Whilst muscle strength, quality or quantity may be used in the diagnosis of sarcopenia, physical performance is generally used to categorise the severity of sarcopenia (33).

Categories of sarcopenia

Sarcopenia may be 'primary' or 'secondary'. Primary sarcopenia is an age-related process where there is gradual loss of total muscle mass with increasing age. Maximal levels of muscle mass are usually achieved in young adulthood, with gradual loss of muscle mass (1-2% per

year) and strength (1.5-5% per year) beyond the age of 40. Peak muscle mass is higher for males, as compared to females, which also necessitates the need for sex-based cut-off levels in sarcopenia.

Secondary sarcopenia is an accelerated loss of muscle mass (not proportional to age related loss alone) due to an underlying systemic disease (such as malignancy), inadequate nutritional availability (either due to inadequate intake or malabsorption) or due to lifestyle modifications resulting in a more sedentary lifestyle with physical inactivity (33).

Sarcopenia may also be 'acute' or 'chronic' with acute sarcopenia lasting less than 6 months, and usually due to an underlying 'secondary' cause, with reversibility noted on treating the underlying cause (33).

What clinical presentations are associated with sarcopenia?

A patient may report one of several symptoms to suggest sarcopenia. This may include, but is not limited to, recurrent falls, weakness, slow walking speed, difficulty or slowness when attempting to stand up from a sitting position, weight loss or muscle wasting.

EWGSOP2 recommends the use of the SARC-F questionnaire to further assess these patient reported symptoms. This is a 5-item questionnaire that assesses a patient's perception of individual strength, walking ability, ability to rise from a chair, ability to climb stairs and susceptibility to falls. The SARC-F has a low-to-moderate sensitivity, but a very high specificity in predicting low muscle strength (33). Another alternative is the Ishii screening test that calculates the probability of sarcopenia based on 3 variables (age, grip strength and calf circumference) to provide an overall score (34).

This then enables the clinician to assess the 3 key domains of sarcopenia diagnosis mentioned previously:

1) Muscle strength
2) Muscle quality and / or quantity
3) Physical performance

Broadly speaking, low muscle strength is suggestive of sarcopenia. Low muscle quality and / or quantity confirms the diagnosis of sarcopenia. The combination of muscle strength, quality and quantity, along with individual physical performance delineates the severity of underlying sarcopenia.

How should we diagnose sarcopenia?

Diagnosing sarcopenia may be a complex process. Whilst I have described the use of cross-sectional psoas and / or skeletal muscle area previously, there exist multiple other diagnostic tools. These can be further simplified based on the 3 key domains of sarcopenia diagnosis:

1) Diagnosing muscle strength – Measuring grip strength or performing the chair stand test (or chair rise test)
2) Measuring muscle quality or quantity – Dual energy x-ray absorptiometry (DXA), bioelectrical impedance analysis (BIA), CT or MRI imaging
3) Measuring physical performance – Gait speed, short physical performance battery (SPPB) test, timed-up-and-go test (TUG), 400-metre walk, or long-distance corridor walk test

Muscle strength as a diagnostic tool

Grip strength can be measured using a calibrated handheld dynamometer (e.g. Jamar dynamometer) and can be used as a surrogate marker of muscle strength across the body. Grip strength has been shown to be an independent predictor of outcomes such as hospital length of stay, functional limitations, quality of life and mortality.

Lower limb muscle strength can be measured using the chair stand test (or chair rise test) which measures the time taken for an individual to rise 5 times from a seated position to a standing position without using their upper limbs for support. A variation of this is the timed chair stand test where the same action is performed, but the total number of completed movements (sit to stand and back) in 30 seconds measured (33).

Muscle quantity as a diagnostic tool

Common methods of measuring muscle quantity include the use of CT or MRI to measure cross sectional muscle area, particularly in the hospital setting or for research purposes. In particular, skeletal muscle area at the L3 or L4 cross sectional level or psoas muscle area at these levels have been commonly used and studied to measure sarcopenia. The skeletal muscle area (SMA) is then normally adjusted for height or BMI to provide a more standardised measure that can then be used for comparison amongst different populations. However, despite multiple studies looking at these measurements, standardised or internationally recognised cut-off points and threshold values for ‘sarcopenic’ and ‘non-sarcopenic’ patients still do not exist.

Dual energy X-ray Absorptiometry (DXA) is a more widely available diagnostic tool that is commonly used in clinical practice to assess bone mineral density (BMD) in the lumbar spine, hip and / or forearm to diagnose osteoporosis. It can also be used to measure body composition to calculate lean mass and fat mass in the whole body, in addition to BMD. These lean mass measurements can be used to diagnose sarcopenia, and fat mass / distribution be utilised to produce an individual patient risk profile. Excessive amounts of visceral adipose tissue as compared to subcutaneous tissue increases the risk of mortality usually through increased cardiovascular risk, in addition to conferring worse lipid, glycaemic, haematocrit and renal blood profiles.

Another method of measuring muscle quantity is the use of bioelectrical impedance analysis (BIA). This measures the body's electrical conductivity to estimate the total muscle quantity / mass, but can therefore be influenced by the hydration status of the patient.

Calf circumference has previously been used as a surrogate for muscle quantity, but has been found to no longer be as accurate as the above modalities(33).

Physical performance as a diagnostic tool

Some of the tests that can be used to measure physical performance as a diagnostic tool include the Short Physical Performance Battery (SPPB), the Timed-Up and Go (TUG) test and measurement of gait speed. These tests not only measure muscle performance, but also serve as a marker of a patient's neurological function and balance.

Gait speed is usually measured over a distance of 4 metres, with a cut-off value of ≤ 0.8 m/s used as an indicator of severe sarcopenia by EWGSOP2 guidelines (33).

The SPPB tests a combination of gait speed, balance and performance of the chair stand test (measurement of lower limb muscle strength). It has a maximum score of 12, with a score of ≤ 8 suggestive of sarcopenia.

The TUG test measures the ability of an individual to rise from a chair, walk a distance of 3 metres, turn back and return to sit down in the chair.

The 400-metre walking test assesses the endurance of a patient by asking them to walk 20 laps of 20 metres as quickly as possible, while allowing only 2 rest stops. As a result, this test is time consuming, requires a corridor or walking space at least 20 metres long and may be hard to incorporate into routine clinical practice, particularly in the NHS (33).

Novel techniques of diagnosing sarcopenia

Considering sarcopenia is associated with poor outcomes, and has some reversibility to allow individual patient optimisation, there are ongoing efforts to identify new tests or chemical biomarkers to diagnose sarcopenia more easily, and assess response to risk factor modification, optimisation and intervention. However, any new test or screening tool needs to be cost effective, easily available and acceptable to the patient, in addition to being reproducible in a variety of settings. It should also, ideally, have internationally standardised and recognised cut-off values.

CT or MR imaging of the mid-thigh is one such diagnostic tool that has demonstrated good predictability as a surrogate marker of total skeletal muscle mass, and may have a greater correlation to a patient's total skeletal muscle mass as compared to cross-sectional muscle area (CSMA) at the L3 or L4 level. Additionally, there is also some evidence that it is more sensitive to change as compared to truncal skeletal muscle (33). This may be explained by poor mobility

or a less active lifestyle preferentially resulting in muscle atrophy of the lower limbs, as compared to skeletal muscle in the trunk or upper limb, due to disuse or reduced use.

Another novel area of research is the role of muscle 'quality' rather than 'quantity' in sarcopenia. This can usually be determined from cross sectional imaging like CT or MRI, which characterises muscle density or attenuation (measured in Hounsfield units). Factors like excessive fatty infiltration of the muscle results in poor 'quality' muscle which may exacerbate underlying frailty and impact an individual's functional status. Muscle quality may also refer to the ratio of muscle strength to appendicular skeletal muscle mass or volume (33).

In addition to cross-sectional imaging, other imaging modalities like ultrasound may also be used to assess sarcopenia, and have been shown to be useful in assessing quadriceps femoris muscle area and thickness. It may also be used to assess muscle quality, by delineating underlying muscle echogenicity. Fatty infiltration (myosteatosis) resulting in poor quality muscle can be seen to be hyper-echogenic on ultrasound imaging (33). Ultrasound, therefore, has the benefits of assessing muscle quality and quantity, without the underlying radiation risks of CT or costs associated with MRI. However, it's use is operator dependent, and may be more time consuming in day-to-day clinical practice where resources and scanning slots remain limited. Appropriate training of the ultrasound operator is also of paramount importance.

Moving away from imaging as a diagnostic tool for sarcopenia, several attempts have been made to identify a biomarker to diagnose sarcopenia. Because of the complexities surrounding sarcopenia as a diagnosis, it is unlikely that a single biomarker will provide a rapid and precise diagnostic tool, but may require a battery of tests to give an overall diagnostic score.

Creatine is a metabolite produced by the liver and kidney, and may also be ingested by the consumption of meat or certain other food products / supplements. It is partially taken up by muscle cells and converted to phosphocreatine, but excess creatine in the bloodstream is

converted into creatinine and excreted by the kidneys. Creatinine excretion rate is another potential marker of overall muscle mass, but is yet to be fully validated (33).

Suggested EWGSOP algorithm for screening and diagnosing sarcopenia (33)

The EWGSOP2 guidelines suggest the use of a pathway titled F-A-C-S to screen and diagnose sarcopenia. The F-A-C-S pathway stands for the following:

1) Find-cases – Use of the SARC-F questionnaire and / or clinical suspicion to identify potential at-risk individuals
2) Assess – Use of a sarcopenia diagnostic test that measures muscle strength e.g. grip strength or chair stand test
3) Confirm – Confirm sarcopenia by use of a diagnostic test that measures skeletal muscle mass / area / volume such as DXA, BIA, CT or MRI
4) Severity – Assess the severity of underlying sarcopenia by use of a diagnostic test that measures physical performance e.g. gait speed, SPPB or TUG test.

What is the role of skeletal muscle index (SMI) in sarcopenia?

Muscle mass correlates with body size (33). This means a 'larger' individual will have more skeletal muscle mass or volume, as compared to a 'smaller' individual. If threshold or cut-off values for sarcopenia were selected based on skeletal muscle mass / volume alone, then a 'larger' sarcopenic individual may be considered 'non-sarcopenic' due to their total muscle

mass / volume and a ‘smaller’ non-sarcopenic individual may be considered ‘sarcopenic’ for the opposite reasons.

The same principle can be applied to body weight. Absolute body weight alone does not determine whether an individual is underweight, ‘healthy’ or overweight. Absolute body weight normalised for height is used to produce a measure, commonly known as body mass index (BMI), which is utilised to classify individuals on the basis of weight.

BMI = Weight (in kgs) / Height^2 (in metres)

The same principles can be used in diagnosing sarcopenia, where muscle mass can be adjusted for body size using height, weight or body mass index (33,35).

This has given rise to the term called ‘skeletal muscle index (SMI)’ where skeletal muscle mass / volume is normalised for height. Considering axial slices from cross-sectional imaging such as CT or MRI are used to measure cross-sectional skeletal muscle area (CSMA) at a certain pre-specified level, SMI has also been extended to incorporate skeletal muscle area (SMA) normalised for height.

Skeletal Muscle Index (SMI) = Skeletal Muscle Area (SMA – in cm^2) / Height^2 (in metres)

Skeletal muscle index (SMI) is now considered as a suitable surrogate marker for sarcopenia assessment by several expert groups, and is extensively used in sarcopenia research, particularly in oncological studies (35).

The role of SMI in Vascular Surgery

Oncology patients have regular CT scans following diagnosis to monitor disease progression, treatment response and long-term surveillance once curative-intent treatment is completed. These CT scans have been used in several studies to diagnose sarcopenia and monitor it's associated detrimental effects.

Patients undergoing abdominal aortic aneurysm repair (AAA) in the elective or emergency setting will almost certainly have a pre-operative CT scan to understand the morphology of the aneurysm, and certain key anatomical features such as diameter and location. Pre-operative CT scanning is also essential to understand what endovascular options are available (standard infrarenal EVAR vs more complex endovascular solutions), and if endovascular aneurysm repair is chosen over open surgical repair (OSR), the pre-operative CT scan is utilised for choosing the appropriate endovascular device (based on it's instructions for use – IFU) and associated adjuncts (in addition to bail-out strategies).

Considering this pre-operative CT scan includes the abdomen / pelvis, it can similarly be utilised to calculate the cross-sectional skeletal muscle area (CSMA) at the L3 or L4 level to diagnose sarcopenia and study related outcomes.

As previously explained, this CSMA could then be normalised for height to produce the skeletal muscle index (SMI) to produce a more standardised measure of sarcopenia.

CT imaging (particularly CT angiography) is commonly used in vascular surgery for a variety of other pathologies too such as aorto-iliac occlusive disease, lower limb peripheral arterial disease, aortic dissection and thoracic aortic aneurysmal disease to name a few.

We proposed to undertake a systematic review to identify the existing literature base for the use of skeletal muscle index (SMI) in vascular surgery as a whole, to help take our overall research study forward.

Methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (36). The aims and methodology of the systematic review were pre-specified and registered at the International Prospective Register of Systematic Reviews (PROSPERO) (37). The registration number is CRD42022314191.

Inclusion & Exclusion Criteria

All original studies investigating the use of Skeletal Muscle Index (SMI) as a prognostic factor for surgical outcomes in patients undergoing vascular surgery were included. Conference abstracts, reviews, systematic reviews, and letters were excluded. No restrictions on language or year of publication were applied.

As explained previously, the prognostic factor of interest was SMI (defined as Skeletal Muscle Area (SMA) / Height^2), which serves as a surrogate marker of sarcopenia. The SMA (Skeletal Muscle Area) was usually measured from the cross-sectional area of skeletal muscle on an axial slice of the CT scan of a patient, taken prior to or within 30 days after the vascular surgical intervention. The cross-sectional skeletal muscle area (CSMA) was usually measured using segmentation software, and both manual and automated segmentation methods were accepted for this systematic review. No thresholds or cut-off values for sarcopenia were pre-set by the authors and the individual study's definitions of SMI cut-off values and cohort analyses was used.

Outcome measures

Outcome measures were long-term (all-cause) mortality, 30-day (in-hospital) mortality and post-operative morbidity / complications. Each individual study definition of mortality and morbidity was used for overall analysis in this systematic review, and an attempt was not made to standardise definitions across various studies, due to the potential complexities arising from this.

Search strategy

MEDLINE and EMBASE databases were searched using the Ovid® search engine (https://www.ovidsp.ovid.com). A combination of Medical Subject Headings (MeSH terms) and free-text keywords were used to formulate a comprehensive search strategy. No restrictions on language or year of publication were applied. The main keywords used were 'vascular disease', 'skeletal muscle' and 'CT'. The final literature search was conducted on 9th February 2022. Following title and abstract screening, full text articles were sourced and then selected on the basis of pre-defined inclusion and exclusion criteria. The bibliographies of the final list of studies included in the systematic review were screened for any additional studies that may have been missed using this search strategy. Details of the search strategy and specific keywords used are presented in the table below.

	Embase Classic + Embase <1947 to 2022 February 09>	**Ovid MEDLINE(R) ALL** <1946 to February 09, 2022>
1	exp vascular disease/	exp Vascular Diseases/
	AND	AND
2	computer assisted tomography/ or CT.mp. or exp segmentation algorithm/ or image segmentation/	Tomography, X-Ray Computed/ or CT.mp. or image segmentation.mp. or CT segmentation.mp.
	AND	AND
3	exp skeletal muscle/ or skeletal muscle.mp.	exp Muscle, Skeletal/ or skeletal muscle.mp.
	AND	AND
4	exp mortality/ or treatment outcome/ or exp survival/ or survival analysis/	exp Mortality/ or Mortality.mp. or treatment outcome/ or survival analysis/ or survival Rate/ or survival rate.mp.

Figure 2. Table summarising the search terms used in Embase and Medline. Each row was combined using the 'AND' function.

Study selection and management

Search results were imported and managed on COVIDENCE® (www.covidence.org) for ease of screening and study selection. Covidence is a web-based collaboration software platform that streamlines the production of systematic and other literature reviews (38).

After conducting the literature search and importing the results onto COVIDENCE®, 2 review authors (AS and MY) independently evaluated the eligibility of studies for inclusion in the systematic review. Disagreements were discussed between the 2 reviewers and an appropriate decision regarding eligibility reached, whilst following the inclusion / exclusion criteria,

assessing the study's relation to the prognostic factor in question (SMI) and ensuring the study population was undergoing vascular surgical intervention.

The following data was extracted from the eligible, selected studies - Study characteristics, patient demographics, data related to the prognostic factor (SMI / SMA), outcomes data (morbidity and mortality), and data related to the risk of bias (see sub-section below). The CHARMS checklist (CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies) was used to assist in unbiased data collection (39). This checklist is useful for identifying key data that needs to be extracted from individual studies looking at prognosis and outcomes-based research, and also to critically appraise these studies and formulate the overall research question in a systematic review. All the data collected was recorded in a Microsoft Excel® spreadsheet, and then utilised for further data and statistical analysis.

Assessment of risk of bias

The Quality in Prognosis Studies (QUIPS) tool was used to assess the risk of bias in individual, selected studies in this systematic review (40). The QUIPS tool critically appraises the validity and risk of bias in prognostic factor studies in 6 domains - Participation, attrition, prognostic factor measurement, outcome measurement, confounding, and analysis and reporting.

Each of those 6 domains contains a checklist with 3 possible ratings ('Yes', 'No', or 'Unsure') to be given by the systematic review authors' for individual sub-domains associated with each main domain. On completion of all the necessary ratings, an overall risk of bias ('high', 'moderate', or 'low') is generated for the overall systematic review and for individual research studies forming part of the systematic review.

The modified QUIPS tool was developed in 2013 to provide a focused assessment of risk of bias in prognostic factor studies, by a group of epidemiologists, statisticians and clinicians while attempting to develop a risk of bias tool for prognosis studies in lower back pain. The QUIPS tool was noted to be straightforward to use with a median completion time of 20 minutes per study and reasonable interrater agreement (k statistic) of 0.56 – 0.82 (Median 0.75) (40).

Meta-analysis

Meta-analysis of data from the studies included in this systematic review was conducted using the RevMan Version 5.3 software (41).

Pooled estimates of 30-day outcome data (mortality and morbidity) were calculated using the odds ratio (OR) or standardised mean difference (SMD) between the patient groups ('sarcopenic' vs 'non-sarcopenic') in each study cohort, and associated 95% confidence intervals (CI). For time-to-event data (long-term mortality), a meta-analysis was conducted using the inverse variance method and was presented as summary hazard ratio (HR) and associated 95% confidence intervals (CI). The relevant HR and CI was taken directly from the studies if they had been reported. Otherwise, the HR and CI was indirectly calculated using the reported p-value from the log rank test and total events, or from data extracted from survival curves. Data extraction from survival curves (Kaplan-Meier curves) was conducted using Engauge Digitizer Software published on GitHub (www.markummitchell.github.io/engauge-digitizer/) (42). This software takes graphs or images as input, and produces numerical output based on this input that can then be used for analysis. It is the opposite of using numerical or data points to produce a graph.

No attempt was made to contact the authors of included studies to retrieve any missing data.

Heterogeneity between studies was assessed with the I^2 statistical tool and standardised reference range values were used to assess the degree of heterogeneity (43):

0% - 40% - May represent little to no heterogeneity

30% - 60% - May represent moderate heterogeneity

50% - 90% - May represent substantial heterogeneity

75% - 100% - May represent considerable heterogeneity

To account for any study heterogeneity, the random effects model was chosen for the meta-analysis undertaken. Reasonable study heterogeneity and variability was expected as preliminary, 'scoping' searches had revealed the use of SMI in multiple vascular pathologies, with varied muscle segmentation methods. It was felt that the random effects model would therefore be able to provide an 'average intervention effect' (43).

Forest plots for the primary outcome measures (30-day mortality, post-operative morbidity, and long-term mortality) were created. This was done using Microsoft Excel 365® and RevMan Version 5.3 (41).

Studies attributed to having a high risk of bias in any of the 6 areas as per the QUIPS tool were removed one at a time and the statistical analysis then repeated.

Sub-group analysis was also conducted to compare outcomes data between various vascular pathologies and muscle segmentation methods to identify if there were any significant differences between these.

Results

The literature search yielded a total of 816 articles.

34 duplicate articles were identified that were excluded, resulting in a total of 782 research articles that were screened using their titles and abstracts, based on inclusion / exclusion criteria. 729 articles were excluded following this, leaving a total of 53 research articles for which full text reports were obtained and read through.

46 articles were excluded after the full-text review due to the following reasons: The prognostic factor studied was not SMI (n=40), duplicate studies (n=4) and conference / short abstracts only (n=2). A total of 7 studies met inclusion and exclusion criteria, and were included in the systematic review and meta-analysis (44–50).

Of the 40 studies / articles where SMI was not used as the prognostic factor, the alternative prognostic factor used was: Frailty scores (n=6), Skeletal Muscle Area / SMA (n=26), Skeletal Muscle Density / SMD (n=5), Skeletal Muscle Volume (n=1), Degree of fatty infiltration of skeletal muscle (n=1), and no clear prognostic factor (n=1). Whilst skeletal muscle area (SMA) can be used to calculate SMI, SMA as a prognostic factor on it's own is unlikely to be a useful predictor of outcomes due to the reasons listed in the previous section where the rationale for normalising SMA for height was highlighted and explained.

No additional studies were identified form the reference list of the final list of selected studies.

The summary of the literature search results is presented in the PRISMA flow diagram below.

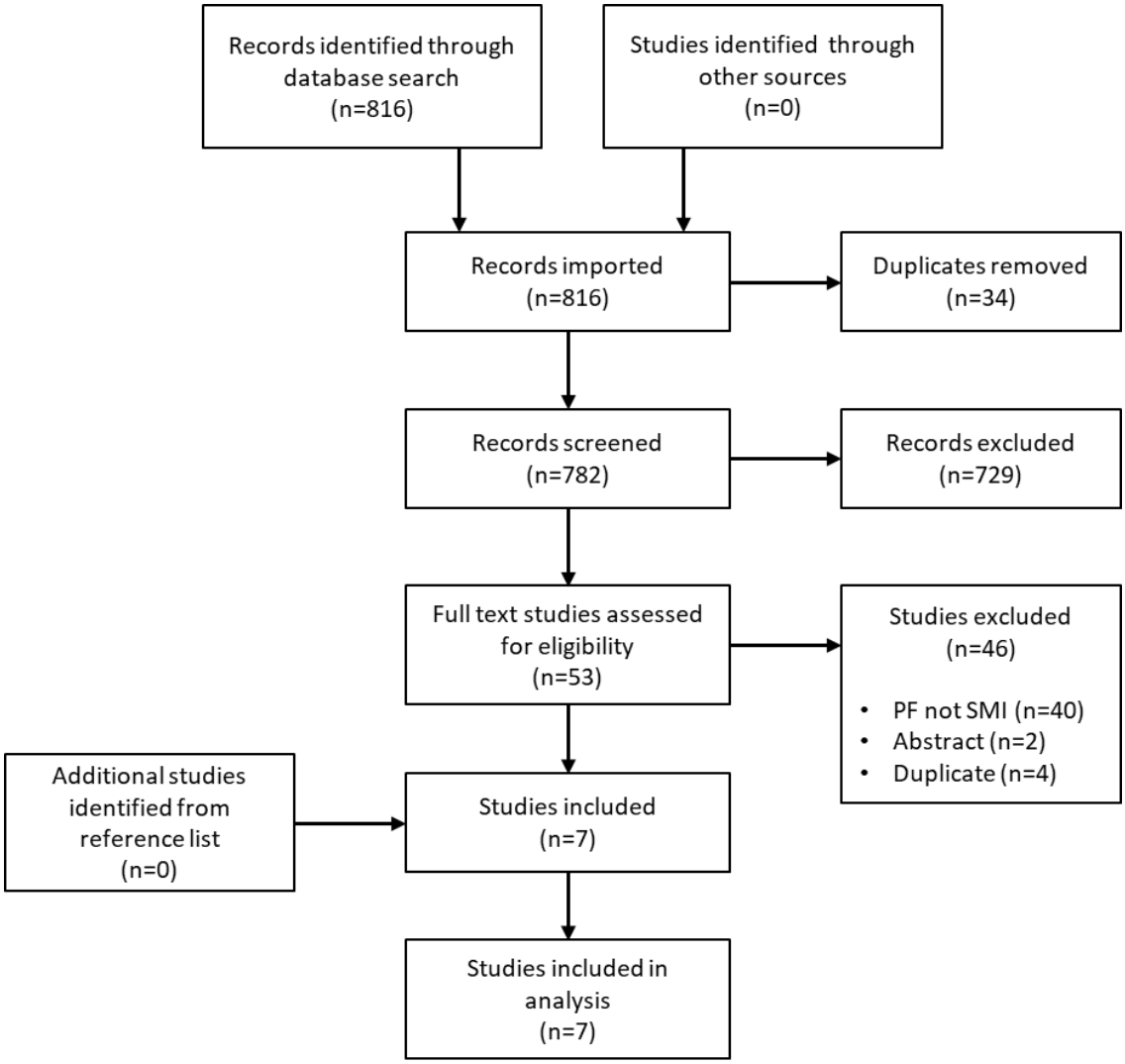


Figure 3. Study flow diagram based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).

Note: PF, Prognostic Factor. SMI, Skeletal Muscle Index.

Study characteristics

The 7 included studies which investigated the prognostic role of SMI in vascular surgery were published between 2017 and 2022. All studies were single centre retrospective studies, with vascular surgical intervention undertaken between the years 1999 and 2020. One study also included a cost analysis (47).

3 studies investigated the prognostic role of SMI in AAA repair (44,45,47), with one study focusing solely on open AAA repair (47) and another study focusing solely on EVAR (45). 4 studies investigated the prognostic role of SMI in patients with peripheral vascular disease (PVD) (46,48–50) with 2 studies focusing specifically on critical limb ischaemia (CLI) (48,49). The summary of the study characteristics is presented in the table below.

Study	Study type	Study design	Vascular pathology	Recruitment period	Total patients (Low SMI/ Normal SMI)	Low SMI cutoff (Gender)	SMI cut-off value source (Method, Source)	CT area source (Location, Muscle)	SMA calculation method
Bang J et al.	Single centre	Retrospective	AAA	Sep 1999 to Dec 2011	379 (104/275)	39.6 cm2/m2 (M) 28.6 cm2/m2 (F)	Below 2 SD, External healthy cohort	L3, SM	Semi-automatic
Cheng B.T et al.	Single centre	Retrospective	AAA	Sep 2002 to Jun 2014	272 (50/222)	500 mm2/m2 (M, F)	Optimal value from ROC, Study cohort	L3, Psoas	Manual
Kim Y.K et al.	Single centre	Retrospective	PAD	Jan 1, 2005 to Dec 31, 2012	167 (112/55)	52.4 cm2/m2 (M) 38.5 cm2/m2 (F)	NR*, external study	L3, SM	Semi-automatic
Koter S et al.	Single centre	Retrospective	AAA	2008 to 2016	156 (39/156)	58.72 cm2/m2 (M, F)	Lowest quartile, Study cohort	L3, SM	Manual
Matsubara Y et al.	Single centre	Retrospective	CLI	Jan 2002 to Dec 2012	114 (53/61)	114.0 cm2 (M) 89.8 cm2 (F)	Below 5th percentile, External healthy cohort	L3, SM	NR
Selcuk N et al.	Single centre	Retrospective	CLI	Oct 2015 to Dec 2020	217 (82/135)	5.5 cm2/m2 (M) 4.0 cm2/m2 (F)	NR*, Study cohort	L3, Psoas	Manual
Vedder I.R et al.	Single centre	Retrospective	PAD	Jan 1, 2005 to Dec 31, 2017	686 (140/546)	41.2 cm2/m2 (M) 31.5 cm2/m2 (F)	Optimal value from ROC, Study cohort	L3, SM	Semi-automatic

Figure 4. Summary table listing the 7 studies included in the systematic review and meta-analysis.

Note: M, Male. F, Female. SD, Standard deviation. ROC, Receiver Operating Characteristic curve. NR, Not reported. SM, Total skeletal muscle. Psoas, Psoas muscle. IQR, Interquartile Range.

A total of 1991 patients were included in the meta-analysis, of which 580 patients were classified as having 'low' SMI ('sarcopenic') and 1411 patients were classified as having 'normal' SMI ('non-sarcopenic').

The SMI threshold values varied between individual studies, and individual study definitions were used in this systematic review and meta-analysis to classify patients as 'low' or 'normal' SMI.

The stratification of 'low' and 'normal' SMI was based on a cut-off point using the lowest quartile in 1 study (46) and based on a receiver operating characteristic (ROC) curve in 2 studies (45,50). 2 studies used a 'healthy' external cohort to determine a cut-off point using percentiles or standard deviations (SD) (44,48).

5 studies stratified the cohorts by gender and provided gender specific cut-off points for 'low' and 'normal' SMI (44,46,48–50). All studies utilised axial CT scans of the abdomen at the L3 (3rd lumbar spine) level to identify and segment the skeletal muscle. 2 studies segmented psoas muscle area only at this level (45,49), whilst the other 5 studies used the total cross-sectional area of all skeletal muscle segmented at the L3 level. All studies included manual segmentation of the SMA and SMI calculation, of which 3 studies incorporated semi-automated segmentation of the SMA from the CT scans based on a segmentation software or program (44,46,50). No fully automated segmentation methods were identified in this systematic review.

Pre-operative patient demographics and characteristics of the study cohort are presented in the table below. Meta-analysis of this data showed that patients with 'low' SMI had significantly higher age (SMD 0.81, 95% CI 0.71- 0.92; p<0.001), lower BMI (SMD -0.81, 95% CI -0.90 - -0.70; p<0.001), and higher prevalence of chronic kidney disease (OR 1.49, 95% CI 1.02 - 2.17; p=0.04).

Study	Male gender (%)	Age	BMI	Height	Smoking (%)	Diabetes (%)	Hypertension (%)	CAD (%)	CVD (%)	COPD (%)	CKD (%)
Bang J et al.	86 vs 91	74 vs 67	21 vs 24	NR	NR	15 vs 17	81 vs 68	30 vs 32	13 vs 10	35 vs 26	11 vs 7
Cheng B.T et al.	76 vs 90	75 vs 70	25 vs 29	NR	NR	22 vs 22	78 vs 72	44 vs 52	NR	26 vs 21	22 vs 11
Kim Y.K et al.	80 vs 51	63 vs 60	21 vs 24	164 vs 162	NR	All	32 vs 33	n/a vs 27	11 vs 13	9 vs 18	46 vs 43
Koter S et al.	80 vs 84	74 vs 70	26 vs 27	173 vs 173	NR	10 vs 7	42 vs 45	21 vs 19	NR	NR	NR
Matsubara Y et al.	62 vs 75	NR	19 vs 23	NR	42 vs 41	38 vs 49	49 vs 59	NR	36 vs 23	NR	NR
Selcuk N et al.	91 vs 93	65 vs 59	23 vs 26	NR	94 vs 98	41 vs 49	56 vs 56	54 vs 44	NR	11 vs 14	2 vs 4
Vedder I.R et al.	71 vs 58	70 vs 64	23 vs 26	NR	54 vs 58	27 vs 34	61 vs 67	38 vs 36	NR	30 vs 23	9 vs 3
Pooled estimate (95% CI)	OR 1.04 (0.80 to 1.34), p=0.79	SMD 0.81 (0.71 to 0.92), p<0.001	SMD -0.81 (-0.90 to -0.70), p<0.001	SMD -0.13 (-0.38 to 0.13), p=0.34	OR 0.88 (0.59 to 1.28), p=0.5	OR 0.8 (0.61 to 1.03), p=0.09	OR 0.99 (0.79 to 1.23), p=0.96	OR 1.04 (0.797 to 1.349), p=0.79	OR 1.4 (0.90 to 2.15), p=0.13	OR 1.12 (0.82 to 1.51), p=0.48	OR 1.49 (1.02 to 2.17), p=0.04

Figure 5. Comparative data on baseline demographics and clinical characteristics (Value shown as: 'Low' SMI vs 'Normal' SMI)

Note: CAD, Coronary Artery Disease. CVD, Cerebrovascular Disease. NR, Not Reported. SMD, Standardised Mean Difference.

Study quality and risk of bias

The general quality of the studies was high with overall bias rated as low or moderate risk. The summary of the 6 risk of bias domains as per the QUIPS tool is presented below. The risk of bias for "Study Attrition" was judged to be high, primarily due loss of data during the follow-up period or missing data (as can be expected from retrospective studies). The implication of the missing data was also not discussed in most studies. The risk of bias for "Prognostic Factor Measurement" was judged low to moderate because the cut-off points or threshold values of the prognostic factor (SMI) resulted in dichotomizing a continuous variable, but this was done differently between the studies. It was derived from the study cohort in 4 studies (45,47,49,50), but derived from an external 'healthy' cohort in the remaining 3 studies (44,48). The risk of bias for "Study Confounding" was judged moderate in some studies because potential confounding factors such as age and gender were not adjusted for. The overall risk of bias of the other domains were judged to be low. All studies had adequate study participation and all, except for one study, clearly explained the methodology surrounding SMI measurement and calculation, including the software used. Appropriate statistical analyses were conducted with clear reporting of the results. One study did not provide a rationale for the SMI cut-off value, and it was not present in any previous study by the same author.

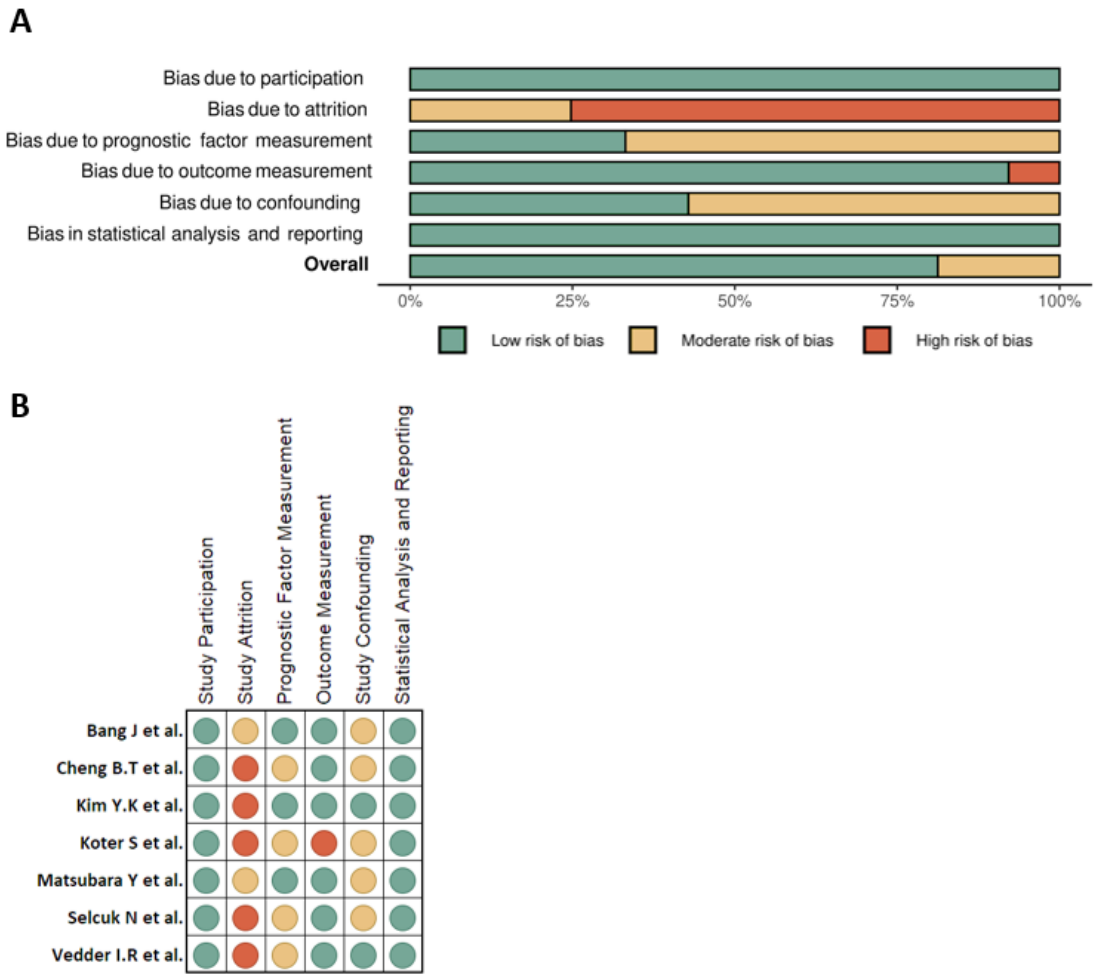


Figure 6. (A) Risk of bias graph: Review of the 6 domains of QUIPS presented as percentiles across all included studies.

(B) Risk of bias summary for individual studies.

Each of the six domains of QUIPS presented as high (red), moderate (yellow), or low (green) risk of bias.

The full QUIPS risk of bias checklist results are presented in the graph below.

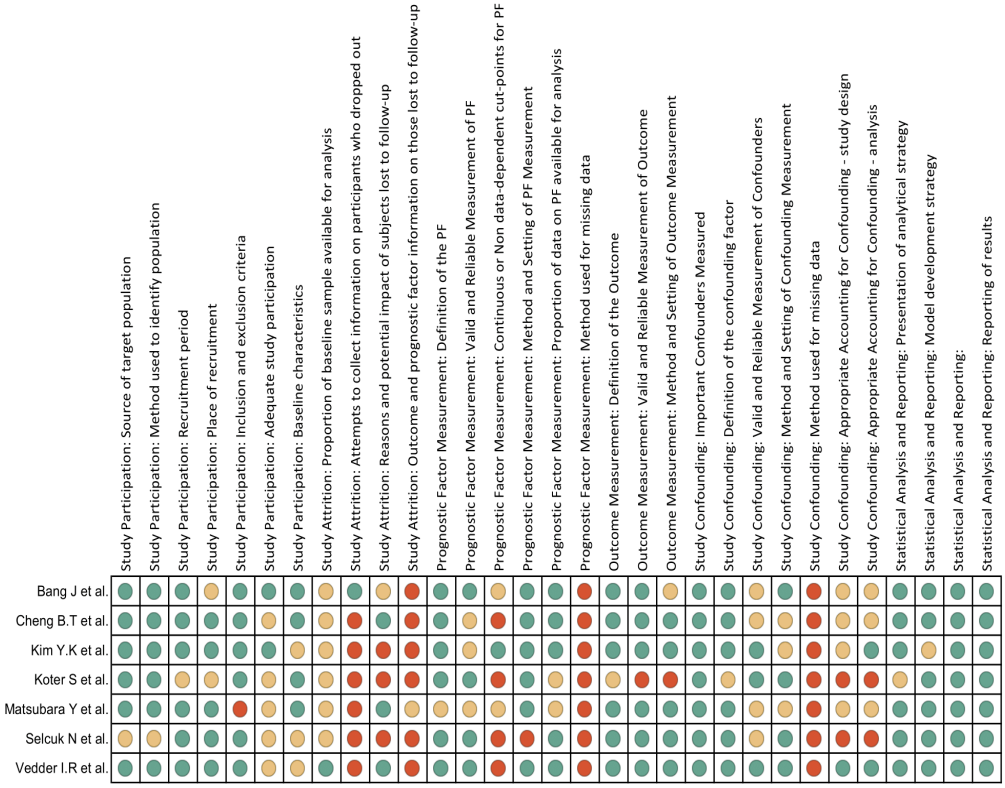


Figure 7. Quality in Prognostic Studies (QUIPS) detailed checklist summary for individual studies risk of bias.

Meta-analysis of 'Normal' Vs 'Low' SMI in Post-operative Outcomes in Vascular Surgery

30-Day or In-Hospital Mortality

4 studies reported 30-day mortality or in-hospital mortality, with 224 patients classified as having 'low' SMI and 535 patients classified as having 'normal' SMI (45,47–49). Patients with low SMI had a significantly higher risk of 30-day or in-hospital mortality as compared to patients with normal SMI (OR 12.74, 95% CI 3.29 – 49.39; $p<0.001$). The statistical heterogeneity was low ($I^2 = 0\%$; $p = 0.94$).

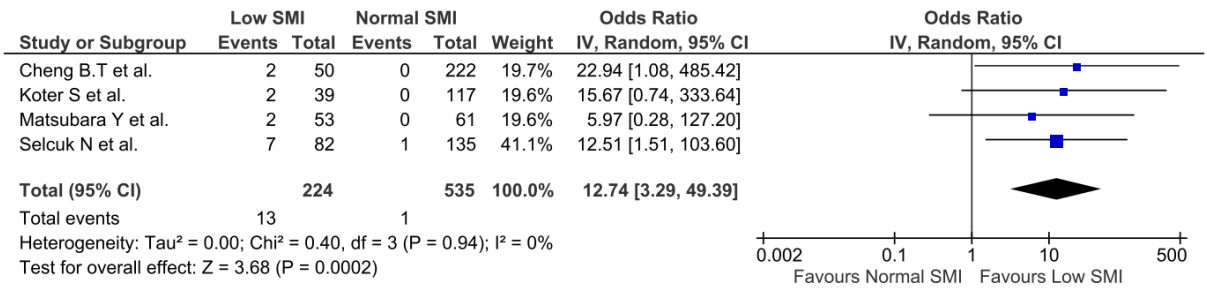


Figure 8. Forest plot of 30-day or in-hospital mortality in patients with Normal vs Low SMI undergoing vascular surgery.

The solid squares denote the odds ratios (ORs), the horizontal lines represent the 95% confidence intervals, and the diamond denote the pooled ORs.

Note: SMI, Skeletal Muscle Index. IV, Inverse Variance. CI, Confidence Interval.

Post-Operative Morbidity / Complications

4 studies reported post-operative morbidity, with 275 patients classified as having 'low' SMI and 749 patients classified as having 'normal' SMI (44,45,47,49). Patients with low SMI had a significantly higher risk of post-operative morbidity compared to patients with normal SMI (OR 6.21, 95% CI 1.31 – 29.94; p<0.001). The statistical heterogeneity was considerably high (I^2 = 86%; p<.001) and this was likely due to the variations in definitions of post-operative morbidity and complications amongst various studies.

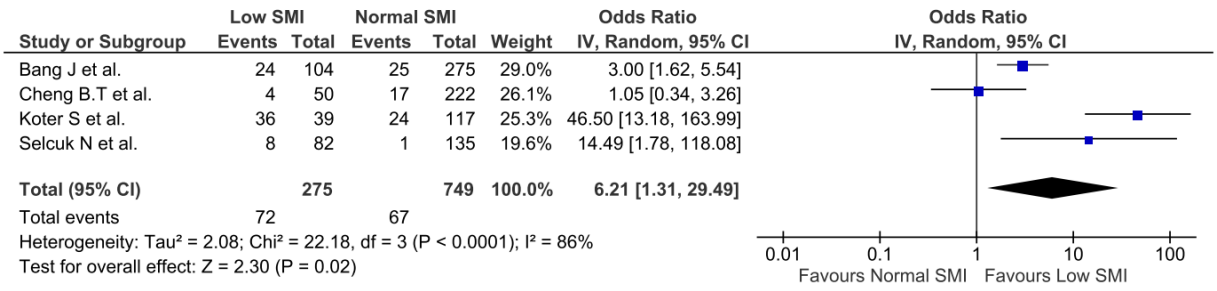


Figure 9. Forest plot of post-operative morbidity in patients with Normal vs Low SMI undergoing vascular surgery.

Note: SMI, Skeletal Muscle Index. IV, Inverse Variance. CI, Confidence Interval.

All-cause Long-term Mortality

6 studies reported time-to-event data for long-term mortality, with 498 patients classified as having 'low' SMI and 1276 patients classified as having 'normal' SMI (45–48,50). Patients with low SMI had a significantly higher long-term mortality risk compared to patients with normal SMI (HR 2.80, 95% CI 2.25 – 3.49; p<0.001). The statistical heterogeneity was low (I^2 = 0%; p = 0.64).

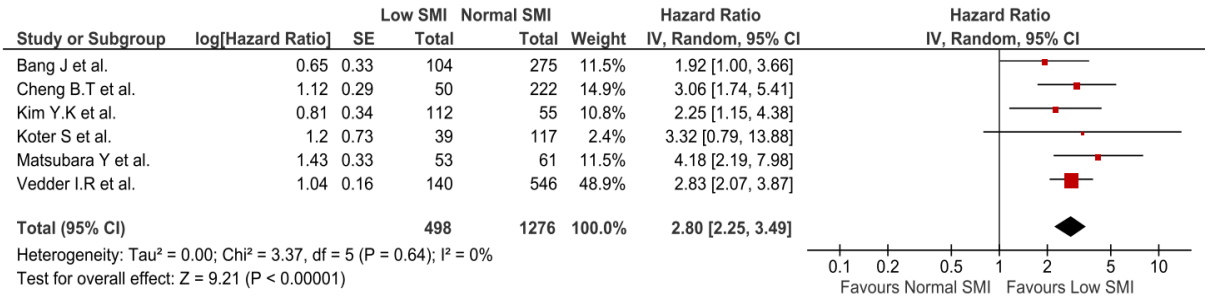


Figure 10. Forest plot of all-cause long-term mortality in patients with Normal vs Low SMI undergoing vascular surgery.

The solid squares denote the hazard ratios (HRs), and the diamond denotes the pooled HR.

Note: SMI, Skeletal Muscle Index. SE, Standard Error. IV, Inverse Variance. CI, Confidence Interval.

Sub-group analysis

Removing one study at a time did not affect the significance of the outcomes. Only one study had a high risk of bias in 2 domains of the QUIPS risk of bias tool, but removal of this study from the analysis did not alter the results of the meta-analysis significantly.

Sub-group analysis was undertaken of patients undergoing surgical intervention for AAA and peripheral arterial disease (PAD), and also to compare methods of muscle segmentation (manual vs semi-automated) to assess if there were any significant differences in these groups.

The meta-analysis results obtained were broadly similar to the overall cohort, but did not achieve statistical significance for 30-day (post-operative) morbidity in patients undergoing AAA repair, and manual segmentation for 30-day (post-operative) mortality.

Detailed results of the sub-group analysis are shown below.

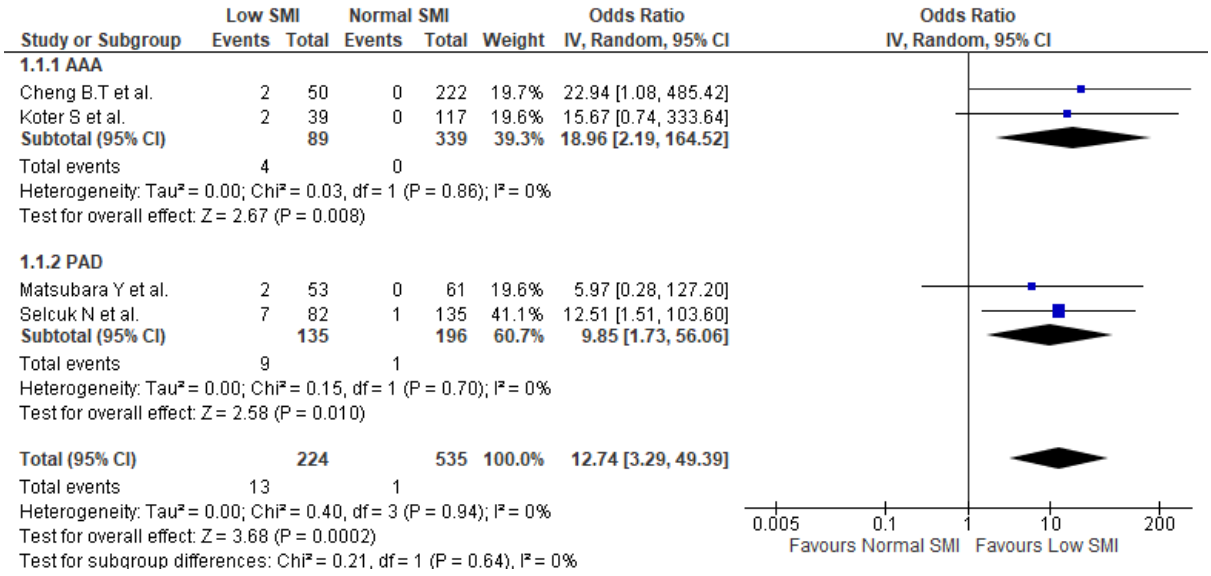


Figure 11. Sub-group analysis of patients undergoing intervention for AAA vs PAD, comparing 30-day mortality in patients with Normal vs Low SMI.

The solid squares denote the odds ratios (ORs), the horizontal lines represent the 95% confidence intervals, and the diamonds denote the pooled ORs. *Note:* SMI, Skeletal Muscle Index. IV, Inverse Variance. CI, Confidence Interval.

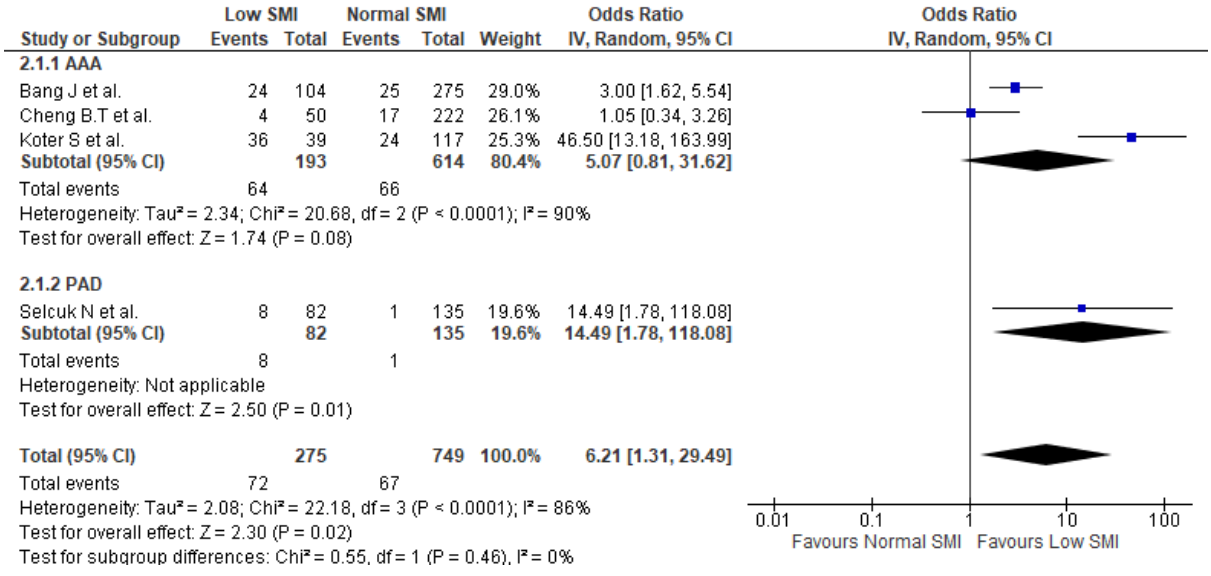


Figure 12. Sub-group analysis of patients undergoing intervention for AAA vs PAD, comparing 30-day morbidity in patients with Normal vs Low SMI.

Note: SMI, Skeletal Muscle Index. IV, Inverse Variance. CI, Confidence Interval.

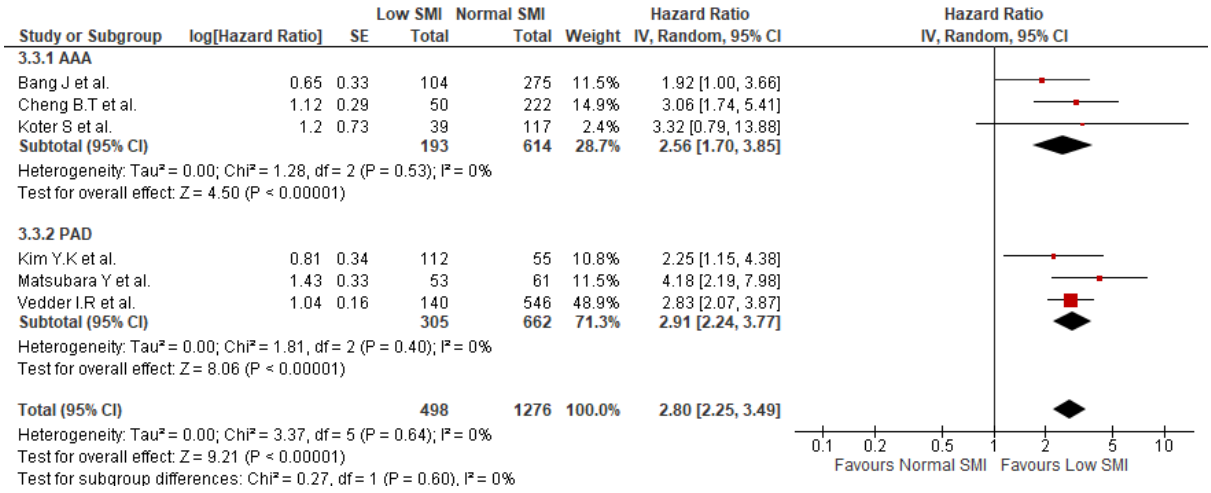


Figure 13. Sub-group analysis of patients undergoing intervention for AAA vs PAD, comparing all-cause long-term mortality in patients with Normal vs Low SMI.

The solid squares denote the hazard ratios (HRs), and the diamonds denote the pooled HRs.

Note: SMI, Skeletal Muscle Index. SE, Standard Error. IV, Inverse Variance. CI, Confidence Interval.

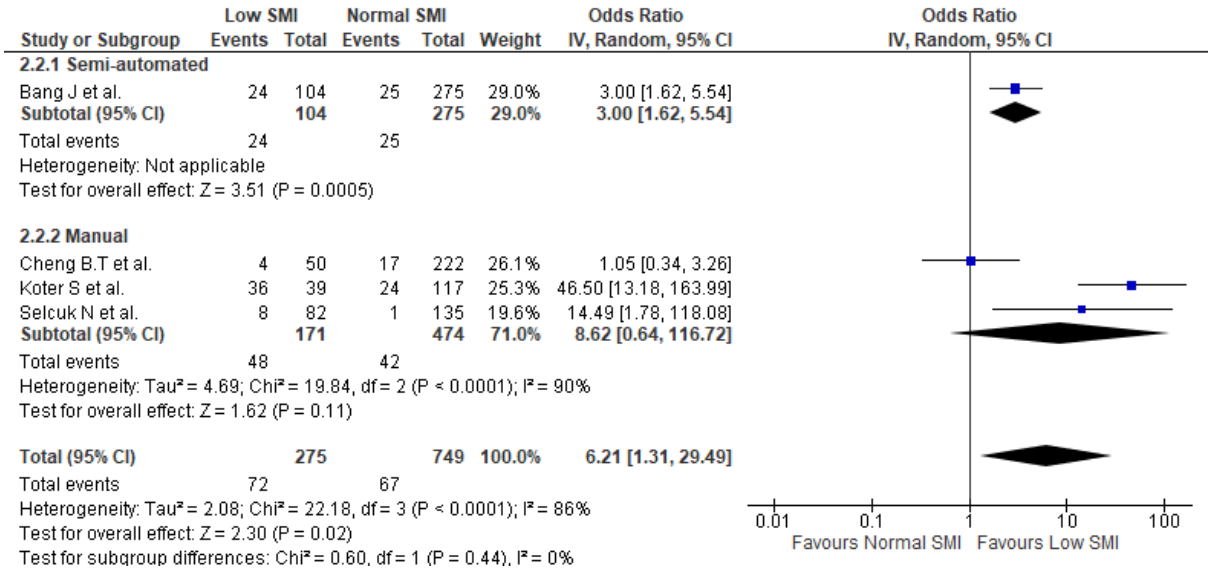


Figure 14. Sub-group analysis of Manual vs Semi-automated SMI calculation, comparing 30-day mortality in patients with Normal vs Low SMI undergoing vascular surgery.

The solid squares denote the odds ratios (ORs), the horizontal lines represent the 95% confidence intervals, and the diamonds denote the pooled ORs. *Note:* SMI, Skeletal Muscle Index. IV, Inverse Variance. CI, Confidence Interval.

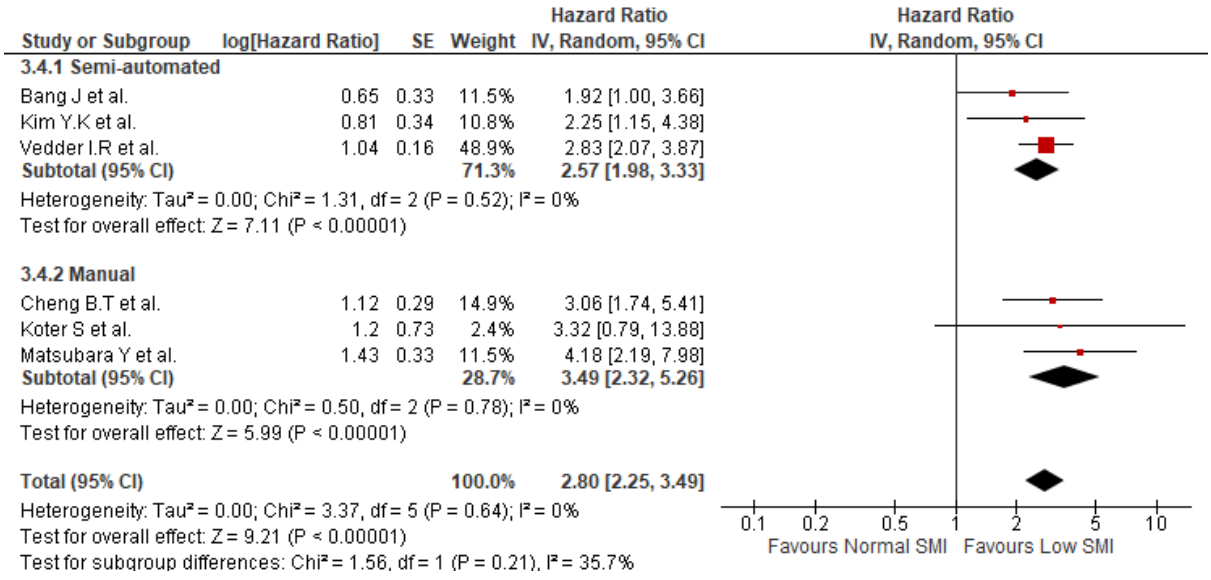


Figure 15. Sub-group analysis of Manual vs Semi-automated SMI calculation, comparing all-cause long-term mortality in patients with Normal vs Low SMI.

The solid squares denote the hazard ratios (HRs), and the diamonds denote the pooled HRs. *Note:* SMI, Skeletal Muscle Index. SE, Standard Error. IV, Inverse Variance. CI, Confidence Interval.

Discussion

Meta-analysis of the 7 included studies in this systematic review demonstrates that patients with 'low' SMI who underwent vascular surgical intervention had a significantly higher risk of short- and long-term mortality, along with an increased risk of post-operative complications.

All the included studies, except one study (47) represented pragmatic, real world vascular surgical practice and directly investigated the prognostic role of sarcopenia and SMI in vascular surgery. Koter et al (47) primarily performed a cost analysis of sarcopenia in patients undergoing open aortic surgery. Though the studies included patients undergoing a variety of vascular surgical interventions for differing pathologies, the primary outcome measures demonstrated similar results with low heterogeneity, suggesting that outcomes were not treatment-specific or disease-specific, but in fact 'sarcopenia-specific'.

3 studies (44,48,49) primarily investigated the prognostic role of SMI and it's association with a specific complication. 2 studies (44,48,49) looked at the association with post-operative cardiovascular events and 1 study looked at the association with post-operative acute kidney injury (44,48,49). As the definitions and 'inclusion / exclusion' criteria for post-operative complications varied between studies, this resulted in considerably high heterogeneity in the meta-analysis for post-operative morbidity. This could be avoided by having standardised,

internationally accepted definitions for these complications, or having internationally accepted and uniform reporting standards in vascular surgical practice.

It should also be noted that all patients included in this systematic review and meta-analysis underwent surgical intervention, and there was no comparative 'control' group with conservatively managed patients to assess differences in survival outcomes. A study investigating the prognostic role of low psoas muscle area (PMA) in patients with asymptomatic infrarenal AAA showed no significant difference in survival between operated and non-operated patients (51). However, SMI was not formally calculated by normalising the psoas muscle area for height and the limitations of using SMA on it's own have been highlighted previously. Regardless, the role of sarcopenia in conservatively managed patients needs to be assessed further. Development of threshold values for sarcopenia may also help in deciding which patients would benefit the most from intervention, and which patients should be conservatively managed.

All the studies in this systematic review chose to calculate SMI by measuring cross-sectional muscle area at the L3 level. Whilst this may be beneficial to improve reproducibility in future studies by standardising the study methodology, there may be more precise alternative measures of diagnosing sarcopenia and measuring SMI, which need further investigation. Use of quadriceps or thigh muscle is gaining more traction as an alternate surrogate marker of sarcopenia, and may even be more sensitive to changes in overall total body muscle quantity and composition, than skeletal muscle in the trunk or torso (33).

However, most vascular surgical patients, particularly those planned for AAA repair will routinely undergo a CT scan prior to surgical intervention that includes the skeletal muscle at the L3 level, and therefore, the measurement of SMI based on skeletal muscle at the L3 level may be a more convenient method to produce a risk prediction score or model of peri-operative

morbidity and mortality (52). However, alternative methods of measuring sarcopenia and SMI may need to be considered for lower limb vascular pathology (or upper limb, thoracic and cerebrovascular pathology) as pre-operative imaging may not include the skeletal muscle at the L3 level.

How do we define a 'low' and 'normal' SMI?

Determination of SMI threshold and cut-off values is one of the areas future research projects need to focus on to allow universal standardisation to derive meaningful conclusions from sarcopenia research studies. As demonstrated in this systematic review, the definition of 'low' SMI varied between all the studies, and cut-off values were determined in a variety of different ways, with some studies deriving it from their internal cohort and some utilizing external 'healthy' cohorts as controls. A population-wide large scale study will help determine optimal cut-off values, but this would require significant resources to set-up and undertake.

Although the European Working Group on Sarcopenia in Older People (EWGSOP) recently revised their definition of sarcopenia and produced a detailed guidelines and review document (33), this document highlighted the lack of consistency in setting sarcopenia cut-off values.

Potential bias in the review process

The PRISMA guidelines were followed to optimize this systematic review and meta-analysis to conform to internationally accepted standards. The extensive literature search aimed to

ensure that all studies assessing the role of SMI in vascular surgery were included. However, some limitations of this systematic review and meta-analysis are still present.

As expected from retrospective studies, some data was lost to follow-up in the included studies in this systematic review and meta-analysis. No prospective studies were identified that evaluated the role of skeletal muscle index (SMI) in vascular surgery. Therefore, future consideration needs to be given to the development of a robust prospective study or randomised controlled trial (RCT) to optimise study methodology, have a complete dataset and produce more reliable, generalisable results.

In the case of missing data from the included studies, no attempt was made to contact the authors of the studies to try and retrieve the missing data.

Comparison with other systematic reviews

There are no previous systematic reviews investigating the prognostic role of SMI in vascular surgery. However, 2 recent reviews have investigated the prognostic role of sarcopenia using skeletal muscle area (SMA) in abdominal aortic aneurysmal disease (53) and peripheral vascular disease (54).

The systematic review investigating the role of skeletal muscle mass in patients undergoing AAA repair included 7 studies, of which only 2 studies normalised skeletal muscle area (SMA) values for height to calculate SMI (these studies were also included in our systematic review). The remaining 5 studies used SMA with varying cut-off values to assess outcomes (not normalising for height), and the limitations of not normalising SMA with height to produce SMI have been explained previously. The meta-analysis included 1440 patients, with 374

patients classified as having ‘low’ skeletal mass. Patients with ‘low’ skeletal mass had a significantly higher risk of long-term mortality compared to those patients with ‘normal’ skeletal muscle mass (HR 1.66, 95% CI 1.15 – 2.40; p = 0.007). However, no significant difference was noted in peri-operative mortality and morbidity, but only 2 studies were included in this meta-analysis, thereby limiting it’s overall applicability.

The systematic review and meta-analysis investigating the role of skeletal muscle mass in patients undergoing lower limb revascularisation for PVD included 6 studies, with a total of 1235 patients, where 440 patients were classified as having ‘low’ skeletal muscle mass. The meta-analysis showed no significant difference in peri-operative or long-term mortality between patients with low and normal skeletal muscle mass, but improved amputation-free survival and limb salvage rates in patients with ‘normal’ skeletal muscle mass. 4 of the included studies in this systematic review and meta-analysis were published in Japan, of which 2 studies were from the same author, potentially leading to some selection bias in the overall results. Variable cut-off values were also used for stratifying patients into ‘low’ and ‘normal’ skeletal muscle mass, and only one study normalised skeletal muscle area for height.

Generally speaking, these similar reviews had broadly similar results to the findings of our systematic review and meta-analysis. Low skeletal muscle mass (whether normalised for height or not) as a surrogate marker for sarcopenia confers poorer outcomes for patients undergoing vascular surgical intervention. Normalising skeletal muscle area for height to produce skeletal muscle index (SMI) probably produces a more accurate measure of sarcopenia, as it accounts for overall patient body size. Further research is needed to produce optimal threshold values for SMI, and also to confirm the superiority of use of SMI over SMA.

Implications for current clinical practice

This systematic review and meta-analysis demonstrates that skeletal muscle index (SMI) may be used as a significant predictor of outcomes after vascular surgical intervention. SMI derived from measuring the cross-sectional skeletal muscle area (CSMA) at the L3 level from a pre-operative CT scan of a patient planned for vascular surgical intervention may be incorporated into surgical decision making and risk prediction models to stratify patients on the basis of peri-operative risk to guide optimal decision making and help assist patients in making informed choices.

However, the optimal threshold value for 'low' SMI remains unclear and further adjustments for possible cofounding factors, such as sex and BMI may be needed to further refine it's use in day-to-day clinical practice.

Avenues for further research

Besides those already mentioned above, another key factor that needs consideration is the time taken to produce SMA / SMI values for each patient.

Manual and semi-automated segmentation of the skeletal muscle at the L3 level were both shown to be appropriate at measuring SMA (and subsequently SMI). However, the time taken to undertake muscle segmentation for each, individual patient is significant and not to be underestimated. Due to the time taken to segment each CT scan, incorporation of this process and calculation is unlikely to be possible in routine clinical workflow, especially when time and existing resources are already stretched to the maximum. In today's rapidly evolving era

of artificial intelligence (AI), the use of automation may help in speed up and streamline this process, and perhaps even produce a value or measurement with every CT scan undertaken in clinical practice.

We, therefore, explored the possibility of using a fully automated segmentation model to assess whether it produced comparable measurements to manual segmentation, and whether it could be used to predict outcomes in patients undergoing AAA repair, as demonstrated in subsequent chapters of this document.

3. THE USE OF AN AUTOMATED SARCOPENIA MEASUREMENT TOOL AS A PREDICTOR OF OUTCOME IN PATIENTS WITH ABDOMINAL AORTIC ANEURYSMS (SPAR)

Abstract

Introduction

Skeletal muscle index (SMI) at the 3rd lumbar vertebral level (L3) is a surrogate marker of sarcopenia and accurately predicts outcomes in a variety of surgical specialties, including vascular surgery. Manual measurement of SMI is a time-consuming process, preventing it's incorporation in routine clinical workflow. We explore the feasibility of using a fully automated deep-learning section-based muscle segmentation tool in calculating SMI and predicting outcomes in patients undergoing AAA repair.

Methods

Pre-operative CT scans of patients undergoing AAA repair between 2016 and 2020 were run through and analysed by the automated sarcopenia segmentation tool, with 20% of scans undergoing manual segmentation to serve as 'ground truth'. The primary outcome measure was the correlation between SMI at L3 and 30-day composite outcome, which was a combination of death and defined post-operative complications.

Results

280 CT scans underwent successful automated segmentation, with excellent IRR noted between the manual and automated segmentations with an ICC 0.96 (95% CI 0.95 – 0.97, $p < 0.05$) for skeletal muscle area and ICC 0.98 (95% CI 0.98 – 0.99, $p < 0.05$) for skeletal muscle density (SMD).

On regression modelling of SMI and 30-day composite outcome, good negative correlation was noted amongst women undergoing open AAA repair ($r = -0.86$ / $R^2 = 0.75$, $p<0.05$).

Men undergoing open AAA repair in the lowest SMD quartile had poorer 30-day composite outcomes as compared to the highest quartile ($t(34) = 2.741$, $p = 0.01$).

Conclusion

This automated sarcopenia segmentation tool provides a rapid and accurate measure of SMI and SMD, and may help produce a sarcopenia score to guide clinicians on optimal treatment strategies and improve post-operative outcomes by identifying high-risk patients.

Background

Following on from the systematic review and meta-analysis which identified skeletal muscle index (SMI) as a key predictor of outcomes in vascular surgery, we planned to design and undertake a research study to further evaluate the use of SMI as a predictor of outcomes in a cohort of patients undergoing surgical repair of abdominal aortic aneurysms (AAA) at our institution.

We aimed to explore whether segmentation of skeletal muscle (manual and automated) at L3 level can be used to predict 30-day composite outcomes (mortality and defined complications as per the NVR) in patients undergoing abdominal aortic aneurysm repair.

As has been demonstrated previously, skeletal muscle segmentation and calculation of skeletal muscle area (SMA) from cross-sectional imaging can be a useful marker for possible pre-

operative risk stratification, determining the optimal choice of intervention (conservative vs endovascular vs open surgery) and predicting post-operative outcomes. However, manual muscle segmentation by the responsible clinician is a time-consuming process and is not feasible in routine clinical workflow, which has mainly confined it's use to research studies. This has also led to attempts at partially or completely automating this process to try and improve it's accessibility / applicability and allow incorporation into routine clinical workflow, though this is still in it's early stages of development.

One such fully automated deep-learning section-based muscle segmentation tool for sarcopenia assessment has been developed and validated at the Comprehensive Cancer Imaging Centre at Imperial College London (55).

This automated segmentation tool was designed to undertake opportunistic screening for frailty and utilised over 1,000 CT scans (a mixture of publicly available data sets and ICHT scans) as part of the development and training process. The segmentation tool has 2 sequential steps – Identification and selection of the appropriate axial CT slice for segmentation, followed by segmentation of all the skeletal muscle on this selected CT slice. Both these processes utilised fully convolutional neural networks (FCNN), based on a U-Net-like architecture.

The middle of L3 (3^{rd} lumbar vertebral level) was chosen as the axial level for muscle segmentation and as demonstrated in the systematic review (Chapter 2), the L3 level is a commonly used reference level for calculating skeletal muscle area and index. This automated segmentation tool segmented all the skeletal muscle at the L3 level, which included the erector spinae, psoas, and rectus abdominis muscle groups, and produced the output in < 1 second per CT examination.

However, this fully automated segmentation tool has been developed and validated using a cohort of oncology patients, and has not been used previously on patients undergoing vascular surgery, either to test validity or to perform opportunistic screening for frailty and obtain a 'sarcopenia score'.

This gave rise to the idea of utilising this segmentation tool on a retrospective cohort of patients who had undergone abdominal aortic aneurysm (AAA) repair to test validity of this automated segmentation tool and whether it could be used to predict outcomes in this cohort. Patients undergoing AAA repair were chosen from amongst all vascular surgical patients as these patients routinely undergo cross-sectional imaging of the abdomen / pelvis, thereby including the L3 level, as compared to other groups of vascular surgical patients such as those with peripheral vascular disease or cerebrovascular disease.

The findings of this feasibility study would also help determine whether this segmentation tool should be incorporated into routine clinical workflow or not. If successful, a prospective multicentre trial could then be designed to evaluate this automated segmentation tool and further improve it's credibility and validity.

Study Objectives

The primary objective of this study was to test the validity and reliability of the automated segmentation tool in a cohort of vascular surgical patients who had already undergone AAA repair. For the purposes of this study, we shall call this the 'sarcopenia segmentation tool'.

In terms of secondary objectives, this study aims to correlate sarcopenia with other patient factors such as age, gender, length of stay (LOS), discharge destination (home or other), blood

tests (Haemoglobin / Albumin protein / CRP / Creatinine / Glucose / Cholesterol), osteoporosis score, and body mass index (BMI).

Another objective of the wider study was to explore the role of change of muscle bulk over time on overall patient outcomes, and whether an alternate muscle group (such as the quadriceps) could be used as a surrogate marker of sarcopenia / frailty. The use of quadriceps muscle bulk on CT imaging to measure frailty and sarcopenia has not been undertaken previously, but is particularly relevant to vascular surgery patients who have undergone AAA repair as cross-sectional imaging usually goes down to the level of the mid-thigh to assess access vessel status (common iliac, external iliac, common femoral, profunda femoris and superficial femoral vessels) in case endovascular intervention is planned. If found relevant and usable as an alternative to skeletal muscle segmentation at the L3 level, the quadriceps could potentially then be used to measure frailty / sarcopenia in patients undergoing lower limb vascular intervention, where cross-sectional imaging may not include skeletal muscle at the L3 level.

In addition to these clinical objectives, there are also some machine learning objectives of this study. The study aims to validate the accuracy of the automated sarcopenia segmentation tool (using DICE similarity coefficient) against ground truth segmentation and also to develop a preliminary tool to automate the calculation of skeletal muscle index (SMI) using patient height and muscle area at the L3 level (derived from the sarcopenia segmentation tool).

The DICE score is defined as twice the area of the overlap between the predicted segmentation and the ground truth segmentation divided by the sum of both areas.

This study will adhere to the principles outlined in the UK Policy Framework for Health and Social Care Research. It will be conducted in compliance with the study protocol, the Data Protection Act and other regulatory requirements as appropriate.

Methodology

The study was titled 'SPAR' – The Use of An Automated Sarcopenia Measurement Tool as a Predictor of Outcome in Patients with Abdominal Aortic AneuRysms.

The study was planned as a retrospective cohort study utilising data (CT scans and clinical / demographic data) from a cohort of 500 consecutive patients who had undergone AAA repair at Imperial College Healthcare NHS Trust (ICHT) between the years 2016 and 2020. The patients included in this study were identified from the National Vascular Registry (NVR) list of AAA repairs undertaken at ICHT, which is submitted and prepared 'in-house' by the ICHT clinical vascular team. The year 2016 was chosen as the cut-off year as this was when electronic patient records commenced at ICHT.

This study's aim was to diagnose sarcopenia on pre-operative CT scans of patients, using the 'sarcopenia segmentation tool' described earlier in this chapter (55). This will allow automated identification of the middle of L3 vertebral level on the axial slices in the CT scan series, followed by fully automated segmentation of the skeletal muscle area at this level. Mean muscle attenuation (in Hounsfield units) will also be determined by this automated tool to assess muscle density, which is thought to be a surrogate marker of muscle quality (as described in Chapter 2).

The skeletal muscle index (SMI) will then be calculated, by normalising skeletal muscle area (SMA) for height. These automated sarcopenia outputs (SMA and skeletal muscle attenuation)

along with the skeletal muscle index (SMI) will then be analysed in conjunction with clinical patient specific data from the NHS trust electronic patient records and the National Vascular Registry (NVR) database records, to assess the validity and applicability of this sarcopenia segmentation tool in clinical workflow.

The following clinical data was collected from patient records:

Pre-intervention data

- Age, Sex, Ethnicity, BMI (Height and Weight)
- Osteoporosis score (calculated from the DEXA scan)
- Co-morbidities (Hypertension, Diabetes, Coronary Artery Disease, Chronic Kidney Disease, Respiratory Disease or Cerebrovascular Disease)
- Smoking status (Current smoker, Ex-smoker, Never smoked, Not known)
- Pre-operative abdominal aortic aneurysm size (derived from the last scan on record, prior to intervention)
- Blood test results (Haemoglobin / Albumin protein / CRP / Creatinine / Glucose / Cholesterol / Triglycerides) – The blood test immediately prior to the date of surgical intervention was recorded in case there was more than 1 blood test result available on electronic patient records in a 12-month period prior to the date of surgical intervention.
- Medications (specifically looking at antiplatelet, anticoagulant or statin therapy)

Intervention data

This included data on the date and 'type' of aortic intervention. Each individual patient's AAA repair was classified into one of the following 5 categories –

- Open infrarenal repair

- Open juxtarenal repair (proximal aortic control required the placement of the aortic clamp above one or both of the main renal arteries)
- Open thoraco-abdominal aneurysm repair (proximal aortic control required the placement of the aortic clamp in the thoracic aorta)
- Infrarenal EVAR
- Complex EVAR (including fenestrated – FEVAR and branched – BEVAR repairs, and all EVARs necessitating some form of renal intervention / stenting such as the use of chimneys)

Patients who had only undergone TEVAR (Thoracic Endovascular Aortic Repair) for treatment of thoracic aortic aneurysms, trauma or dissections were excluded unless the TEVAR was performed as part of a complex EVAR to treat a complex thoracoabdominal aneurysm.

In patients undergoing staged intervention for treatment of AAA (particularly in complex endovascular repair for thoracoabdominal aneurysms) clinical, imaging and complication metadata was collected around the 1st procedure, and associated hospital stay.

<u>Post-intervention data</u>

- Hospital length of stay (days)
- Need for critical care and length of stay in critical care (days)
- In-hospital post-operative mortality
- NVR defined complications (Cardiac, Respiratory, Hemorrhage, Visceral ischaemia, Limb ischaemia, Renal failure, Stroke, Post-operative Confusion / Delirium or Wound Infection / Dehiscence) or absence of these complications
- Re-admission within 30-days (and the associated reason/s for this)
- Infection
- All-cause (long-term) mortality

- Presence or absence of ruptured AAA
- Need for re-intervention (and the associated type e.g. open / endovascular, and reason)
- Post-intervention CT scan data

Study data was collected and managed using REDCap electronic data capture tools hosted at Imperial College London. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources (56,57).

A REDCap database was set up specifically for this study, under the auspices of Imperial College London, for relevant data collection and analysis. Each patient included in this study was given a unique patient enrolment number, specific to this study. The clinical data collected for each individual patient in this study will, therefore, be anonymised by incorporating this within the umbrella of each patient's unique patient enrolment number.
Clinical data collection commenced in August 2021, and was completed in May 2023.

CT scan data

The CT scan of the abdominal aorta with contrast, immediately prior to intervention was identified and pseudonymised for analysis. There was no set time period prior to surgical intervention during which this CT scan should have happened as this study was retrospective in nature reflecting 'real-world' practice.

If the included patient had a CT scan of the abdominal aorta with contrast at least 6 months post date of surgical intervention, this was also identified and pseudonymised for analysis. This CT scan would be utilised to assess change in muscle bulk in the post-operative period and it's associated implications in patient outcomes, which was another secondary objective of this study.

Pseudonymisation of identified CT scans (pre- and post-intervention) was undertaken to enable its utility for research and segmentation purposes, and protect patient confidentiality as required for ethics approval of this study. Individual patient consent was not undertaken or possible in this study due to it's retrospective nature, but all patient data was collected, stored, and analysed as per university (Imperial College London) and NHS hospital trust (ICHT) protocols.

Each patient included in this study would, therefore, have at least 1 pseudonymised scan, but may have 2 pseudonymised scans if a suitable CT scan 6 months post-intervention was available. Some patients may not have a post-intervention scan at the 6-month mark, particularly if they underwent open repair of their AAA where CT surveillance is usually not required or if they had an early CT scan before the 6-month mark and were switched to ultrasound surveillance in the case of endovascular repair. Some patients, particularly those presenting emergently with ruptured aneurysms, may not have survived to the 6-month mark, and would, therefore, not have suitable post-operative CT imaging for pseudonymisation.

These pseudonymised scans were labelled with the unique patient enrolment number as well as the date of the CT scan. Pseudonymisation of CT scans commenced in November 2021 and was completed in July 2023.

As part of this pseudonymisation process, all patients who underwent AAA repair at ICHT between 2016 and 2020 (as identified from the NVR) had their individual cross-sectional CT imaging reviewed on the ICHT clinical picture archive (PACS) and compared to their clinical data, particularly the date of surgical intervention for AAA repair. Following identification of the relevant pre-intervention and / or post-intervention CT scan, a secure Microsoft Excel® spreadsheet was completed with these details along with patient identifiers and the unique patient enrolment number and sent to the imaging research department at ICHT. As a significant proportion of CT scans in routine clinical practice in vascular surgery are undertaken in 'spoke' hospitals (or non-tertiary hospitals with no provision for undertaking vascular surgical intervention), the location of where the relevant CT scan was undertaken was also provided to try and retrieve these scans for pseudonymisation purposes. Whilst these CT scans may not have been undertaken at one of ICHT's hospitals, they are usually available or can be made available on the ICHT PACS server using the Image Exchange Portal (IEP) system.

Pseudonymisation was thereafter undertaken by the ICHT imaging research team according to a pre-specified standard operating procedure (SOP) and de-identified CT scans sent back using an encrypted folder. These pseudonymised scans were returned in the DICOM (Digital Imaging and Communications in Medicine) format which utilises a non-proprietary data interchange protocol, digital image format and file structure for medical images. This DICOM format allows inter-operability and usability between various imaging systems and software, whether for clinical or research purposes (58).

The de-identified pseudonymised CT scans (DICOM data) were then run through the automated sarcopenia segmentation tool, based at the Hammersmith Hospital (part of ICHT). Manual identification of errors was performed, and final corrected segmentations were used as automated outputs for the study analysis of results and correlation with patient specific clinical

outcomes. Failure cases and corrected cases were also recorded, and may be used to improve and train the existing deep-learning algorithm and sarcopenia segmentation tool, to further enhance it's accuracy and useability. Automated segmentation using the sarcopenia segmentation tool was undertaken between June 2023 – August 2023.

A proportion of cases (20% of the study cohort - 100 patients) were planned to undergo manual segmentation of skeletal muscle on an axial CT slice at the middle of L3 level to provide 'ground truth' for evaluation of the accuracy of the sarcopenia segmentation tool. A decision was made not to perform manual segmentation of CT scans of all patients included in the study as the sarcopenia segmentation tool has previously been validated (using and when compared to manual segmentation) in a cohort of over 1000 CT scans, both for appropriate axial slice detection and segmentation area and attenuation cross-validation (55). Time taken for manual segmentation of each individual CT scan (10 – 15 minutes / scan) was also a factor in the decision not to perform manual segmentation of all scans, along with the need for pseudonymisation and large volume data transfer between the researcher/s and the imaging research department.

Manual segmentation was performed using semi-automated software (ITK - SNAP). ITK-SNAP (59) is a free, open-source, multi-platform software application used to segment structures in cross-sectional medical images using active contour methods, along with manual delineation and image navigation. Training in the use of ITK-SNAP software was provided using online tutorials and the expertise of one of the project supervisors (AR). Manual segmentation was undertaken between January 2023 and July 2023, by 2 independent members of the clinical research team (a higher surgical trainee in vascular surgery and an intercalating BSc student).

The CLAIM guidelines (Checklist for Artificial Intelligence in Medical Imaging) were used to ensure a clear, transparent and reproducible communication framework for the artificial intelligence aspect of this study, as has been recommended in the literature (60).

Study outcomes measures

The primary and secondary outcome measures of the SPAR study are as listed below:

Primary outcome measure - Correlation between skeletal muscle index (SMI) at L3 and 30-day composite outcome.

30-day composite outcome is defined as a combination of death (mortality) and defined post-operative complications (cardiac, respiratory, haemorrhage, visceral ischaemia, limb ischaemia, renal failure, stroke, delirium / confusion and wound dehiscence / infection) as per the NVR.

Secondary outcome measures –

Correlation between skeletal muscle index (SMI) and:

1. Hospital length of stay (days)
2. Length of stay in critical care (days)
3. In-hospital post-operative mortality
4. Re-admission within 30 days
5. All-cause & aneurysm related mortality

The role of skeletal muscle 'quality' measured as skeletal muscle attenuation or density (in Hounsfield units) will also be compared as a prognostic factor in outcomes analysis (in a similar manner to SMI), as this measure is also produced in all manual or automated segmentation

outputs, to assess if this provides any correlation to outcomes, and if so, whether this correlation is stronger than SMI or not.

Inclusion and exclusion criteria

1. All patients aged 50 years and over, with a diagnosis of AAA (symptomatic or asymptomatic) at the time of the pre-operative CT scan and subsequently underwent infra-renal endovascular aneurysm repair (EVAR), complex endovascular aneurysm repair or open aneurysm repair, whether undertaken electively or emergently, were eligible to be included in this study.
2. The pre-operative CT scan of the abdomen / aorta must include the L3 vertebral body, which serves as the reference level for manual and automated muscle segmentation.
3. The CT scan should also be undertaken using the intravenous administration of a suitable CT contrast agent.

Exclusion criteria included:

- Patients with no available electronic patient records within ICHT
- Patients with no suitable CT scan with contrast
- Patients with a 'corrupted' CT scan (and therefore cannot be used for automated or manual segmentation)
- Patients with a confirmed diagnosis of connective tissue disease (e.g. Marfan syndrome, Ehlers-Danlos syndrome) – This cohort of patients generally tend to have surgical intervention at a younger age due to non-atherosclerotic aneurysmal disease usually with a genetic component, and their underlying aortopathy is usually not associated with frailty / sarcopenia.

The inclusion of only patients without confirmed connective tissue disease ensured relative homogeneity in the study cohort.

Statistical analysis

As this is a feasibility study, no formal power and sample size calculation was completed. However, a pragmatic study to assess and train the deep-learning sarcopenia segmentation tool requires a significant case load of approximately 500 patients in order to identify failure cases for purposes of re-training the existing model for future work.

Descriptive statistics were used for computation of basic values for descriptive purposes, such as mean, variance, dispersion, skewness, kurtosis.

Computation of appropriate correlation coefficients (Pearson's r, Spearman's Rho) will be used to discern association of sarcopenia (skeletal muscle index and attenuation) with age, gender, length of stay, location of discharge (home or other), blood results (Haemoglobin / Albumin protein / CRP / Creatinine / Glucose / Cholesterol), osteoporosis score and body mass index (BMI), and also to discern association of thigh muscle bulk (skeletal muscle index and attenuation) with the same parameters.

The outputs of the AI algorithm were evaluated for accuracy against ground truth using the DICE similarity co-efficient. Failure cases, either with poor DICE co-efficient or output generated will be reviewed, corrected and used for further training of the AI algorithm.

Inter-rater and intra-rater reliability of manually segmented CT scans were assessed using Bland-Altman plots.

As the primary outcome measure for this study looked at 30-day composite post-operative outcomes defined as a combination of death (mortality) and NVR defined post-operative complications (cardiac, respiratory, haemorrhage, visceral ischaemia, limb ischaemia, renal

failure, stroke, delirium / confusion and wound dehiscence / infection), we needed a strategy for grading the severity of various complications to produce an overall score that would reflect the post-operative outcome and complication severity for each patient post AAA repair. This score could then be compared against the sarcopenia segmentation tool's outputs such as skeletal muscle index (SMI – derived from the L3 skeletal muscle area) and skeletal muscle density (in Hounsfield units) to assess for any underlying correlation of statistical significance.

One of the earliest severity grading systems for post-operative complications was the T92 system (61), presented initially in 1992 with a subsequent modification by Dindo et al (62) in 2005. This amended system by Dindo et al described the severity of post-operative complications in greater detail, by accounting for the need for a general anaesthetic and a return to the operating theatre to manage the complication, or the need for HDU / ITU admission for single or multiple organ support as a result of the post-operative complication. A further modification to the T92 system incorporating the modifications suggested by Dindo et al resulted in the development of the "Accordion" Severity Grading System in 2009, which is the severity grading system we have chosen to use for our study (63).

The expanded "Accordion" Severity Grading System (63) classifies post-operative complications into 6 categories on the basis of complication severity, which are listed below:

1. Mild complication – Management requires only minor invasive procedures that can be done at the bedside such as insertion of intravenous lines, urinary catheters, nasogastric tubes, and bedside / superficial drainage of wound infections. Management of the complication may include physiotherapy. The following drugs may be used in the

management of this complication - Antiemetics, antipyretics, analgesics, diuretics, and / or electrolytes.

2. Moderate complication – Management requires pharmacological treatment with drugs other than those allowed for minor complications, such as the use of antibiotics. Administration of blood (and it's associated products) and total parenteral nutrition may also be a part of the management of this post-operative complication severity.
3. Severe complication (Level I) – Management requires an invasive procedure without general anaesthesia. This may be endoscopic, interventional or re-operation, provided no general anaesthesia is used (for example, wound re-exploration under conscious sedation and / or local anaesthesia).
4. Severe complication (Level II) – Management requires an invasive procedure or intervention under general anaesthesia.
5. Severe complication (Level III) – A complication resulting in organ system failure, usually requiring further management in a higher acuity setting, such as Level 2 (HDU) or Level 3 (ITU) care.
6. Mortality – Post-operative death

Using the above Accordion severity grading system, the post-operative complications (as per the NVR) and mortality were allocated a score, the sum total of which would form a composite 'Accordion' score that would provide a quantitative measure of post-operative outcomes for each individual patient.

The NVR post-operative complications and mortality were given a score according to the table below, using the "Accordion" Severity Grading System.

Type of Complication	Accordion Score
Cardiac	3
Respiratory	2
Haemorrhage	4
Visceral Ischaemia	4
Limb Ischaemia	4
Renal Failure	1
Stroke	2
Delirium / Confusion	1
Wound Dehiscence / Infection	1
Death	6

Figure 16. Table demonstrating Accordion scoring system to grade the severity of post-operative complications of patients undergoing AAA repair

Using this Accordion scoring system, patients undergoing AAA repair were given an Accordion score ranging from 0 – 28, depending on their individual post-operative course and complications suffered (if any).

Whilst there may be some variability in the management of these post-operative complications which may differ from patient-to-patient depending on clinical need and clinician preference,

the Accordion scores were allocated on the basis of how these post-operative complications would be broadly managed in day-to-day clinical practice.

Study Approvals

Ethics approval

The SPAR study obtained ethics approval from the Health Regulatory Authority (HRA) and Health and Care Research Wales (HCRW) in May 2021. The IRAS project ID is 285472 and it's associated protocol number is 21SM6746. As this ethics approval was for an initial period of 2 years, an extension was requested and applied for in April 2023 with no change in study protocol to enable completion of data collection and analysis. This extension was granted, with ethics extension now in place until 31/10/2024. This study ethics extension did not have any implications on already existing resources or funding, nor did it require further regulatory review.

The study also received confirmation of capacity and capability from ICHT before any research activity commenced. Re-confirmation of capacity and capability from ICHT was also obtained and received following confirmation of ethics extension in April 2023.

The overall study coordination centre was Imperial College London.

The study was conducted in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th World Medical Assembly, Helsinki (1964) and later revisions.

In accordance with the UK Policy Framework for health and social care, research projects carried out in the Trust (ICHT) may be randomly chosen by the Joint Research Compliance Office (JRCO) for auditing.

Consent

Specific patient consent was not obtained. The SPAR study is a retrospective study in which patients will only be identified by members of the direct clinical care and / or clinical research team. Patient identifiable data will not be accessible outside these clinical or research teams. Pseudonymised imaging data was used by the clinical and research team members for inputting imaging data into the sarcopenia segmentation tool, held securely at the Hammersmith Hospital, London (part of ICHT) and Comprehensive Cancer Imaging Centre at Imperial College London. Outputs were also produced and held in the same, secure research network database and shared with members of the clinical research team only. Manual imaging segmentation using ITK-SNAP was also undertaken using pseudonymised imaging on secure ICHT IT systems.

Confidentiality

The study investigators will preserve the confidentiality of patient information as per the Data Protection Act (2018).

Patient data was securely collected from NHS password and smartcard protected electronic patient records, pseudonymised by assigning a unique patient enrolment number to patient records and entered onto the REDCap database (hosted by Imperial College London) by the clinical research team. The pseudonymised clinical data was analysed in password protected Imperial College or ICHT computers.

The look-up table, which links the patient identifier to the unique patient enrolment number will be stored behind the clinical firewall in a NHS password protected computer.

Indemnity and Sponsorship

Imperial College London holds negligent harm and non-negligent harm insurance policies which apply to this study.

Imperial College London will act as the main sponsor for this study. Delegated responsibilities will be assigned to the NHS trusts (ICHT) taking part in this study.

Funding

In view of the retrospective nature of this study, no significant funding requirements were needed for undertaking this study.

Internal funding was provided by the 2 research co-supervisors and chief investigators of this study (Professor Alun H Davies and Professor Andrea Rockall) to facilitate pseudonymisation of CT scans included in this study and utilisation of the automated sarcopenia segmentation tool as this required the input and assistance of the imaging research team.

Adverse events and follow-up

In this retrospective study, we do not expect any adverse events in patients. Furthermore, no significant issues are envisaged. If any additional, unexpected results are discovered during imaging review, which may have an impact on patient management or outcomes, the responsible clinical team will be informed and an addendum attached to the relevant CT imaging report and / or electronic patient records.

Follow-up visits are not applicable in this retrospective study collecting data on patients who have previously undergone an intervention for AAA repair, as it utilises CT images that were undertaken as part of the pre- and post-operative assessment and management of patients undergoing AAA repair. Patients are not individually contacted at any point of this study, as only their existing clinical metadata and imaging are utilised for data collection and analysis, and no additional information, assessments or intervention would be required of them.
Data and all appropriate documentation will be stored for a minimum of 10 years after the completion of this study.

Results

528 patients were identified from the ICHT NVR database records as having undergone AAA repair, or having been coded for the same between January 1st 2016 and March 31st 2020, and therefore this group of patients was selected as our study cohort to test the sarcopenia segmentation tool.

However, only 334 patients met inclusion / exclusion criteria for the SPAR study and were included in the CT scan pseudonymisation process, testing of the sarcopenia segmentation tool and subsequent statistical analysis. The reasons for inclusion / exclusion are given in the patient flowchart below.

Clinical / demographic data for these 334 patients was collected and stored on the REDCap database, and CT scan pseudonymisation requests sent to the imaging department with the relevant CT scan dates. Successfully pseudonymised CT scans were received for 280 patients and these were inputted through the sarcopenia segmentation tool, with successful outputs

received for all 280 patients. The imaging research department were unable to produce pseudonymised scans for the other 54 patients as the relevant, associated CT scans were done in a non-ICHT institution and then transferred over to the ICHT PACS via IEP causing some technical issues in the pseudonymisation process. Due to timeline constraints, an attempt was not made to find an alternative method to pseudonymise these scans in conjunction with the imaging research department.

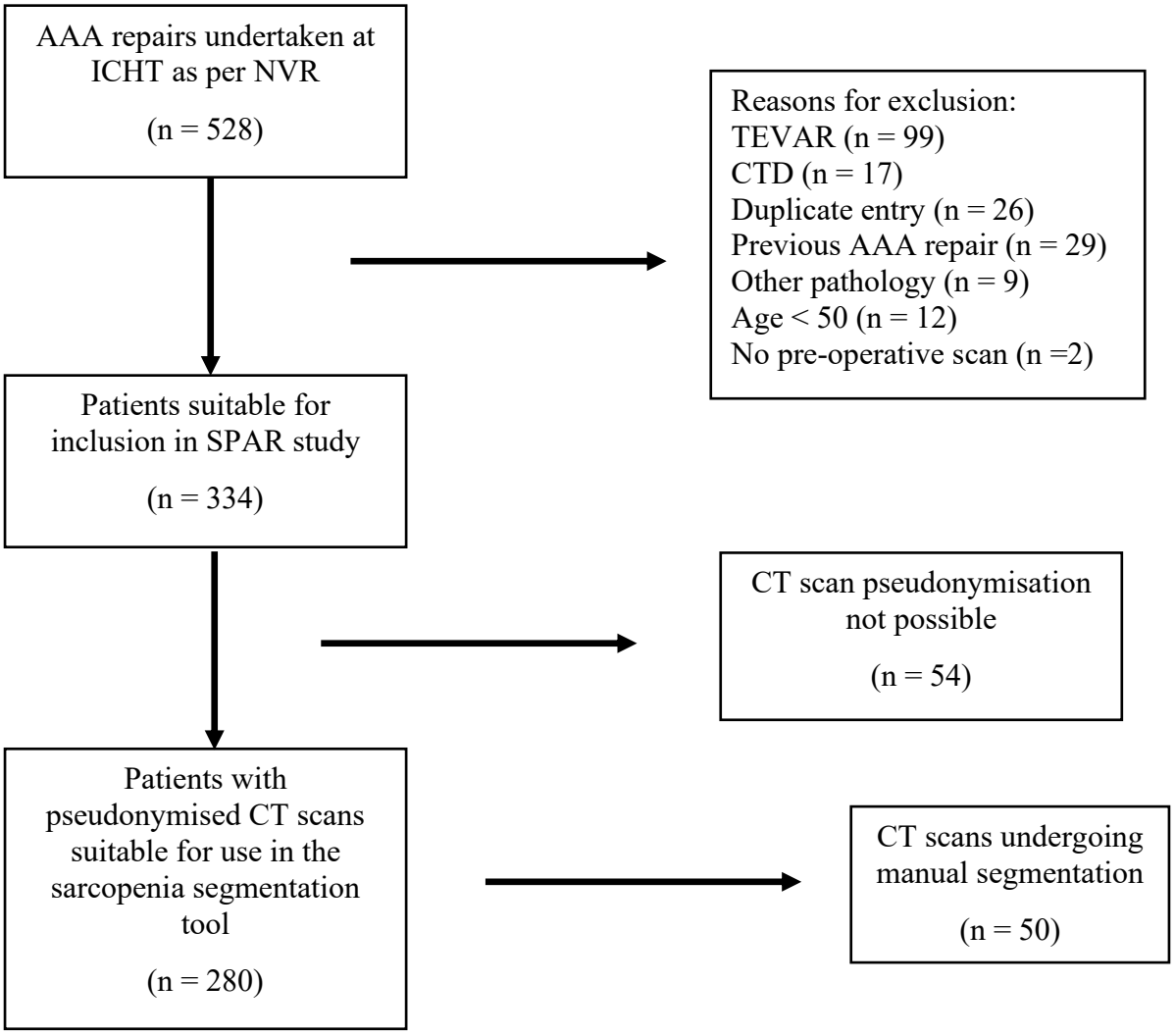


Figure 17. Flow diagram demonstrating study patient cohort and reasons for inclusion / exclusion, and final sample size for testing the sarcopenia segmentation tool, and associated outcomes analysis. (CTD = Connective Tissue Disease)

A significant number of patients (n = 99) from the initial study cohort were found to have undergone TEVAR for a variety of thoracic aortic pathologies, such as type B aortic dissection (TBAD), traumatic aortic transection or thoracic aneurysmal disease and had been included in the ICHT NVR database as patients who had undergone 'AAA repair'. This resulted in nearly 1 in 5 patients from the original study cohort having to be excluded from the SPAR study as these patients did not have a treated abdominal aortic aneurysm (AAA) and, therefore, did not meet inclusion criteria, resulting in significant study attrition.

Other significant reasons for study exclusion included duplicate patient entries (n = 26) and patients having initial AAA repair elsewhere or prior to the study period (n = 29). Duplicate patient entries onto the NVR database usually occurred when patients had re-intervention, particularly in the case of staged procedures in complex endovascular repair of thoracoabdominal aneurysms or required proximal extension with a fenestrated cuff or similar for ongoing type 1A endoleak (ongoing high pressure flow in the aortic aneurysm sac due to inadequate seal at the proximal aspect of the endograft and inadequate apposition to the neck of the aortic aneurysm). Re-intervention for patients who had their index AAA repair elsewhere or outside the study window happened for similar reasons, or in the case of previous open AAA repair were usually due to progressive dilatation / aneurysmal degeneration of the visceral segment necessitating re-intervention.

Study Demographics

The key study demographic data of the entire patient cohort (n = 334) are as listed in the table below.

Sex	Male (%)	275 (82.33)

	Female (%)	59 (17.66)
Type of AAA repair	Open infrarenal repair (%) Open juxtarenal repair (%) Infrarenal EVAR (%) Complex EVAR (%) TAAA Repair (%)	71 (21.26) 35 (10.48) 107 (32.03) 73 (21.86) 48 (14.37)
Mean Age (IQR)		74.2 (68 – 80)
Mean Height (IQR) - cm		172.5 (167 – 179)
Mean Weight (IQR) - kgs		79.57 (68.76 – 90)
Mean BMI (IQR)		26.67 (23.53 – 29.84)
Co-morbidities (%)	Diabetes Hypertension Coronary Artery Disease Respiratory Cerebrovascular Chronic Kidney Disease	48 (14.37) 236 (70.66) 117 (35.03) 108 (32.33) 33 (9.88) 56 (16.77)
Smoking status (%)	Never smoker Ex-smoker Current smoker	38 (11.38) 187 (55.99) 81 (24.25)

	Not known	28 (8.38)
Medications (%)	Antiplatelet/s Anticoagulant Antihypertensive/s Statin	197 (58.98) 46 (13.77) 225 (67.36) 247 (73.73)

Figure 18. SPAR study demographics table

Manual Segmentation

50 pre-operative CT scans (17.86% of pseudonymised pre-operative CT scans) underwent manual segmentation using the ITK-SNAP software by 2 independent members of the clinical research team to provide 'ground truth' for evaluation of the accuracy of the sarcopenia segmentation tool.

Each, individual pseudonymised pre-operative CT scan usually had multiple series of scans of varying thickness (usually one series of thin 1 mm axial slices and one series of thick 2 – 3 mm axial slices) along with varying contrast phases (pre-contrast, arterial or portal venous phases as per local protocols) associated with it. To ensure uniformity and standardisation of study outputs and results, the arterial phase CT scan was chosen for both manual and automated segmentation outputs. Whilst skeletal muscle area is not affected by the contrast phase, significant differences in skeletal muscle attenuation (in Hounsfield units – a marker of skeletal muscle density) are noted on the basis of contrast phase, due to varying degrees of CT contrast

uptake by the skeletal muscle, particularly when comparing non-contrast phase images to arterial phase images.

Manual segmentation outputs produced from ITK-SNAP by the 2 independent clinical researchers then underwent 'thresholding' by the imaging research team to produce a final output to serve as 'ground truth' and to be used for comparison purposes against the outputs of the automated sarcopenia segmentation tool. Thresholding is done to exclude macroscopic fat within the segmented muscle to obtain a more accurate estimate of skeletal muscle area and mean attenuation (in Hounsfield units). The same thresholds ([-29 HU, 150 HU]) were applied as those used by the automated sarcopenia segmentation tool to standardise the process and outputs (55).

ITK-SNAP also produces skeletal muscle volumes (in mm^3) on segmentation, and this was converted to skeletal muscle area (SMA) by dividing the skeletal muscle volume by the associated CT slice thickness (in mm).

The demographics of this patient cohort (n=50) are as listed in the table below.

Sex	Male (%) Female (%)	44 (88) 6 (12)
Type of AAA repair	Open infrarenal repair (%) Open juxtarenal repair (%) Infrarenal EVAR (%) Complex EVAR (%)	13 (26) 5 (10) 10 (20) 17 (34)

	TAAA Repair (%)	5 (10)
Mean Age (IQR)		74.76 (69 – 81.75)
Mean Height (IQR) - cm		173.64 (168 – 178)
Mean Weight (IQR) - kgs		76.97 (66 – 83)
Mean BMI (IQR)		25.48 (22.52 – 27.58)
Co-morbidities (%)	Diabetes Hypertension Coronary Artery Disease Respiratory Cerebrovascular Chronic Kidney Disease	7 (14) 37 (72) 16 (32) 16 (32) 4 (8) 7 (14)
Smoking status (%)	Never smoker Ex-smoker Current smoker Not known	5 (10) 26 (52) 13 (26) 6 (12)
Medications (%)	Antiplatelet/s Anticoagulant Antihypertensive/s Statin	36 (72) 4 (8) 34 (68) 44 (88)

Figure 19. Study demographics table of patients undergoing manual segmentation of pseudonymised CT scans for 'ground truth'

Automated Segmentation

Automated segmentation outputs were produced for all 280 pseudonymised pre-operative CT scans in this study cohort by the sarcopenia segmentation tool (including the 50 pseudonymised CT scans that underwent manual segmentation), rendering a 100% output success for the use of this sarcopenia segmentation tool.

Each automated segmentation report by the sarcopenia segmentation tool included the following information – Unique trial ID number, sex, scan date / time, scan description (contrast phase and slice thickness), patient age, L3 slice detection model confidence (in percentage), muscle estimation area (of skeletal muscle at the model selected axial slice at the L3 level) and mean muscle attenuation (in Hounsfield units). The deep-learning algorithm of the automated sarcopenia segmentation tool has been trained to identify the middle of the 3rd lumbar vertebral body (L3) and utilise the axial slice at this level for the purposes of muscle segmentation to provide skeletal muscle area and mean muscle attenuation at this level. The mean skeletal muscle attenuation (in Hounsfield units) is a measure of skeletal muscle density, which serves as a surrogate marker of skeletal muscle quality as has been previously described.

4 images were also provided as part of every automated segmentation report. These included coronal and sagittal images of the CT scan with a horizontal line drawn through the L3 vertebral body demonstrating the level at which automated segmentation was undertaken. This was accompanied by 2 identical axial slices at this level, one without any muscle segmentation and one with skeletal muscle segmentation highlighted in purple.

A sample automated segmentation output is provided in the figure below.

Figure 20. Sample automated segmentation output produced by the sarcopenia segmentation tool demonstrating patient and scan identifiers, along with automated muscle segmentation (in purple) and relevant outputs

Multiple outputs were usually produced for each pseudonymised CT scan for different contrast phases and series. The arterial contrast phase output with the highest L3 slice detection confidence percentage was chosen for each patient in the study cohort when there were multiple outputs produced.

Manual and automated segmentation outputs for each patient included in the SPAR study were collected and stored on the REDCap database for the purposes of outcomes related statistical analysis.

Correlation between Manual and Automated Segmentation

Inter-rater reliability (IRR) between the 2 independent clinical researchers undertaking manual segmentation was measured using intraclass correlation coefficient (ICC) as the data to be compared was continuous quantitative data. One of the clinical researchers (HS) did not undertake manual segmentations on 3 pseudonymised CT scans as she felt the quality of the CT scan was inadequate or there were no clear margins on skeletal muscle at the L3 level due to AAA rupture and associated retroperitoneal haematoma obscuring these margins.

Excellent IRR was noted between the manual segmentations (n = 47) undertaken by the 2 researchers with an ICC 0.96 (95% CI 0.95 – 0.97, $p < 0.05$) for skeletal muscle area (SMA) at the L3 level. A scatterplot demonstrating this is shown in the figure below.

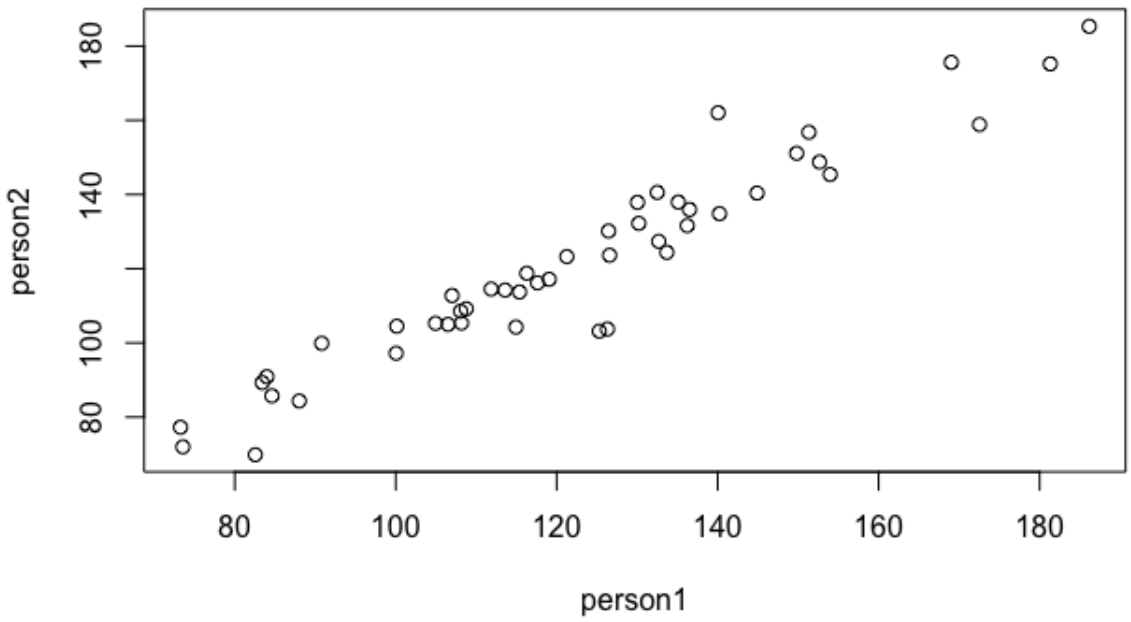


Figure 21. Scatterplot demonstrating excellent IRR between manual segmentations undertaken by 2 independent clinical researchers (ICC 0.96, 95% CI 0.95 – 0.97, $p < 0.05$) for skeletal muscle area (SMA)

Excellent IRR was noted between the manual segmentations (n = 47) undertaken by the 2 researchers with an ICC 0.98 (95% CI 0.98 – 0.99, $p < 0.05$) for mean skeletal muscle attenuation (in Hounsfield units) at the L3 level. A scatterplot demonstrating this is shown in the figure below.

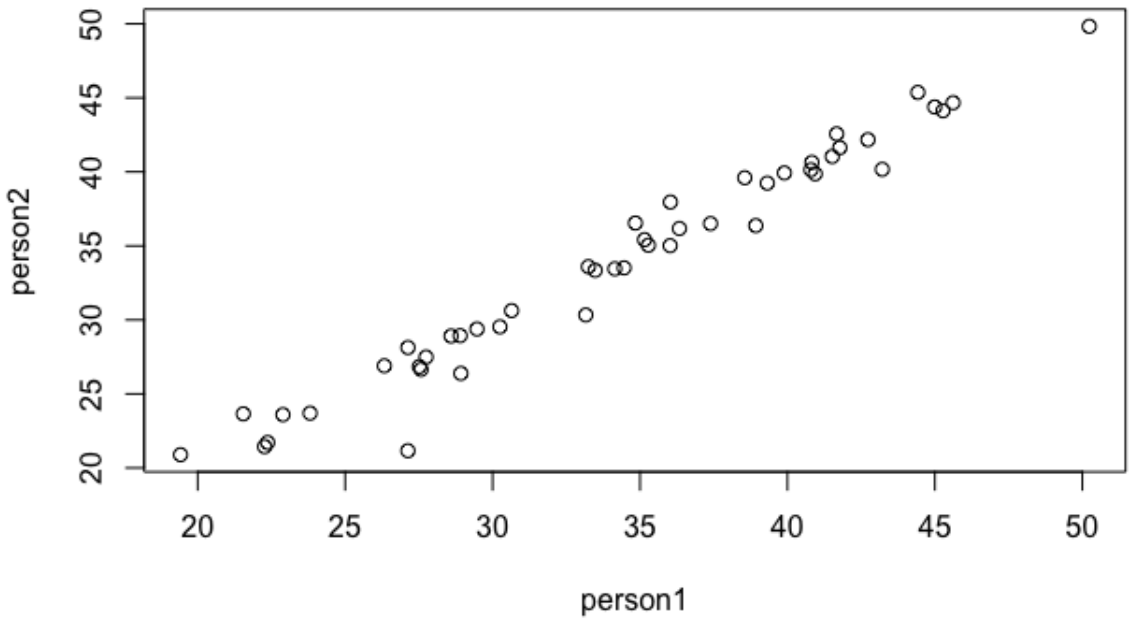


Figure 22. Scatterplot demonstrating excellent IRR between manual segmentations undertaken by 2 independent clinical researchers (ICC 0.98, 95% CI 0.98 – 0.99, $p < 0.05$) for skeletal muscle attenuation (in Hounsfield units)

Inter-rater reliability (IRR) was then tested between the automated sarcopenia segmentation tool and the 2 independent clinical researchers undertaking manual segmentation using intraclass correlation coefficient (ICC). One automated output was deemed unsatisfactory and not included in the IRR calculations.

Excellent IRR was noted between the manual (utilising segmentations from both clinical researchers) and automated segmentations ($n = 46$) with an ICC 0.93 (95% CI 0.91 – 0.94, $p < 0.05$) for skeletal muscle area (SMA) at the L3 level. Scatterplots demonstrating this are shown in the figure below.

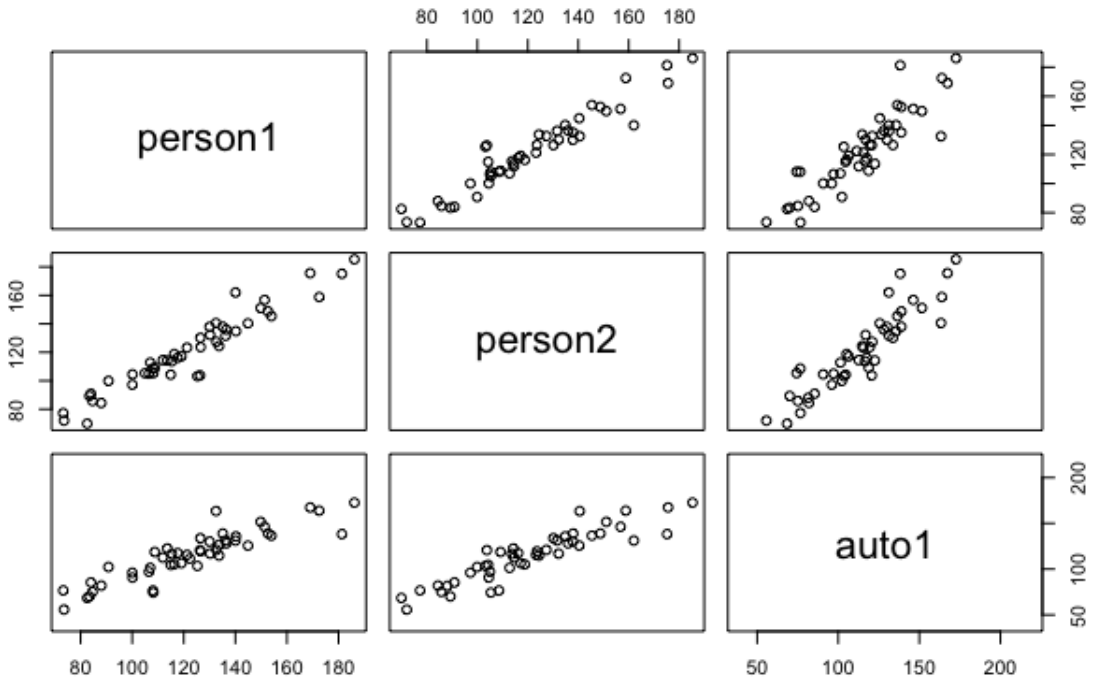


Figure 23. Scatterplots demonstrating excellent IRR between manual and automated segmentations (ICC 0.93, 95% CI 0.91 – 0.94, $p < 0.05$) for skeletal muscle area (SMA)

Excellent IRR was noted between the manual (utilising segmentations from both clinical researchers) and automated segmentations (n = 46) with an ICC 0.90 (95% CI 0.88 – 0.92, $p < 0.05$) for mean skeletal muscle attenuation (in Hounsfield units) at the L3 level. Scatterplots demonstrating this are shown in the figure below.

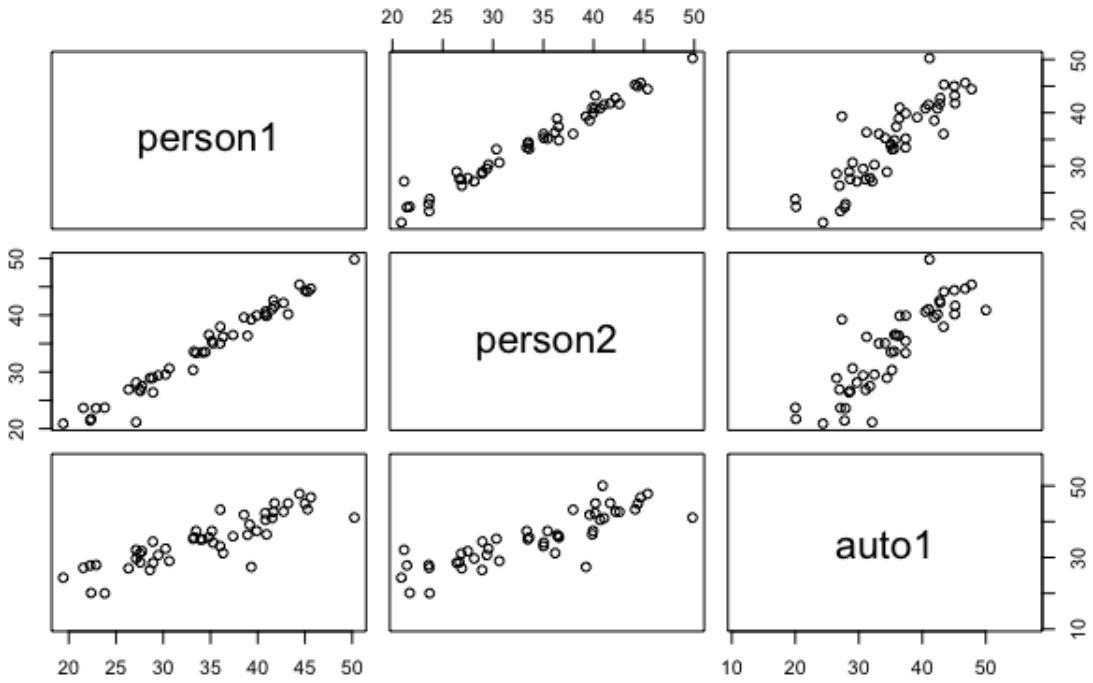


Figure 24. Scatterplots demonstrating excellent IRR between manual and automated segmentations (ICC 0.90, 95% CI 0.88 – 0.92, $p < 0.05$) for skeletal muscle attenuation (in Hounsfield units)

Following calculation of skeletal muscle index (SMI) from the skeletal muscle area (SMA) and patient's height using the relevant numerical formula ($SMI = SMA / Height^2$), normality testing was undertaken using the Shapiro-Wilk normality test. A histogram depicting the SMI values of the entire patient cohort is demonstrated below.

Shapiro-Wilk normality testing confirmed the SMI values to be normally distributed ($W = 0.99143$, $p = 0.1239$). p-value > 0.05 implies that distribution of the SMI data is not significantly different from normal distribution allowing us to assume normality.

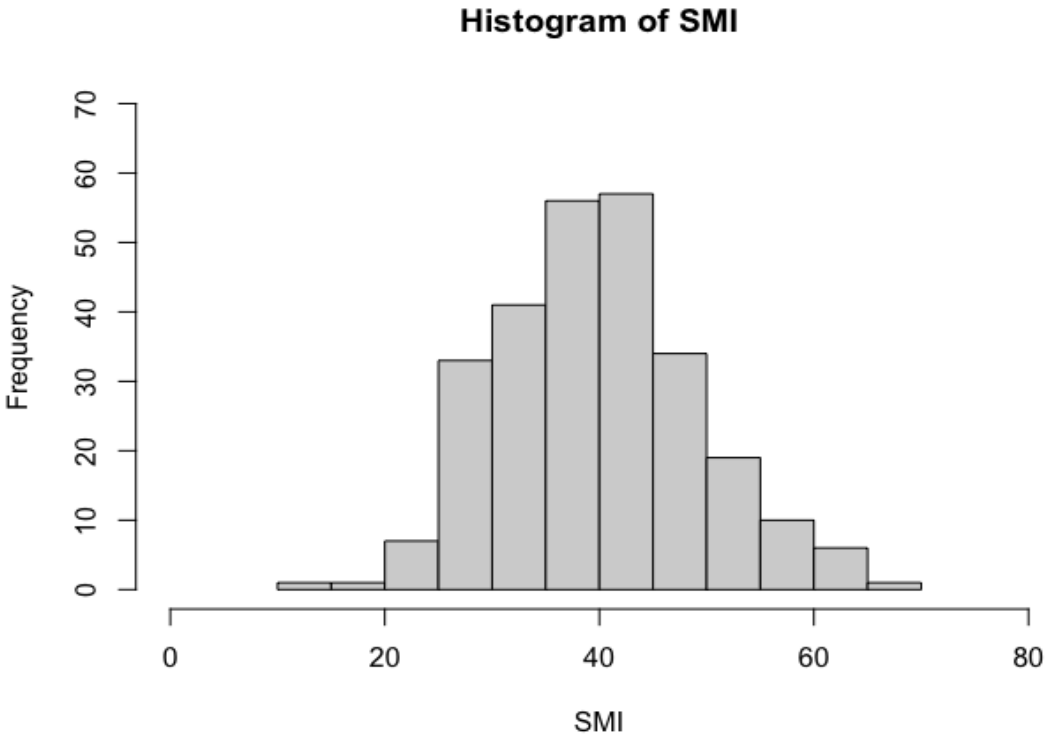


Figure 25. Histogram demonstrating SMI values of the entire SPAR patient cohort

Similar normality testing was undertaken for skeletal muscle density (SMD) utilising the values for mean skeletal muscle attenuation (in Hounsfield units) for the entire patient cohort. A histogram depicting these values is demonstrated below.

Shapiro-Wilk normality testing confirmed the SMD values to be normally distributed ($W = 0.99584$, $p = 0.6646$). p-value > 0.05 implies that distribution of the SMD data is not significantly different from normal distribution allowing us to assume normality.

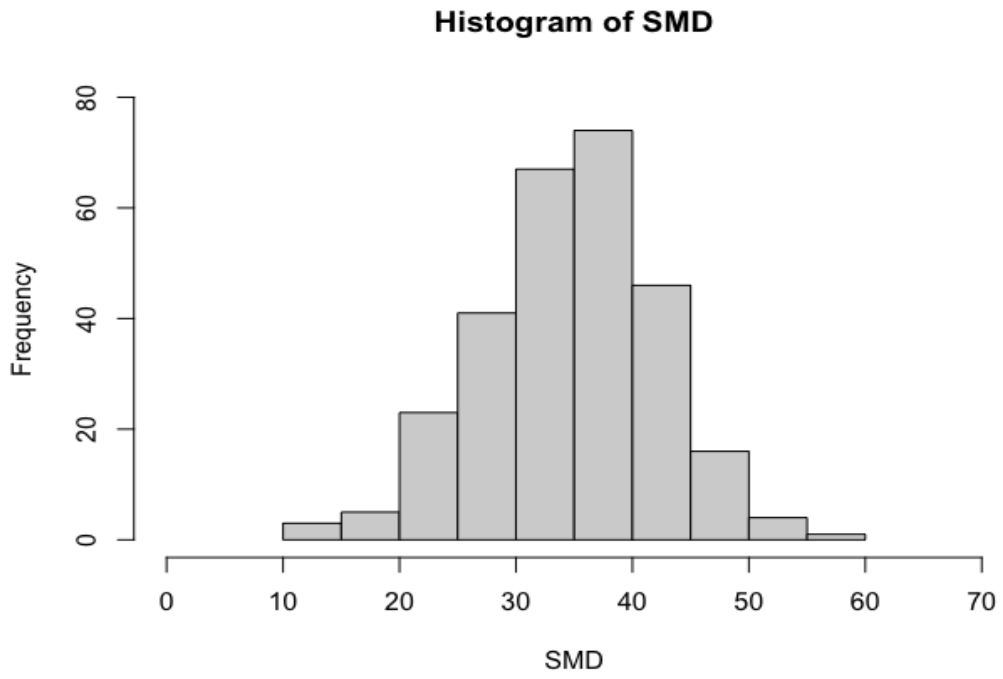


Figure 26. Histogram demonstrating SMD values (in Hounsfield units) of the entire SPAR patient cohort

Outcomes Analysis

Skeletal muscle mass varies significantly on the basis of sex, with men having significantly higher muscle mass on average as compared to women. This was also reflected in our patient cohort with men having higher SMI scores (median male SMI = 40.8, IQR = 35.26 – 46.85) as compared to women (median female SMI = 30.88, IQR = 27.05 – 35.71).

As has been alluded to in chapter 1, the risk profiles for open surgery versus endovascular surgery are also significantly different, with open surgery carrying a higher risk of 30-day morbidity and mortality as compared to endovascular intervention, and therefore usually reserved for younger patients with better cardiorespiratory fitness, reduced risk profiles, lesser

co-morbidities and / or hostile aortic anatomy rendering it unsuitable for endovascular intervention. The risk profiles for ruptured aortic aneurysm repair are also significantly different to elective intervention, with ruptured aneurysms carrying much higher risk of mortality / morbidity, regardless of age, co-morbidities or type of intervention (open vs endovascular repair).

On the basis of these factors, it was decided to undertake outcomes analysis separately for men and women. For similar reasons, outcomes for patients undergoing unruptured aortic aneurysm repair were also analysed separately from ruptured aortic aneurysm repairs.

Furthermore, patients were then grouped into 3 types of repairs for the purposes of statistical analysis to compare post-operative outcomes on the basis of varying risk profiles –

1) Open surgical repair group (included open infrarenal and juxtarenal AAA repairs)
2) Endovascular AAA repair group (included standard and complex EVAR)
3) Thoracoabdominal aortic aneurysm repair group

The Accordion score served as a surrogate measure of 30-day composite outcomes (death and defined complications as per the NVR).

As demonstrated in the systematic review in chapter 2, most individual studies that investigated the use of SMI to compare post-operative outcomes divided their patient cohorts into sub-groups of quartiles or tertiles to undertake statistical outcomes analysis. We undertook the same and divided each intervention sub-group / cohort into quartiles to compare 30-day composite outcomes, using Accordion scores, against skeletal muscle index (SMI) and skeletal muscle density (SMD – using mean skeletal muscle attenuation).

A) Open surgical repair for abdominal aortic aneurysm (AAA) – Males

93 men underwent open surgical repair for aortic aneurysmal disease in our patient cohort, with 84 men having a pseudonymised scan that underwent automated segmentation. The mean Accordion score was 2.53 (SD = 3.89), and the mean SMI was 43.17 (SD = 9.41).

There was no correlation noted between Accordion score and SMI on Pearson's correlation testing ($r = -0.132$) or regression analysis ($R^2 = 0.017$, $p = 0.239$).

21 deaths were noted in this group during the follow-up period with mean time to mortality being 36.85 months (SD = 29.07 months). Moderate negative correlation was noted between the Accordion score and overall mortality on Pearson's correlation testing ($r = -0.768$, $p < 0.001$) and regression analysis ($R^2 = 0.59$, $p < 0.001$). This may be expected as death / mortality is a contributory factor to the Accordion score, and adds a score of 6 to the overall score.

Mean skeletal muscle attenuation / density (SMD) in this cohort was 36.79 Hounsfield units (SD = 7.28 HU). There was no correlation noted between Accordion score and SMD on Pearson's correlation testing ($r = -0.304$) or regression analysis ($R^2 = 0.092$, $p = 0.005$).

Men undergoing open surgical repair for unruptured aneurysms (n=72) were then divided into quartiles using SMI scores. A two-sample t-test was performed to compare Accordion scores between the lowest SMI quartile and highest SMI quartile. There was no significant difference in Accordion scores between the lowest SMI quartile (Mean = 1.61, SD = 2.40) and the highest SMI quartile (Mean = 1.33, SD = 2.06); $t(34) = 0.372$, $p = 0.712$.

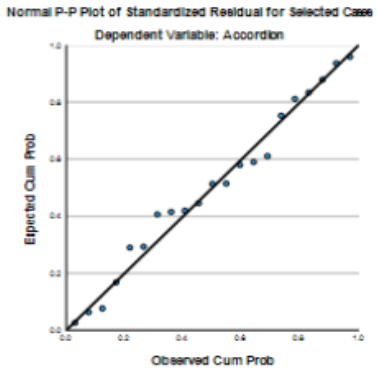
Men undergoing open surgical repair for unruptured aneurysms (n=72) were also divided into quartiles using SMD scores (mean skeletal muscle attenuation). A two-sample t-test was performed to compare Accordion scores between the lowest SMD quartile and highest SMD quartile. There was significant difference in Accordion scores between the lowest SMD quartile

(Mean = 3.17, SD = 3.63) and the highest SMD quartile (Mean = 0.67, SD = 1.33); t(34) = 2.741, p = 0.01.

Number of patients	93
Mean Accordion Score	2.53 (SD = 3.89)
Mean SMI	43.17 (SD = 9.41)
Deaths in F/U period	21
Mean time to mortality	36.85 months (SD = 29.07 months)
Mean SMD	36.79 HU (SD = 7.28 HU)

	Lowest SMI Quartile	Highest SMI Quartile
Mean Accordion Score (SD)	1.61 (2.4)	1.33 (2.06)

	Lowest SMD Quartile	**Highest SMD Quartile**
Mean Accordion Score (SD)	**3.17 (3.63)**	**0.67 (1.33)**
t(34) = 2.741, p = 0.01		



Regression modelling for Accordion score and overall mortality

Moderate negative correlation noted ($r = -0.768$ / $R^2 = 0.59$, $p < 0.001$)

Figure 27. Summary of outcomes analysis for men undergoing open AAA repair

B) Open surgical repair for abdominal aortic aneurysm (AAA) – Females

Only 13 women underwent open surgical repair for aortic aneurysmal disease in our patient cohort, limiting any meaningful statistical analysis, with only 9 women having a pseudonymised scan that underwent automated segmentation.

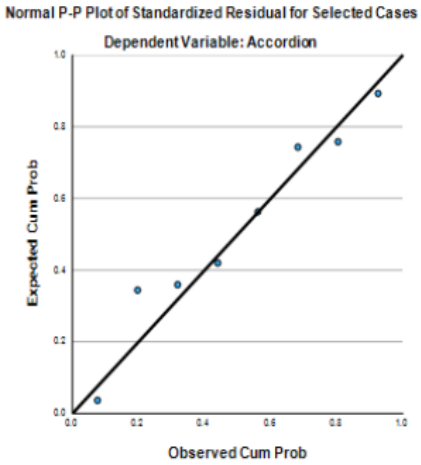
The mean Accordion score was 7.15 (SD = 6.07), and the mean SMI was 32.27 (SD = 6.72) in this cohort.

Good negative correlation was noted between Accordion score and SMI on Pearson's correlation testing ($r = -0.86$, $p = 0.003$). Good correlation was noted between Accordion score and SMI on regression analysis ($R^2 = 0.75$, $p = 0.006$).

9 deaths were noted in this group during the follow-up period with mean time to mortality being 9 months (SD = 18.61 months). Strong negative correlation was noted between the Accordion score and overall mortality on Pearson's correlation testing ($r = -0.90$, $p < 0.001$) and regression analysis ($R^2 = 0.81$, $p < 0.001$).

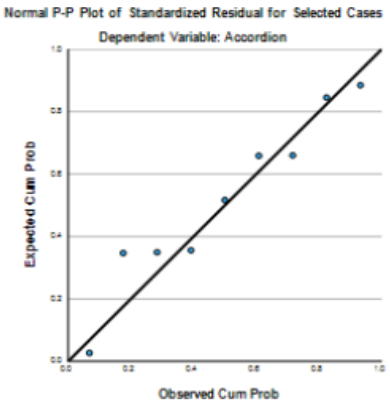
Mean skeletal muscle attenuation / density (SMD) in this cohort was 33.67 Hounsfield units (SD = 4.91 HU). There was no correlation noted between Accordion score and SMD on Pearson's correlation testing ($r = 0.097$, $p = 0.402$) or regression analysis ($R^2 = 0.009$, $p = 0.805$).

Number of patients	13
Mean Accordion Score	7.15 (SD = 6.07)
Mean SMI	32.27 (SD = 6.72)
Deaths in F/U period	9
Mean time to mortality	9 months (SD = 18.61 months)
Mean SMD	33.67 HU (SD = 4.91 HU)



Regression modelling for Accordion score and SMI

Good negative correlation noted
($r = -0.86$ / $R^2 = 0.75$, $p < 0.05$)



Regression modelling for Accordion score and overall mortality

Moderate negative correlation noted
($r = -0.9$ / $R^2 = 0.81$, $p < 0.001$)

Figure 28. Summary of outcomes analysis for women undergoing open AAA repair

C) Endovascular abdominal aortic aneurysm (AAA) repair – Males

147 men underwent endovascular abdominal aortic aneurysm repair in our patient cohort, with 127 men having a pseudonymised scan that underwent automated segmentation. The mean Accordion score was 1.36 (SD = 2.54), and the mean SMI was 39.95 (SD = 8.47) in this cohort.

There was no correlation noted between Accordion score and SMI on Pearson's correlation testing (r = -0.01) or regression analysis (R^2 = -0.008, p = 0.911).

54 deaths were noted in this group during the follow-up period with mean time to mortality being 34.82 months (SD = 23.45 months). Moderate negative correlation was noted between the Accordion score and overall mortality on Pearson's correlation testing (r = -0.409, p = 0.001), but no correlation was noted on regression analysis (R^2 = 0.17, p = 0.002).

The mean skeletal muscle attenuation / density (SMD) in this cohort was 33.92 Hounsfield units (SD = 7.18 HU). There was no correlation noted between Accordion score and SMD on Pearson's correlation testing (r = 0.047, p = 0.302) or regression analysis (R^2 = 0.002, p = 0.604).

Men undergoing EVAR for unruptured aneurysms (n=140) were then divided into quartiles using SMI scores. A two-sample t-test was performed to compare Accordion scores between the lowest SMI quartile and highest SMI quartile. There was no significant difference in Accordion scores between the lowest SMI quartile (Mean = 1.48, SD = 2.38) and the highest SMI quartile (Mean = 1.13, SD = 2.85); t(60) = 0.532, p = 0.596.

Men undergoing EVAR for unruptured aneurysms (n=140) were also divided into quartiles using SMD scores (mean skeletal muscle attenuation). A two-sample t-test was performed to compare Accordion scores between the lowest SMD quartile and highest SMD quartile. There was no significant difference in Accordion scores between the lowest SMD quartile (Mean =

1.00, SD = 1.612) and the highest SMD quartile (Mean = 0.84, SD = 1.53); t(60) = 0.404, p = 0.688.

Number of patients	147
Mean Accordion Score	1.36 (SD = 2.54)
Mean SMI	39.95 (SD = 8.47)
Deaths in F/U period	54
Mean time to mortality	34.82 months (SD = 23.45 months)

	Lowest SMI Quartile	Highest SMI Quartile
Mean Accordion Score (SD)	1.48 (2.38)	1.13 (2.85)
t(60) = 0.532, p = 0.596		

	Lowest SMD Quartile	Highest SMD Quartile
Mean Accordion Score (SD)	1 (1.61)	0.84 (1.53)
t(60) = 0.404, p = 0.688		

Figure 29. Summary of outcomes analysis for men undergoing EVAR

D) Endovascular abdominal aortic aneurysm (AAA) repair – Females

33 women underwent endovascular abdominal aortic aneurysm repair in our patient cohort, with 27 women having a pseudonymised scan that underwent automated segmentation. The mean Accordion score was 1.57 (SD = 3.26), and the mean SMI was 29.29 (SD = 6.53).

There was no correlation noted between Accordion score and SMI on Pearson's correlation testing (r = -0.217, p = 0.149) or regression analysis (R^2 = 0.047, p = 0.298).

13 deaths were noted in this group during the follow-up period with mean time to mortality being 33.76 months (SD = 22.64 months). Moderate negative correlation was noted between the Accordion score and overall mortality on Pearson's correlation testing (r = -0.469, p = 0.053), but no correlation was noted on regression analysis (R^2 = 0.22, p = 0.11).

Mean skeletal muscle attenuation / density (SMD) in this cohort was 27.56 Hounsfield units (SD = 7.03 HU). There was no significant correlation noted between Accordion score and SMD on Pearson's correlation testing ($r = 0.302$, $p = 0.063$) or regression analysis ($R^2 = 0.091$, $p = 0.126$).

Women undergoing EVAR for unruptured aneurysms (n=22) were then divided into quartiles using SMI scores. A two-sample t-test was performed to compare Accordion scores between the lowest SMI quartile and highest SMI quartile. Whilst there was a significant difference in Accordion scores between the lowest SMI quartile (Mean = 2.5, SD = 6.12) and the highest SMI quartile (Mean = 0.33, SD = 0.82), this did not achieve statistical significance ($t(10) = 0.859$, $p = 0.205$) probably due to the small numbers.

Women undergoing EVAR for unruptured aneurysms (n=22) were also divided into quartiles using SMD scores (mean skeletal muscle attenuation). A two-sample t-test was performed to compare Accordion scores between the lowest SMD quartile and highest SMD quartile. There was no significant difference in Accordion scores between the lowest SMD quartile (Mean = 0.4, SD = 0.89) and the highest SMD quartile (Mean = 0.6, SD = 1.34); $t(8) = -0.277$, $p = 0.394$.

Number of patients	33
Mean Accordion Score	1.57 (SD = 3.26)
Mean SMI	29.29 (SD = 6.53)
Deaths in F/U period	13
Mean time to mortality	33.76 months (SD = 22.64 months)

	Lowest SMI Quartile	Highest SMI Quartile
Mean Accordion Score (SD)	2.5 (6.12)	0.33 (0.82)
t(10) = 0.859, p = 0.205		

	Lowest SMD Quartile	Highest SMD Quartile
Mean Accordion Score (SD)	0.4 (0.89)	0.6 (1.34)
t(8) = -0.277, p = 0.394		

Figure 30. Summary of outcomes analysis for women undergoing EVAR

E) Thoracoabdominal aortic aneurysm (TAAA) repair

48 patients underwent thoracoabdominal aortic aneurysm repair in our patient cohort, with 33 patients having a pseudonymised scan that underwent automated segmentation. The mean Accordion score was 6.5 (SD = 6.01), and the mean SMI was 40.56 (SD = 7.81).

Amongst men, there was no correlation noted between Accordion score and SMI on Pearson's correlation testing ($r = 0.175$, $p = 0.236$) or regression analysis ($R^2 = 0.031$, $p = 0.473$). There was no correlation was noted between Accordion score and SMD on Pearson's correlation testing ($r = 0.158$, $p = 0.236$) or regression analysis ($R^2 = 0.025$, $p = 0.471$).

9 deaths were noted amongst men during the follow-up period with mean time to mortality being 12.32 months (SD = 21.19 months). There was no correlation noted between the Accordion score and death on Pearson's correlation testing ($r = -0.112$, $p = 0.388$) or regression analysis ($R^2 = 0.012$, $p = 0.775$).

Amongst women, there was no correlation noted between Accordion score and SMI on Pearson's correlation testing ($r = 0.321$, $p = 0.219$) or regression analysis ($R^2 = 0.103$, $p = 0.439$). There was no correlation noted between Accordion score and SMD on Pearson's correlation testing ($r = 0.062$, $p = 0.433$) or regression analysis ($R^2 = 0.004$, $p = 0.865$).

6 deaths were noted amongst women during the follow-up period with mean time to mortality being 5.78 months (SD = 8.59 months). Moderate negative correlation was noted between the Accordion score and death on Pearson's correlation testing ($r = -0.431$, $p = 0.197$), but there was no correlation noted on regression analysis ($R^2 = 0.186$, $p = 0.393$).

Patients undergoing TAAA repair for unruptured aneurysms (n=37) were then divided into quartiles using SMI scores. A two-sample t-test was performed to compare Accordion scores

between the lowest SMI quartile and highest SMI quartile. There was no significant difference in Accordion scores between the lowest SMI quartile (Mean = 4.5, SD = 3.73) and the highest SMI quartile (Mean = 5.33, SD = 4.76); t(10) = -0.338, p = 0.371.

Patients undergoing TAAA repair for unruptured aneurysms (n=37) were also divided into quartiles using SMD scores (mean skeletal muscle attenuation). A two-sample t-test was performed to compare Accordion scores between the lowest SMD quartile and highest SMD quartile. There was no significant difference in Accordion scores between the lowest SMD quartile (Mean = 5.5, SD = 6.5) and the highest SMD quartile (Mean = 4.5, SD = 2.51); t(10) = 0.351, p = 0.366.

Number of patients	48
Mean Accordion Score	6.5 (SD = 6.01)
Mean SMI	40.56 (SD = 7.81)
Deaths in F/U period	15
Mean time to mortality (SD)	Men – 12.32 months (21.19) / Women – 5.78 months (8.59)

Figure 31. Summary of outcomes analysis for patients undergoing TAAA repair

Discussion

This feasibility study confirms that the sarcopenia segmentation tool is a reliable and accurate tool to undertake skeletal muscle segmentation, and automating this process using this deep learning algorithm produces an automated segmentation output in a few seconds, as compared to the tedious, time-consuming process associated with manual segmentation. Integration into routine clinical workflow would allow these outputs (SMA and SMD) to be produced for every patient undergoing a CT scan to provide a marker of sarcopenia / frailty and aid in the clinical decision-making process.

There is, however, still some scope for improvement of this sarcopenia segmentation tool to further improve it's accuracy and reliability. The production of a single automated output for every CT scan would greatly improve it's ease of use as the researcher had to go through multiple outputs generated by the sarcopenia segmentation tool to identify the output with the highest slice detection confidence in the appropriate CT contrast phase. Selection of the appropriate contrast phase depending on the relevant specialty (for example, arterial phase in vascular surgical patients) or standardising the mean skeletal muscle attenuation output (in Hounsfield units) based on contrast phasing would also improve the ease of use the sarcopenia segmentation tool, rather than producing multiple outputs for multiple phases of the same CT scan series.

Whilst the segmentations undertaken by the sarcopenia segmentation tool were largely accurate, there were some instances of segmentation 'bleed out' to non-skeletal muscle tissue as demonstrated in the images below. This is likely to improve with further, ongoing training of the sarcopenia segmentation tool and it's associated deep learning algorithm.

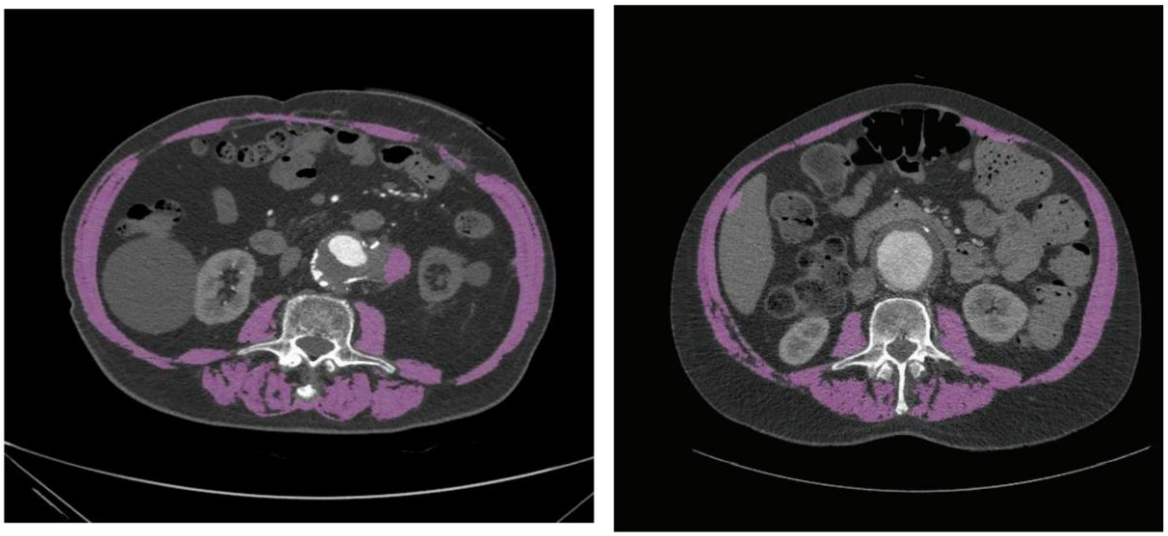


Figure 32. Examples of segmentation 'bleed out' by the sarcopenia segmentation on the tool. On the left, segmentation bleed out onto haematoma associated with a ruptured abdominal aortic aneurysm and on the right, segmentation bleed out onto the liver.

Data quality

The original cohort of patients for this retrospective feasibility study was chosen from the ICHT NVR records. It was felt that 500 patients would serve as an appropriate cohort size for this study and the initial list of patients associated with our study time period from the NVR database did produce over 500 patients. However, it was quickly realised after going through each individual patient's electronic records that several patients had been incorrectly coded for as having undergone AAA repair, and there was significant attrition in sample size on the basis of our pre-defined inclusion / exclusion criteria. This did have some impact on our overall study findings in terms of deriving statistically significant results. Whilst it is clear that the sarcopenia segmentation tool is as reliable and accurate as manual segmentation based on the current patient cohort, having a larger sample size may have helped perform more robust and significant outcomes analysis.

The most common reason for exclusion was patients having a TEVAR being coded for as AAA repair. There certainly does appear to be some scope in undertaking a separate study looking at this group of patients alone and performing outcomes analysis, with the advantage of being a relatively more homogeneous group with a similar post-operative risk profile, as compared to the wide variation seen in varying types of AAA repair.

The optimal method of improving data quality and having a more reliable dataset to perform robust outcomes analysis would be to undertake a prospective trial, ideally using multiple centres to capture more patients to assess the utility of this sarcopenia segmentation tool. This would come at no significant change to the pre-, peri- or post-operative course of each individual patient as there would be no change in patient management / intervention, but merely a standardised database to capture real time patient data and automated segmentation of pre-operative CT scans.

Is SMI or SMD a better predictor of post-operative outcomes?

From this study, it remains unclear as to what measure / marker of sarcopenia is a better predictor of post-operative outcomes. The automated sarcopenia segmentation tool has the benefit of producing both the skeletal muscle area (SMA) and mean skeletal muscle attenuation (in Hounsfield units) at the L3 level. SMA can then be used to calculate the SMI, and mean skeletal muscle attenuation is a measure of SMD, giving the clinician the flexibility of choosing one or both sarcopenia metrics whilst evaluating each individual patient. Whilst SMI / SMA are measures of skeletal muscle quantity, SMD provides a measure of skeletal muscle 'quality'. Further studies are required to interpret and clarify the superiority of one sarcopenia metric as a superior predictor of outcomes as compared to the other.

Within our study cohort, no differences were seen between SMI or SMD as predictors of outcomes in patients undergoing endovascular AAA repair or open TAAA repair. However, in open AAA repair, SMD appeared to be a better predictor of post-operative outcome in men, whereas SMI appeared to be a better predictor of post-operative outcome in women (though this was a small sample size).

Outcomes analysis

Statistically significant results were not obtained in most aspects of outcomes analysis due to the limited sample size and significant loss of patients from our original patient cohort of over 500 patients due to a variety of reasons as previously described. Further 'sub-group' analysis on the basis of sex and type of intervention further compounded this problem, and even precluded comparisons between SMI / SMD quartiles due to small numbers in certain patient

groups. However, performing statistical analysis on the entire cohort as one group, without stratification, would not have provided scientifically robust results as patients with wildly different characteristics and risk profiles would have been grouped together, introducing significant bias.

Further studies to perform robust outcomes analysis would benefit from a prospective design and selection of a relatively homogeneous group of patients such as patients undergoing standard infrarenal EVAR only, as these tend to be the most common intervention for AAA repair undertaken worldwide, which would likely enable easier and faster patient recruitment to achieve an adequate sample size based on power calculations.

However, this must also be balanced against the findings of this study where both SMI and SMD did not appear to have implications on the outcomes of patients undergoing EVAR (both men and women). This could possibly be explained by a couple of key limitations of this study. Firstly, patients undergoing both 'standard' and 'complex' EVAR were grouped together in the EVAR sub-group, though it is well known that the risks associated with complex EVAR intervention (such as fenestrated or branched repairs) are greater than standard infrarenal EVAR.

Secondly, the use of the Accordion score and 30-day composite outcomes as the primary outcome measure for an intervention that is well known to be 'low' risk in the immediate peri- and post-operative period, but carries a higher rate of re-intervention and 'late' rupture risk may not adequately reflect the appropriate post-operative outcome measure for outcomes analysis, and may be better assessed by the use of a longer term outcomes measure such as 'survival at 2-years' or long-term all-cause or aneurysm-related mortality.

What we can safely conclude from this feasibility study is that the automated sarcopenia segmentation tool is rapid and reliable, and can easily be utilised to measure sarcopenia in patients undergoing AAA repair, though the true effects of sarcopenia on post-operative outcomes needs further validation.

4. USE OF QUADRICEPS MUSCLE AS A SURROGATE MEASURE OF MUSCLE AREA AND DENSITY IN PATIENTS UNDERGOING ABDOMINAL AORTIC ANEURYSM (AAA) REPAIR

Abstract

Introduction

Skeletal muscle area (SMA), including psoas, at the level of the 3rd lumbar vertebral body (L3) has been extensively used to measure sarcopenia and predict outcomes in patients undergoing vascular surgery. However, this may be of limited use in patients undergoing surgical intervention for peripheral arterial disease (PAD) due to potential exclusion of L3 in cross-sectional imaging. We investigated the feasibility of using the quadriceps muscle as a surrogate measure of muscle area and density, as an alternative to psoas / L3 skeletal muscle area.

Methods

Manual segmentation of pre-operative CT scans of 48 patients, who underwent open or endovascular AAA repair at our institution between 2017 to 2021, were undertaken by two independent clinical researchers, to obtain cross-sectional skeletal muscle area (SMA) and skeletal muscle density (SMD) at L3, psoas, left (LQ), right (RQ) and bilateral quadriceps (BQ). A restricted analysis (n=38) was performed excluding patients with confounding factors on muscle segmentation.

Results

Excellent inter-rater reliability was noted between the manual segmentations undertaken by the 2 researchers (ICC $\geq$ 0.9, CI = 0.78 - 0.99, $p < 0.05$). Moderate correlation was noted between SMA at L3 and quadriceps muscle area (QMA) in both the overall and restricted analysis, though the degree of correlation was noted to be higher in the restricted analysis. Weak correlation was noted between SMD at L3 and quadriceps muscle density (QMD) in the overall

analysis, but the degree of correlation improved dramatically to at least moderate to good correlation in the restricted analysis.

Conclusion

The quadriceps muscle appears to be a reliable surrogate measure of skeletal muscle area and density. This may have utility in pre-operative risk stratification, particularly in PAD. Automation, using machine learning techniques, may further improve its accessibility. However, future studies with larger datasets are required to further validate this study's findings.

Background

As has been well established in the previous chapters, an objective or quantitative measure of sarcopenia and frailty would be useful in surgical risk prediction models and pre-operative risk stratification to aid the overall clinical decision making process. This would help identify patients at increased risk of post-operative morbidity and mortality, and potentially identify those patients who may benefit from prehabilitation to improve post-operative outcomes and / or also help choose the optimal type of intervention, particularly when there are several options available to the clinician and the patient. For instance, in vascular surgery, an infrarenal abdominal aortic aneurysm (AAA) may be treated via open surgical repair (OSR) or endovascular aneurysm repair (EVAR). The option to manage conservatively with no surgical intervention may also be considered in the patient that is considered too high risk, particularly in a disease process that is largely asymptomatic until rupture. However, making this decision can sometimes be challenging and frailty scoring or assessing the degree of sarcopenia in a patient may help ease this decision making process.

The risk of developing frailty increases with age, and this is particularly relevant to vascular surgical practice, where patients are commonly over the age of 60 years. According to the National Vascular Registry's (NVR) 2022 annual report (64), 80% and 94% of patients who underwent elective OSR and EVAR were older than 65 years respectively. Additionally, the Vascular Society of Great Britain and Ireland (VSGBI) recently reported that peripheral arterial disease (PAD) affects approximately 20% of the population over the age of 60 years.

The role of frailty and sarcopenia resulting in adverse post-operative outcomes and long-term survival is well documented in Chapter 2, along with the need to produce standardised and internationally accepted cut-off values for defining quantitative measures of sarcopenia. Whilst this thesis focusses mainly on the role of sarcopenia in post-operative outcomes of patients undergoing AAA repair, there are other areas of vascular surgery where this may be extended to, such as patients with lower limb peripheral vascular disease or cerebrovascular disease. It has been shown that increased frailty and sarcopenia have been independently associated with poorer overall survival and amputation-free survival in patients with critical limb ischaemia undergoing revascularisation (open, endovascular or hybrid intervention) (65).

Whilst skeletal muscle at the L3 level will routinely be included in all pre-operative CT scans of patients undergoing AAA repair, this may not be the case for patients undergoing intervention for peripheral vascular disease, as the CT scan may commence from a lower level to assess the lower limb vasculature. This would, therefore, require an alternative method of quantifying skeletal muscle area or density.

The quadriceps muscle is usually included in cross-sectional imaging to assess the lower limb vascular as part of work-up for lower limb revascularisation. This resulted in the consideration

of it's use as an alternative to skeletal muscle at the L3 level in this cohort of patients, as a surrogate measure of skeletal muscle area and density.

On conducting a scoping review of the literature, it was noted that only one study to date has investigated the use of quadriceps muscle segmentation in outcomes analysis. Segmentation of the quadriceps muscle was undertaken at an axial slice at the midpoint between the greater trochanter of the femur and superior border of the patella in this study (26). However, CT scans for patients undergoing AAA repair would routinely not go down to this level, and therefore, an alternative anatomical level of muscle segmentation would need to be considered in our patient cohort.

Aims and objectives

Following on from the SPAR study, a smaller feasibility study was planned to be undertaken to determine the feasibility of using the quadriceps muscle as an alternative measure of skeletal muscle area and density (as compared to skeletal muscle at the L3 level) in patients who have undergone AAA repair.

The specific objectives were:

1. To look for evidence of a correlation between skeletal muscle area of the right, left and bilateral quadriceps muscle (RQ, LQ and BQ), and skeletal muscle and psoas at L3 level.
2. To look for evidence of a correlation between skeletal muscle density of the right, left and bilateral quadriceps muscle (RQ, LQ and BQ), and skeletal muscle and psoas at L3 level.

3. To analyse inter-observer reliability of manual segmentations of right, left and bilateral quadriceps muscle (RQ, LQ and BQ), and skeletal muscle and psoas at L3 level.

Methods

As this study was undertaken under the auspices of the larger SPAR study, no further ethical approvals or registration were required.

50 consecutive adult patients who underwent endovascular or open abdominal aortic aneurysm repair from December 2017 to April 2021 at ICHT were identified from the SPAR database and chosen as the patient cohort for this feasibility study. This was the same group of patients who underwent manual segmentation of the skeletal muscle at the L3 level to serve as the 'ground truth' to compare and validate the outputs of the automated sarcopenia segmentation tool. The inclusion and exclusion criteria remained the same as the larger umbrella SPAR study. Pre-intervention, intervention and post-intervention data for each patient included in this feasibility study was extracted from the SPAR database on REDCap.

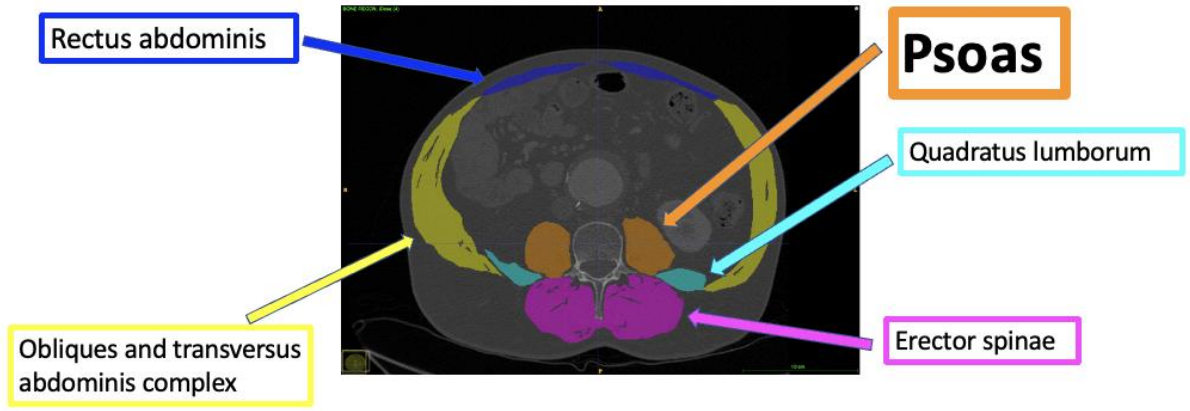
Muscle segmentation

Manual segmentation was carried out using ITK-SNAP software (59), which was also used for manual segmentation in the SPAR study. Skeletal muscle, including the psoas, was segmented on an axial slice on the pre-operative CT scan of each patient at the middle of the 3rd lumbar vertebra (L3). A threshold detection filter between -29 and 150 Hounsfield units (HU) was applied to these segmentations to remove macroscopic deposits of intramuscular fat and obtain a more accurate measure of skeletal muscle area and density.

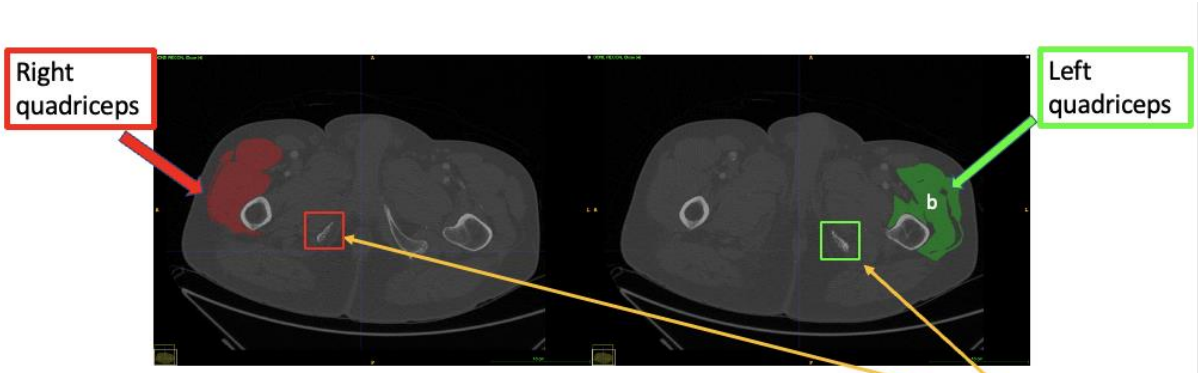
Quadriceps muscle was segmented at the most caudal axial slice (most inferior point) of the ischial spine. The most caudal / inferior aspect of the ischial spine would need to be visible on the axial slice chosen for quadriceps muscle segmentation. This was a purely arbitrary point, and it was felt that an easy to identify bony landmark would reduce observer / operator bias and help improve uniformity and study reproducibility. The right and left quadricep muscles were segmented separately and sometimes on slightly different axial slices to account for patient rotation / tilt whilst on the CT scanner, at the most caudal / inferior aspect of the respective ischial spine.

ITK-SNAP software generated the volume (mm^3) of each quadriceps muscle complex, and this was then converted into cross-sectional quadriceps muscle area(mm^2), based on slice thickness (skeletal muscle area = skeletal muscle volume / CT slice thickness in mm). This segmentation software also produced skeletal muscle density in Hounsfield units.

Manual segmentations on ITK-SNAP for each CT scan included in this study were performed by two independent clinical researchers. These included a specialist registrar in vascular surgery and an intercalating medical student, who were both trained and supervised by the overall project supervisor (AR - Professor of Radiology).



Manual measurement of abdominal muscle bulk using computed tomography (CT). CT Abdomen at level of L3 vertebra.



Manual measurement of quadriceps muscle bulk using computed tomography (CT). CT Abdomen at level of lowest visible point of ischial spine.

Figure 33. Examples of L3 skeletal muscle and quadriceps muscle segmentation at anatomical levels defined by the study methodology

Statistical analysis

Statistical analysis was performed using Prism 9 (Version 9.5.1 (528), January 24, 2023) and R (Version 2023.03). Study data distribution and normality was assessed visually using histograms and Q-Q plots, and statistically using the Shapiro-Wilk test. Pearson's correlation coefficient (r) was used to assess the correlation between manual segmentation of quadriceps muscle and skeletal muscle / psoas at the L3 level. Heatmaps were generated based on the Pearson correlation coefficient matrix for visualization of the significance of linear correlation.

Intraclass correlation coefficients (ICC) estimates along with their 95% confidence intervals (CIs) were calculated using a two-way random, absolute agreement, single measure model with 95% CI for the values from manual segmentation between both observers. ICC values less than 0.5 suggest poor agreement, 0.5–0.75 moderate agreement, 0.75–0.9 good and greater than 0.9 excellent agreement (66). Bland-Altman plots were generated (difference *against* means) for measurements of SMA comparing both observers.

Results

Upon commencement of manual segmentation of CT scans, it became apparent that certain scans could not be segmented appropriately (both in terms of skeletal muscle area and attenuation), and it was decided to perform a restricted analysis by excluding these scans from the overall analysis. This was usually due to the presence of metal implants (usually from hip arthroplasty) in close proximity to the area of muscle segmentation, causing streak artefact. Whilst this not only made the segmentation difficult due to inability to trace muscle borders accurately, the streak artefact resulted in unreliable and inaccurate skeletal muscle attenuation values (in Hounsfield units). Another common cause for exclusion from the overall analysis was due to inability to manually segment psoas and skeletal muscle borders at the L3 level due to presence of haematoma from ruptured AAAs obscuring tissue planes in the retroperitoneum and it's associated musculature. This prevented accurate statistical correlation analysis between skeletal muscle segmentation at the L3 level and quadriceps.

Therefore, a restricted analysis was performed excluding such cases as they presented as significant outliers in the overall analysis as can be seen in the scatter plots in the figures below.

A study flow diagram is provided below demonstrating the overall study sample size and reasons for inclusion / exclusion, and final sample size (restricted analysis) for testing the utility of quadriceps muscle segmentation as a surrogate measure of skeletal muscle area / density as compared to skeletal muscle at the L3 level (including psoas muscle).

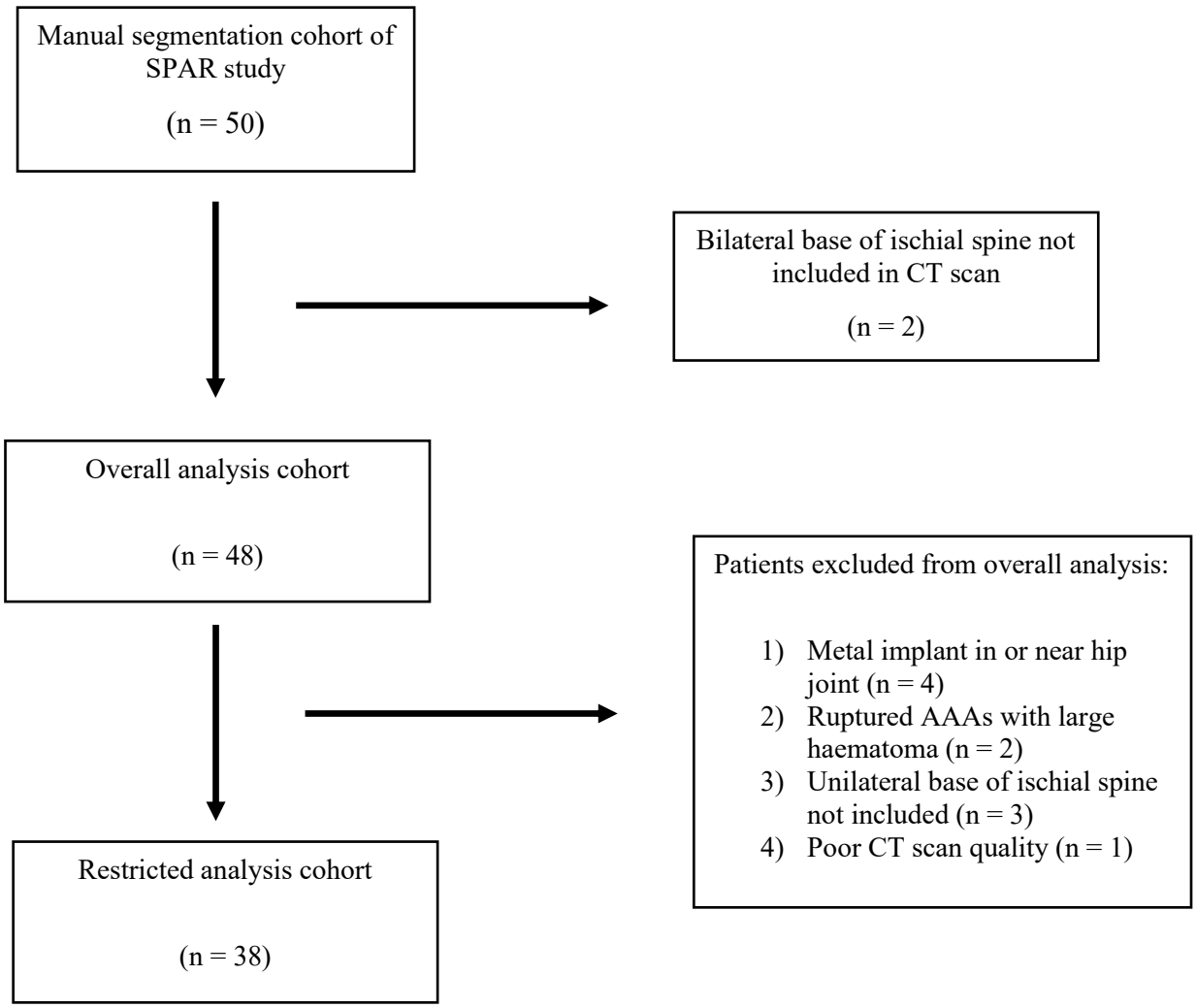


Figure 34. Study flow diagram demonstrating sample size, and inclusion / exclusion criteria for overall and restricted analysis

Study Demographics

The patient demographics / characteristics were the same as those of the manual segmentation cohort ('ground truth' cohort) in the main SPAR study, but for ease of access, these have been provided again in the table below.

Sex	Male (%) Female (%)	44 (88) 6 (12)
Type of AAA repair	Open infrarenal repair (%) Open juxtarenal repair (%) Infrarenal EVAR (%) Complex EVAR (%) TAAA Repair (%)	13 (26) 5 (10) 10 (20) 17 (34) 5 (10)
Mean Age (IQR)		74.76 (69 – 81.75)
Mean Height (IQR) - cm		173.64 (168 – 178)
Mean Weight (IQR) - kgs		76.97 (66 – 83)
Mean BMI (IQR)		25.48 (22.52 – 27.58)
Co-morbidities (%)	Diabetes Hypertension Coronary Artery Disease	7 (14) 37 (72) 16 (32)

	Respiratory Cerebrovascular Chronic Kidney Disease	16 (32) 4 (8) 7 (14)
Smoking status (%)	Never smoker Ex-smoker Current smoker Not known	5 (10) 26 (52) 13 (26) 6 (12)
Medications (%)	Antiplatelet/s Anticoagulant Antihypertensive/s Statin	36 (72) 4 (8) 34 (68) 44 (88)

Figure 35. Study demographics of the quadriceps segmentation study to assess the utility of quadriceps muscle as a surrogate muscle marker to skeletal muscle at the L3 level

Inter-rater reliability between Manual Segmentations

Inter-rater reliability (IRR) between the 2 independent clinical researchers undertaking manual segmentation was measured using intraclass correlation coefficient (ICC) as the data to be compared was continuous quantitative data.

As demonstrated in the table below, overall excellent inter-rater reliability was seen in manual segmentation measurements of the quadriceps muscle (RQ, LQ and BQ) and skeletal muscle at the L3 level, whilst good inter-rater reliability was seen in psoas muscle segmentation.

In the restricted analysis, excellent inter-rater reliability was seen between manual segmentations of all investigated muscle groups.

	Overall Analysis			**Restricted Analysis**		
Muscle Group	ICC	95% CI (Lower, upper bound)	P-value	ICC	95% CI (Lower, upper bound)	P-value
RQ	0.98	0.97, 0.99	<0.05	0.98	0.97, 0.98	<0.05
LQ	0.97	0.95, 0.98	<0.05	0.97	0.95, 0.99	<0.05
BQ	0.98	0.97, 0.99	<0.05	0.98	0.97, 0.98	<0.05
Psoas	0.88	0.78, 0.93	<0.05	0.90	0.81, 0.95	<0.05
L3	0.97	0.96, 0.99	<0.05	0.98	0.96, 0.99	<0.05

Figure 36. Intraclass correlation coefficient (ICC) to compare inter-rater reliability of manual muscle segmentations undertaken by the 2 clinical researchers, in both overall and restricted analyses

Bland-Altman plots (below) to compare inter-rater reliability of manual segmentations between the 2 independent clinical researchers also confirmed good inter-rater reliability, in both the overall and restricted analysis. This may improve reproducibility of a similar segmentation method and model in future studies assessing the use of quadriceps muscle segmentation, including in frailty assessments and outcomes analysis.

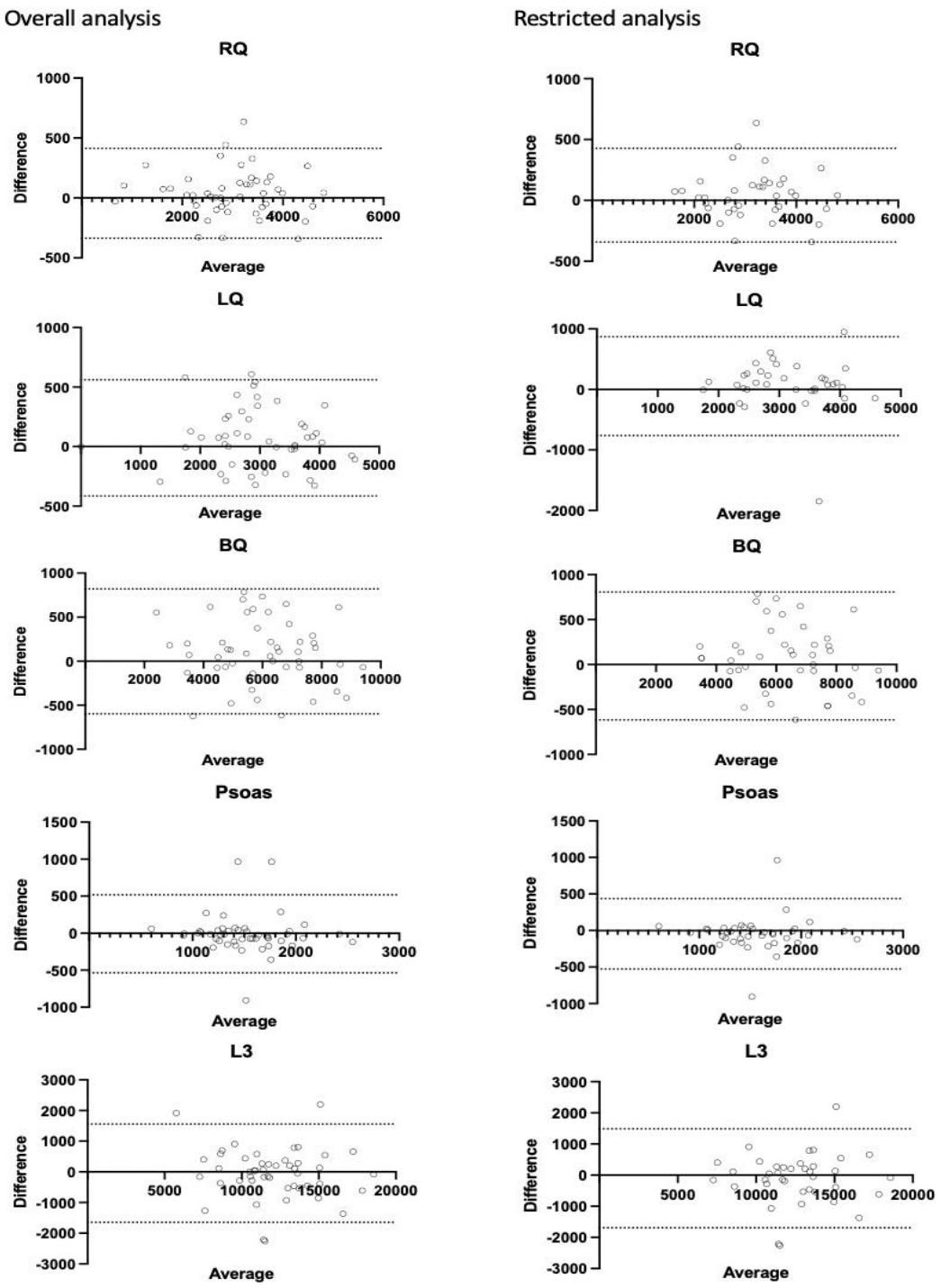


Figure 37. Bland-Altman plots demonstrating excellent IRR between manual segmentations undertaken by the 2 clinical researchers, in both overall and restricted analyses.

Correlation Heatmaps – Skeletal Muscle Area

Correlation heatmaps were generated after calculating the Pearson's correlation coefficient to assess if there was any linear correlation between quadriceps muscle area segmented on an axial slice at the lowest visible point (or most inferior aspect) of the ischial spine, and skeletal muscle area (SMA) segmented on an axial slice passing through the middle of the 3rd lumbar vertebral body (L3).

The correlation of SMA at L3 with quadriceps muscle area was significant ($p<0.05$) both when considering unilateral quadriceps muscle groups (LQ or RQ) or bilateral quadriceps (BQ).
Moderate correlation was noted between SMA at L3 and quadriceps muscle area (QMA) in both the overall and restricted analysis, though the degree of correlation was noted to be higher in the restricted analysis.
However, only weak correlation was noted between psoas muscle area (PMA) at L3 and quadriceps muscle area (QMA) in both the overall and restricted analysis, though the degree of correlation was noted to be higher in the restricted analysis.
The relevant scatterplots and correlation heatmaps are provided below.

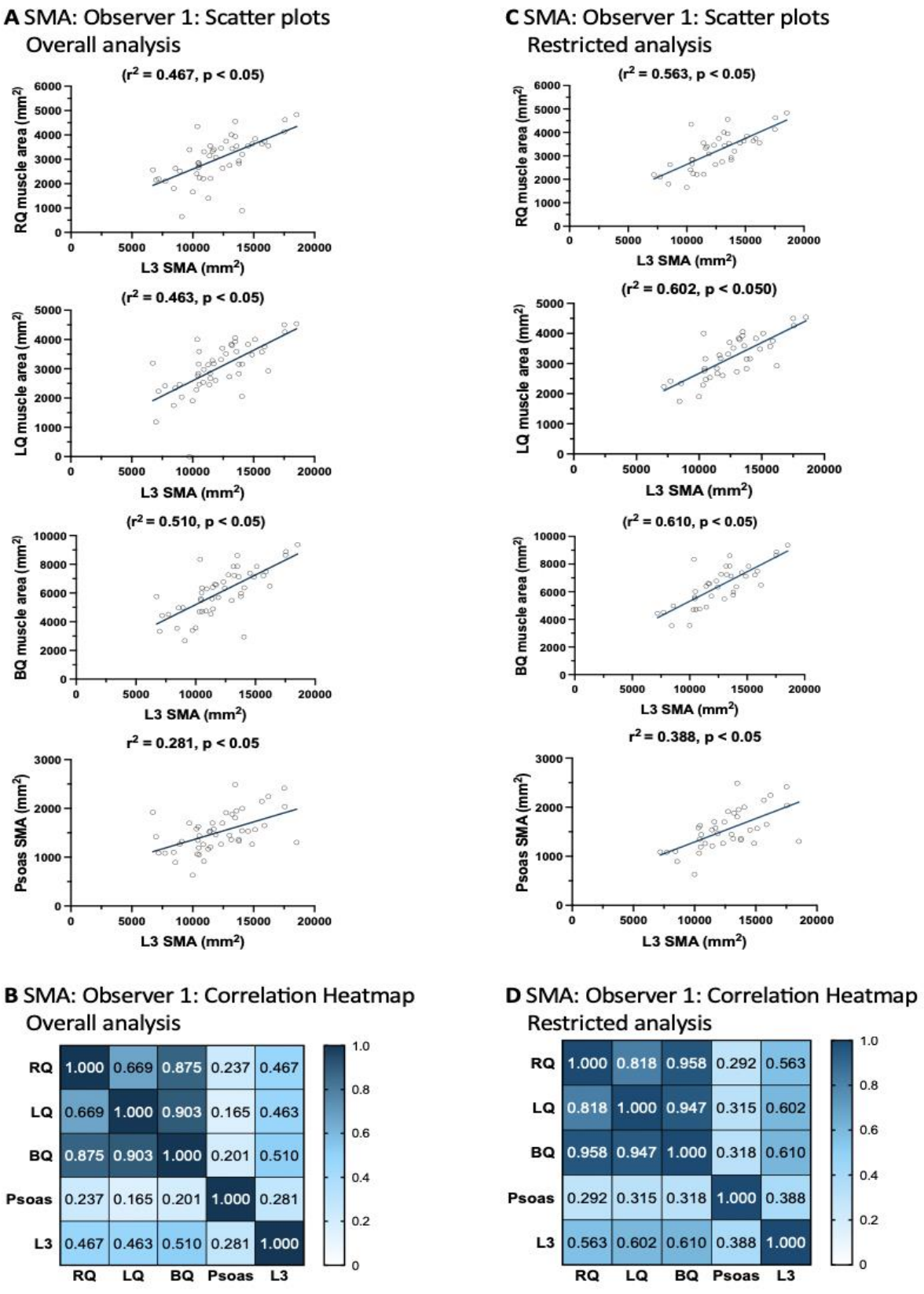


Figure 38. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle area against SMA at L3 for observer 1

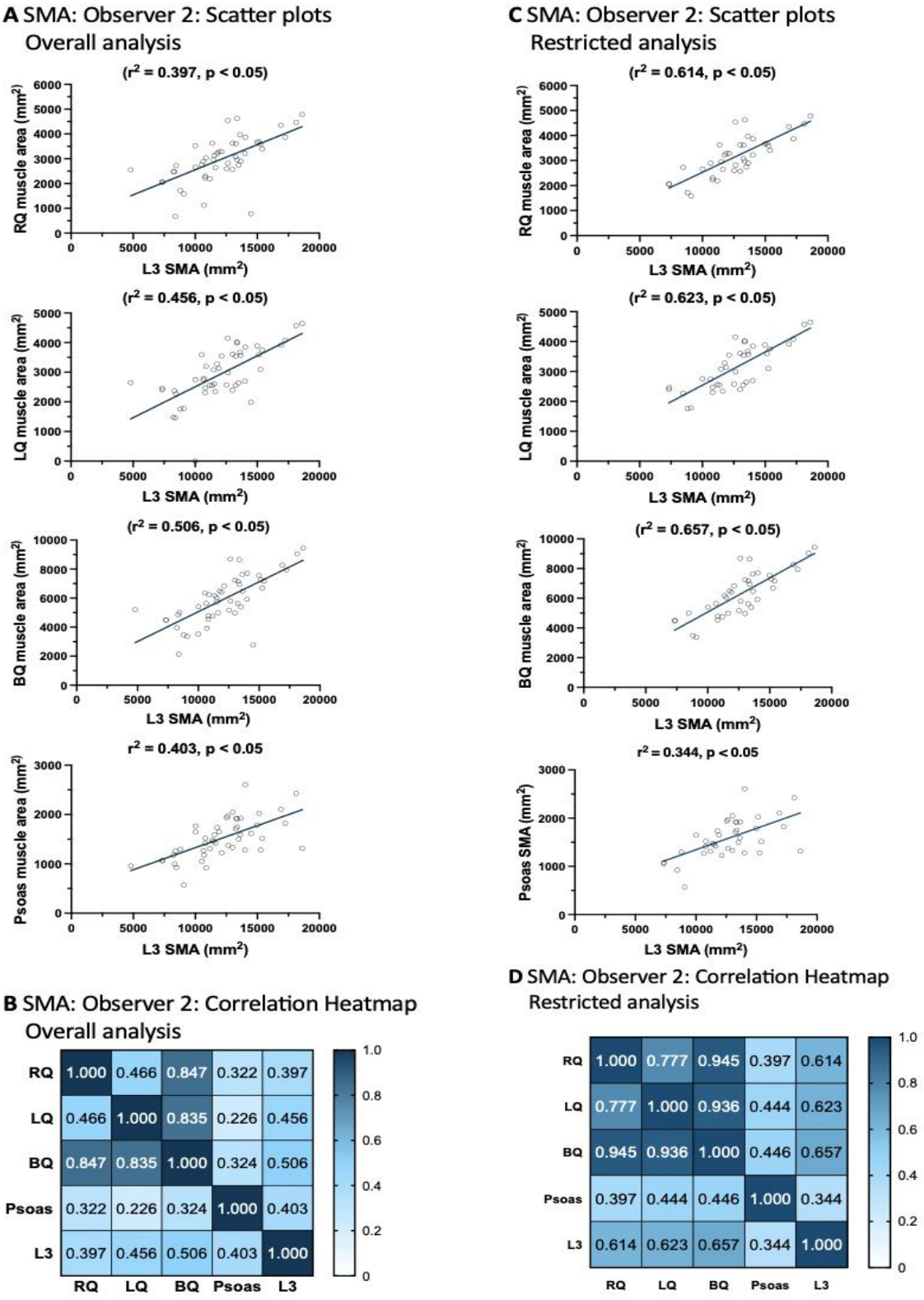


Figure 39. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle area against SMA at L3 for observer 2

Correlation Heatmaps – Mean Skeletal Muscle Attenuation

Correlation heatmaps were generated after calculating the Pearson's correlation coefficient to assess if there was any linear correlation between quadriceps muscle attenuation segmented on an axial slice at the lowest visible point (or most inferior aspect) of the ischial spine, and skeletal muscle attenuation (including psoas muscle attenuation at this level) segmented on an axial slice passing through the middle of the 3rd lumbar vertebral body (L3). As alluded to in previous chapters, mean skeletal muscle attenuation (in mean Hounsfield units – MHU) which is also produced by the muscle segmentation tool is a measure of skeletal muscle density (SMD), which is a surrogate marker of skeletal muscle 'quality'.

The correlation of SMD at L3 with quadriceps muscle density was significant ($p<0.05$) both when considering unilateral quadriceps muscle groups (LQ or RQ) or bilateral quadriceps (BQ).

Weak correlation was noted between SMD at L3 and quadriceps muscle density (QMD) in the overall analysis, but the degree of correlation improved dramatically to at least moderate to good correlation in the restricted analysis.

Weak correlation was noted between psoas muscle density (PMD) at L3 and quadriceps muscle density (QMD) in the overall analysis, but the degree of correlation improved dramatically to at least moderate correlation in the restricted analysis.

The relevant scatterplots and correlation heatmaps are provided below.

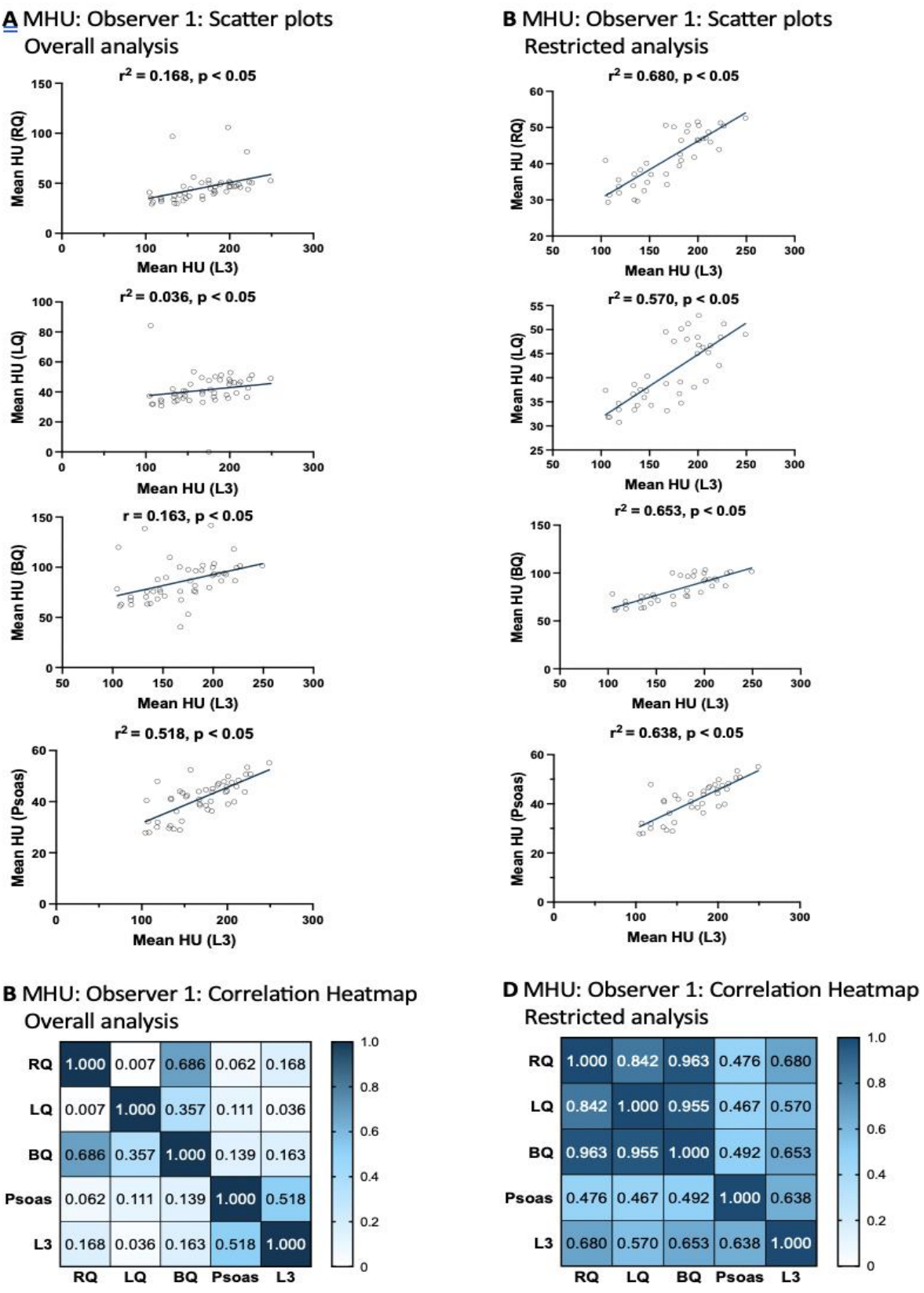


Figure 40. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle density against SMD at L3 for observer 1

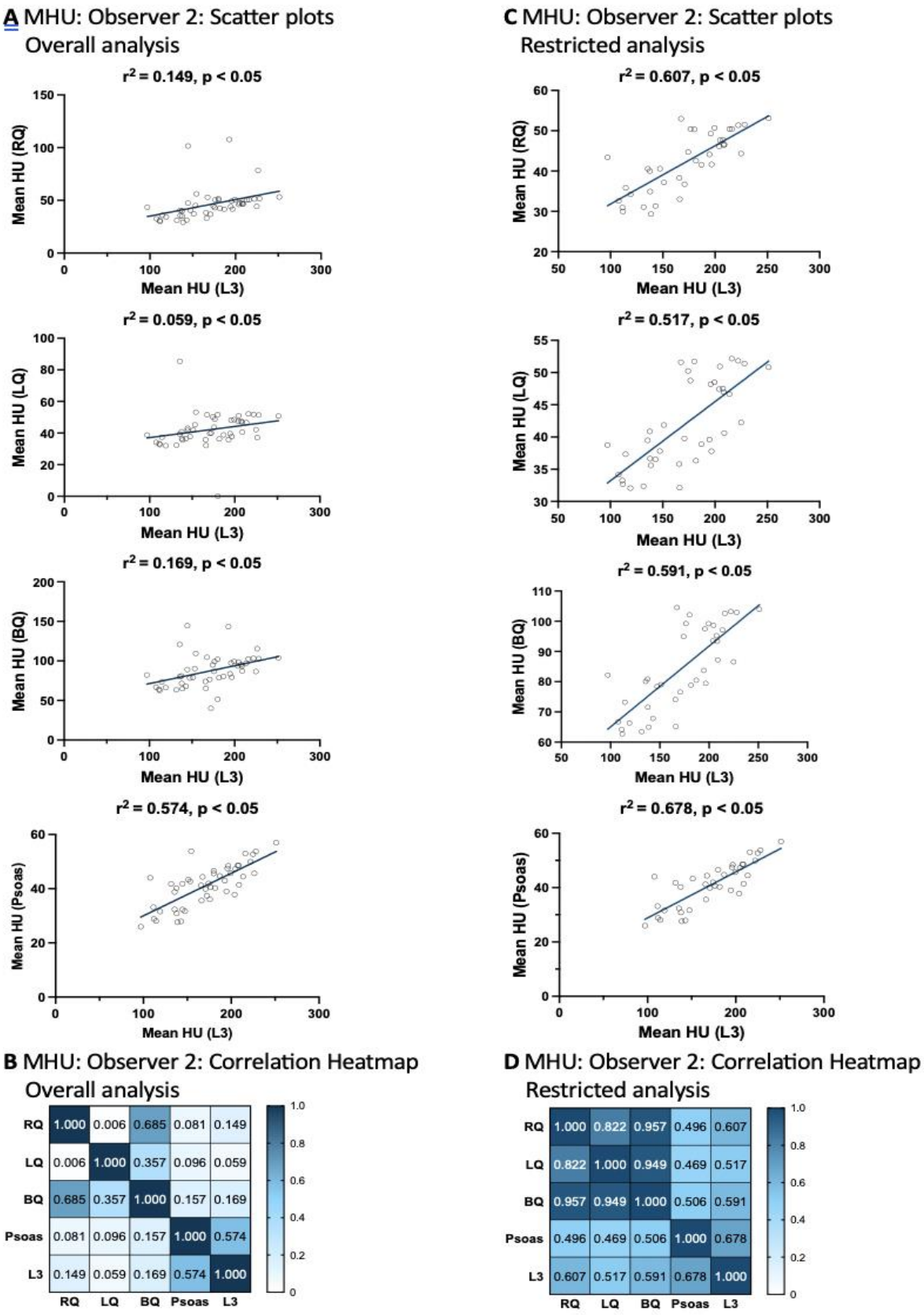


Figure 41. Scatterplots and correlation heatmaps comparing quadriceps and psoas muscle density against SMD at L3 for observer 2

Discussion

This small feasibility study was undertaken to assess the use of alternative muscle groups to skeletal muscle at the L3 level. While axial / truncal skeletal muscle has been studied extensively in frailty and sarcopenia studies, and has also been utilised in outcomes analysis, the use of skeletal muscle from the limb to assess for frailty and sarcopenia and perform outcomes analysis is still in it's infancy. This is particularly relevant when there is no suitable imaging or other similar investigation that maps out axial / truncal skeletal muscle, but cross-sectional imaging of the limb may be available, such as might be the case in lower limb peripheral vascular disease imaging and work-up.

Whilst the use of skeletal muscle at the L3 level has been validated in the previous chapter of this thesis, this small study has demonstrated that a linear correlation does exist between SMA and SMD of the quadriceps muscle and skeletal muscle at the L3 level. A larger cohort study would be required to further validate these findings, and a prospective study could then be considered to assess whether it would be a suitable marker for frailty or sarcopenia, and whether it could be utilised to perform outcomes analysis. If the overall findings are encouraging, it may also be considered as an alternative to axial / truncal skeletal muscle to work out which muscle group is a more sensitive predictor of frailty and long-term outcomes.

Further assessment in other surgical and medical specialties would also help accept or reject these hypotheses.

This feasibility study has also helped identify patient groups in which the utilisation of quadriceps segmentation would not be recommended.

This is particularly the case with any metal implants in and around the quadriceps muscle bulk, for example, in the case of hip arthroplasty or intramedullary nails. Whilst the streak artefact

significantly impairs the ability to perform accurate muscle segmentation, it also alters muscle attenuation significantly giving false / incorrect readings for skeletal muscle density (in Hounsfield units). This is clearly seen on the restricted analysis as removal of these patients in correlation modelling significantly improved the linear correlation between quadriceps muscle attenuation and skeletal muscle attenuation at the L3 level.

This restricted analysis will help guide future feasibility large cohort studies by helping define more accurate inclusion / exclusion criteria in future study methodology to help obtain more accurate and robust results.

This study also demonstrates that quadriceps muscle area and attenuation appears to have a better linear correlation with skeletal muscle area / attenuation at the L3 level as compared to psoas muscle area / attenuation, highlighting potential limitations of using psoas muscle area / attenuation alone as a marker of frailty and predictor of post-operative outcomes.

Excellent interrater reliability (IRR) was noted between manual segmentations of both clinical researchers, though both clinical researchers were in different stages of their medical careers with varying experience in the use and interpretation of cross-sectional imaging, and in muscle segmentation using open access software. Whilst one clinical researcher was an intercalating medical student, the other clinical researcher was a MD (Res) student and vascular surgical trainee with almost 10 years of clinical experience post qualification as a medical practitioner. The intercalating medical student needed manual segmentation training for about 120 minutes before she was able to undertake independent manual segmentations using ITK-SNAP open access software.

This suggests that manual segmentation is easy to learn and can be undertaken after a very short training period. Excellent IRR also suggests that this segmentation method is easily reproducible should large cohort studies be considered in the future.

With further validation in larger prospective studies, this quadriceps segmentation model may be utilised to help accurately identify patients with frailty and sarcopenia and / or help in frailty scoring. This would, in turn, identify those patients at significant risk of poor outcomes and help stratify patients into appropriate treatment groups and also help identify those patients who may benefit from prehabilitation to improve post-operative outcomes.

Strengths and limitations of this study

This study has several strengths. It has identified and validated to some extent the use of the lowest or most inferior point of the ischial spine as an appropriate level to segment the quadriceps muscle. The utilisation of a clear, well defined bony landmark would also help generalise the study methodology for future similar studies, and help improve reproducibility.

Excellent interrater reliability was established by using intraclass correlation coefficients and Bland-Altman plots. This also shows that the study methodology and manual segmentation methods are highly reproducible even amongst those researchers with little to no experience of manual segmentation or medical imaging.

Key inclusion and exclusion criteria for similar future studies have been identified and the reasons for these, particularly with regards to skeletal muscle attenuation in particular have been demonstrated.

The study has also shed some light on the potential limitations of using the psoas muscle complex on it's own as a reliable muscle marker, and utilisation of all skeletal muscle at the L3 level (including the psoas muscle) may be a more accurate marker of skeletal muscle area and attenuation.

The results of this study must be interpreted within the limitations of this study. The small sample size of this study reduces the reliability and generalisability of it's pertinent findings and prevented any robust outcomes analysis from being undertaken, whilst introducing some selection bias.

Confounding variables such as height, BMI, sex and co-morbidities which may have had effects on skeletal muscle area and density were no accounted for in statistical modelling, as this was beyond the scope of this small feasibility study. Such factors would need to be considered in future studies as certain phenotypes may predispose to higher skeletal muscle quantity and quality in the truncal regions as compared to the limb, or vice versa.

The choice of statistical test for assessing the correlation between quadriceps muscle and skeletal muscle at the L3 level may also be debated on further. Despite data for some muscle groups in terms of muscle density being non-parametric, Pearson's correlation coefficient (r) was used. This is because the strength of linear association was the main outcome measure considered in this study. Whilst a non-parametric statistical correlation test such as Spearman's rank could have been used, this test risks data / information loss (in an already small study) because it would have treated this study's continuous data as ordinal by assigning ranks.

It should also be noted that normality tests, such as the Shapiro-Wilk test are not statistically robust and may be misleading in sample sizes of 50 and less, which was the case in our study.

However, the Q-Q plots and histograms did show that for both overall and restricted analyses, variables determined non-parametric by the Shapiro-Wilk test remained non-parametric, though however, their degree of skewness and kurtosis was not wide-ranging. Nonetheless, the use of a parametric test on non-parametric data remains a limitation of this study, reducing it's overall reliability.

Several steps need to be taken to further validate this small feasibility study. Prospective studies adopting a similar study methodology in larger patient cohorts would need to be undertaken in both, similar and varied patient populations, and in those patients with peripheral vascular disease. Studies on patients with peripheral vascular disease would also help delineate the effects of this pathology on lower limb skeletal muscle loss, and it's associated effects on mobility and prognosis. The loss of skeletal muscle over time would also be another useful avenue of clinical research, and it's resulting effects on long-term morbidity and mortality.

Future studies should also aim to determine threshold cut-off values to define sarcopenia to help identify frail and sarcopenic patients from non-sarcopenic patients as to date no universal cut-off values exist for the same.

If quadriceps muscle is found to be a good marker of frailty and sarcopenia, automating this segmentation process in a similar model as described in the previous chapter using a deep learning algorithm may help integrate it's use into day-to-day routine clinical workflow. This study methodology, particularly with regards to segmentation methods, would also need to be reviewed to determine how it may be adapted for patients with metal implants in and around the quadriceps muscle complex. ways in which to improve the availability and quality of existing CT imaging.

With further validation in future studies, the quadriceps model may be able to more accurately detect occult frailty and sarcopenia, and help identify these patients with frailty to help guide clinical decision making and optimise relevant treatment strategies. Particularly in a field like vascular surgery where both open and endovascular surgery are available to treat similar pathologies with varying risk profiles, it may help the patient and clinician reach a joint decision on which treatment modality will be best suited for their physiological reserve and who would benefit the most from pre-operative prehabilitation. This may be in the form of pre-operative nutritional optimisation, exercise training / physiotherapy and / or lifestyle modifications such as smoking cessation with the overall aim of improving patients' post-operative outcomes.

Automation of this segmentation process (if validated further), may help provide clinicians with an objective, quantitative frailty scores from pre-operative CT scans to use in the clinical decision making process.

5. CONCLUSIONS & FUTURE WORK

Sarcopenia as a disease process

As described in this thesis, frailty and sarcopenia is a complex pathophysiological process that can have detrimental effects on patient well-being and functional status. Development of sarcopenia further exacerbates the problem as it limits overall patient mobility and activity, leading to further muscle loss / atrophy due to disuse. The effects on psychological and mental health wellbeing are also well documented (19). Early recognition of frailty and sarcopenia may help reverse or slow down this process, by enabling nutritional, functional, psychosocial and physical optimisation or pre- / rehabilitation to optimise the patient's overall condition.

However, the diagnosis of sarcopenia may be challenging and includes multiple domains such as muscle strength, muscle quality, muscle quantity and physical performance. Assessing muscle strength and physical performance is challenging in day-to-day clinical practice due to limited capacity, staffing, testing space and financial constraints and the lack of a universally accepted 'gold standard' test to measure / diagnose sarcopenia and frailty.

Whilst there is no gold standard test to assess muscle quality or quantity either, the diagnosis of sarcopenia using these domains may be more straightforward as it potentially does not require increased capacity, staffing or testing space. Various studies (as described in chapter 2 of this thesis) have demonstrated the role of skeletal muscle index (SMI) as a suitable surrogate marker of sarcopenia, and whilst it was initially used in the field of oncology, it has been extensively studied and adopted in a variety of surgical specialties including vascular surgery.

Our systematic review and meta-analysis on the utility of SMI in vascular surgery clearly demonstrates that patients with low SMI do worse as compared to patients with 'normal' SMI in terms of 30-day or in-hospital mortality, post-operative morbidity / complications and all-cause long-term mortality.

This must be tempered against the limitations of this systematic review and meta-analysis, which primarily included the inclusion of retrospective single-centre cohort studies and the lack of standardised cut-off values for SMI.

Future research should aim to address these limitations by providing high quality evidence using prospective multi-centre studies or randomised control trials, though the associated time taken to produce this evidence and the financial implications of these would also need to be considered.

Whilst guidelines exist on a myriad of medical and surgical pathologies, including sarcopenia (such as the EWGSOP guidelines), the lack of standardised cut-off values for sarcopenia and SMI was clearly highlighted by our systematic review where all included studies defined sarcopenia differently, by having varying cut-off values for low and normal SMI. This introduces significant bias as what one study may consider 'sarcopenic' or 'low SMI' may be considered entirely normal by another study. The urgent development of standardised cut-off values for sarcopenia / SMI, ideally using population-wide epidemiological studies, would enable future research studies to adopt these values and stratify patients into appropriate groups to enable more robust statistical and outcomes analysis.

Though we have recognised sarcopenia and low SMI to have poor outcomes in patients undergoing surgical intervention, the measurement of SMI is another challenge. Whilst most studies use manual segmentation using a variety of segmentation tools and software, this is

clearly a time consuming process and not a practical use of a clinicians' limited time in today's day and age, particularly with the ongoing system-wide pressures in a healthcare system like the NHS.

From my personal experience of manual segmentation of CT scans for the purposes of validation and ground truth in the SPAR study, it was clear that manual segmentation of each CT scan would take at least 10-15 minutes, despite adequate training and repetitive experience. This was not including the preliminary steps, time and effort associated with pseudonymisation of CT scans for research purposes.

It is clear that this is not a practical use of a clinicians' time in the out-patient clinic or pre-assessment clinic setting and, therefore, manual segmentation of CT scans is likely to be confined to the realms of research studies and clinical trials.

Automation of this segmentation process using modern artificial intelligence technologies may be the solution to this problem to allow incorporation into clinical practice to aid the clinician and patient in the decision-making and risk-stratification process, and help recognise patients at high risk from surgical intervention, with the aim of negating and managing these risks in the peri- and post-operative period.

The automated sarcopenia segmentation tool

With the advent of artificial intelligence and it's rapidly evolving and increasing uptake and utility in today's fast-paced world, the Comprehensive Cancer Imaging Centre at Imperial College London designed a deep-learning section-based muscle segmentation tool for

sarcopenia assessment, and validated it's use in oncology patients. Considering it's uptake and utilisation had not been undertaken by other medical or surgical specialties, our study (SPAR) was designed to test the feasibility of using this fully automated sarcopenia segmentation tool in a cohort of vascular surgical patients to test validity and reliability. Our study also aimed to evaluate whether the outputs so produced by this sarcopenia segmentation could be utilised to predict outcomes in patients undergoing vascular surgical intervention.

The SPAR study further confirms the reliability and validity of the sarcopenia segmentation tool as excellent inter-rater reliability was noted between manual segmentations of CT scans that served as ground truth, against the automated segmentations produced by the sarcopenia segmentation tool, both for skeletal muscle area (SMA) and skeletal muscle density (SMD). The outputs produced by the sarcopenia segmentation tool appeared very 'user-friendly' and easy-to-read with all the key data in terms of patient characteristics and segmentation outputs presented succinctly. Whilst there is still some scope for improving the deep-learning algorithm to further improve the accuracy of segmentation outlines and outputs, the results are overall very encouraging.

However, the outcomes analysis performed utilising the automated sarcopenia segmentation outputs produced some mixed results. In patients undergoing open AAA repair (men and women), low SMI or SMD was associated with poor outcomes. Recognition of these patients in the pre-operative clinical assessment would potentially enable further optimisation or prehabilitation to improve sarcopenia and / or frailty with a view to reducing peri-operative risks and improving post-operative outcomes. The optimal treatment strategy could also be reconsidered in these patients with low SMI / SMD as they may benefit from a less invasive

treatment option such as endovascular repair (if technically feasible) or conservative / non-operative management.

In patients undergoing endovascular repair of their AAA, SMI or SMD did not appear to have a significant effect on post-operative outcomes in our patient cohort, though this may be limited by the heterogeneity in the nature of endovascular repair undertaken in this cohort, as repairs included the entire spectrum of endovascular intervention from standard infrarenal EVAR to complex FEVARs / BEVARs.

Whilst the overall outcomes analysis did produce some interesting results, it was limited by the smaller sample size than originally planned, largely due to incorrect coding of AAA repairs on the institutional and NVR database. The sample size could potentially have been increased by including patients undergoing AAA repair prior to 2016, but this would have involved retrieving physical patient paper notes and sifting through these, which would have been more time-consuming and of limited value, with questionable accuracy.

Future studies should ideally be of prospective nature and larger sample size, and include multiple centres undertaking AAA repair to reduce institutional bias from practices followed by a single centre, and also improve homogeneity of intervention by having stricter inclusion / exclusion criteria and perhaps including patients undergoing only a specific type of aortic intervention, such as standard infrarenal EVAR. Another alternative would be to move away from including patients undergoing AAA repair, and consider an alternate intervention such as TEVAR, where the risk profile from the intervention would be more uniform. Another option would also to be look at long-term outcomes / survival rather than 30-day post-operative composite outcomes, particularly with regards to endovascular intervention, where the risks tend to be low in the immediate post-operative period, but increase with time, particularly with regards to re-intervention and late ruptures / mortality.

It also remains unclear as to whether SMI or SMD is a better surrogate measure of sarcopenia, and this is another avenue of future research. Whilst SMI provides a marker of skeletal muscle quantity, SMD provides a marker of skeletal muscle 'quality'. It is well known that sarcopenia results in fatty infiltration of skeletal muscle, reducing it's overall functionality (33). Whilst this may not result in reduced SMA (and SMI), it would significantly affect SMD and, therefore, this may suggest that SMD is a more useful marker of sarcopenia than SMA. However, thresholding during both manual and automated segmentation, which was utilised in our study, aimed to exclude macroscopic fat deposits within skeletal muscle during the segmentation process and this may help mitigate incorrectly high SMA / SMI values in these patients.

Use of quadriceps muscle as a marker of sarcopenia

Whilst cross-sectional skeletal muscle area at the L3 or L4 muscle has been extensively studied in sarcopenia and frailty analyses, there is an increasing drive to recognise other muscle groups that may be more accurate surrogate markers of sarcopenia, particularly when imaging of the lumbar spine and it's surrounding area is not included.

This gave rise to the idea of investigating the feasibility of using quadriceps muscle area and density, particularly in patients undergoing lower limb vascular intervention, where imaging may not extend up to the lumbar spine.

As there is very limited existing literature on quadriceps muscle segmentation, we formulated our own level of segmentation utilising an easily identifiable bony landmark (base of ischial

spine) to improve reproducibility for future studies that may investigate the role of quadriceps muscle in frailty analyses.

Manual segmentation of the quadriceps muscle was relatively easily performed by both clinical researchers performing these segmentations, and the learning curve did not appear to be too steep. The results were overall encouraging with good correlation noted between SMA / SMD at L3 and QMA / QMD. Early recognition of confounding factors, particularly with regards to streak artefact from metalwork around the hip joint and exclusion of these patients from the restricted analysis improved correlation significantly, and will serve as useful marker for future larger scale studies, particularly with regards to inclusion / exclusion criteria.

Future research should aim to validate the findings of this pilot feasibility study, by utilising a larger cohort of patients, ideally in a prospective manner, modifying these inclusion / exclusion criteria and also investigating which of the 2 muscle groups (quadriceps vs L3 skeletal muscle) is a better surrogate measure of sarcopenia.

What is also interesting to note from this small feasibility study was that quadriceps muscle area / density had a better correlation with L3 skeletal muscle area / density as compared to psoas muscle area / density. Whilst the psoas muscle complex has been extensively studied in the past in sarcopenia and outcomes-based analysis, recent studies have moved away from this by utilising all the skeletal muscle at the L3 level (rectus abdominis, obliques and transversus abdominis complex, erector spinae, quadratus lumborum and psoas) and not the psoas muscle complex alone. This trend was also further validated by our study.

Avenues for future research

Besides the need for further validation of our feasibility studies by using prospective multi-centre large cohort studies or clinical trials, there are several other avenues for future research identified by this thesis encompassing the domains of frailty, sarcopenia, vascular surgery, outcomes-based analysis and artificial intelligence.

The role of change of muscle bulk over time is another area worth investigating further, particularly whilst measuring frailty / sarcopenia and associated post-operative outcomes. Whilst we did collect post-operative CT scans at least 6 months after AAA repair, it was beyond the scope and capacity of the SPAR project at the time to undertake segmentation of skeletal muscle area / density of these scans to compare it against pre-operative CT scan segmentation data. However, the rate of change of muscle area / density over time may serve as a useful marker of frailty / sarcopenia and identify those patients at increased risk of morbidity / mortality, and also identify patients at increased chance of long-term survival and who may benefit the most from re-intervention (if required) particularly in endovascular intervention where the rate of re-intervention tends to be higher.

Conversely, CT scans in the pre-operative phase before and after a period of pre-habilitation and optimisation may help identify which patients benefit the most from these interventions, why certain patients benefit more than others, and which patients are more likely to have poorer peri- and post-operative outcomes if no positive change is noted on their CT scans after a period of optimisation of their underlying sarcopenia.

The role of frailty in other fields of vascular surgery where L3 skeletal muscle or quadriceps muscle imaging is not available is also something that warrants consideration and further

investigation. Patients undergoing carotid intervention (carotid endarterectomy or stenting) for 'symptomatic' or 'asymptomatic' carotid stenosis is one such group that springs to mind where an alternative muscle group included in cross-sectional imaging of the neck (such as the sternocleidomastoid) may need consideration as a surrogate measure of sarcopenia.

Whilst the process of deriving SMI and SMD from cross-sectional CT image slices may have been automated by the sarcopenia segmentation tool, the true value to a clinician lies in ease of interpretation of these results. The development of an imaging-derived sarcopenia or frailty score and associated percentage risk of post-operative morbidity and mortality, which can be included in the report of every pre-operative CT scan of a patient undergoing AAA repair will further aid in the complex decision-making process, and not only benefit the clinician, but also the benefit the patient in improving their awareness of the underlying disease process and help them make an informed decision regarding their own health. Whilst we may be some way off from achieving this, this is what future research should possibly aspire to achieve.

In an increasingly ageing population, the percentage of patients that we encounter on a day-to-day basis with frailty and sarcopenia is only likely to increase. Patient expectations have also increased over time, and we are increasingly treating patients with more complex co-morbidities and pathologies, exposing them to additional overall risk. Identifying those patients at increased risk accurately, optimising their functional status, treating their underlying frailty and correcting / reversing sarcopenia is likely to be key in these patients. We hope that the utilisation of the results and discussion generated by this thesis will have a positive impact in this domain of clinical practice and the health of our patients.

REFERENCES

1. Nordon IM, Hinchliffe RJ, Loftus IM, Thompson MM. Pathophysiology and epidemiology of abdominal aortic aneurysms. *Nature Reviews Cardiology*. 2011. p. 92–102. https://doi.org/10.1038/nrcardio.2010.180.

2. Davis FM, Daugherty A, Lu HS. Updates of recent aortic aneurysm research. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2019. p. E83–E90. https://doi.org/10.1161/ATVBAHA.119.312000.

3. Kristensen KL, Dahl M, Rasmussen LM, Lindholt JS. Glycated hemoglobin is associated with the growth rate of abdominal aortic aneurysms a substudy from the VIVA (Viborg vascular) randomized screening trial. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2017;37(4): 730–736. https://doi.org/10.1161/ATVBAHA.116.308874.

4. Abdulhannan P, Russell DA, Homer-Vanniasinkam S. Peripheral arterial disease: A literature review. *British Medical Bulletin*. 2012. p. 21–39. https://doi.org/10.1093/bmb/lds027.

5. Aboyans V, Ricco JB, Bartelink MLEL, Björck M, Brodmann M, Cohnert T, et al. 2017 ESC Guidelines on the Diagnosis and Treatment of Peripheral Arterial Diseases, in collaboration with the European Society for Vascular Surgery (ESVS). *European Heart Journal*. 2018. p. 763–816. https://doi.org/10.1093/eurheartj/ehx095.

6. Powell JT, Brown LC, Forbes JF, Fowkes FGR, Greenhalgh RM, Ruckley C V., et al. Final 12-year follow-up of surgery versus surveillance in the UK Small Aneurysm Trial. *British Journal of Surgery*. 2007;94(6): 702–708. https://doi.org/10.1002/bjs.5778.

7. Kyriacou H, Mostafa AMHAM, Sumal AS, Hellawell HN, Boyle JR. Abdominal aortic aneurysms part two: Surgical management, postoperative complications and surveillance. *Journal of perioperative practice*. 2021;31(9): 319–325. https://doi.org/10.1177/1750458920947352.

8. Rank F, Ederle AL, Amuel S, Ilson EW, Ohnson ARJ, Einke OBR, et al. Immediate repair compared with surveillance of small abdominal aortic aneurysms. *The New England Journal of Medicine*. 2002;346(19): 1437-1444. https://doi.org/ 10.1056/NEJMoa012573.

9. Cao P, De Rango P, Verzini F, Parlani G, Romano L, Cieri E. Comparison of surveillance versus aortic endografting for small aneurysm repair (CAESAR): Results from a randomised trial. *European Journal of Vascular and Endovascular Surgery*. 2011;41(1): 13–25. https://doi.org/10.1016/j.ejvs.2010.08.026.

10. Patel R, Sweeting MJ, Powell JT, Greenhalgh RM. Endovascular versus open repair of abdominal aortic aneurysm in 15-years' follow-up of the UK endovascular aneurysm

repair trial 1 (EVAR trial 1): a randomised controlled trial. *The Lancet*. 2016;388(10058): 2366–2374. https://doi.org/10.1016/S0140-6736(16)31135-7.

11. Wanhainen A, Verzini F, Van Herzeele I, Allaire E, Bown M, Cohnert T, et al. Editor's Choice – European Society for Vascular Surgery (ESVS) 2019 Clinical Practice Guidelines on the Management of Abdominal Aorto-iliac Artery Aneurysms. *European Journal of Vascular and Endovascular Surgery*. 2019;57(1): 8–93. https://doi.org/10.1016/j.ejvs.2018.09.020.

12. Boden I, Skinner EH, Browning L, Reeve J, Anderson L, Hill C, et al. Preoperative physiotherapy for the prevention of respiratory complications after upper abdominal surgery: Pragmatic, double blinded, multicentre randomised controlled trial. *The BMJ*. 2018;360. https://doi.org/10.1136/bmj.j5916.

13. Drudi LM, Phung K, Ades M, Zuckerman J, Mullie L, Steinmetz OK, et al. Psoas Muscle Area Predicts All-Cause Mortality After Endovascular and Open Aortic Aneurysm Repair. *European Journal of Vascular and Endovascular Surgery*. 2016;52(6): 764–769. https://doi.org/10.1016/j.ejvs.2016.09.011.

14. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. *Journal of Gerontology*: *Medical Science*. 2001;56(3): 146-156. https://academic.oup.com/biomedgerontology/article/56/3/M146/545770

15. Cesari M, Leeuwenburgh C, Lauretani F, Onder G, Bandinelli S, Maraldi C, et al. Frailty syndrome and skeletal muscle: results from the Invecchiare in Chianti study. *The American Journal of Clinical Nutrition*. 2006;83(5):1142-1148. https://doi.org/10.1093/ajcn/83.5.1142

16. Buettner S, Wagner D, Kim Y, Margonis GA, Makary MA, Wilson A, et al. Inclusion of Sarcopenia Outperforms the Modified Frailty Index in Predicting 1-Year Mortality among 1,326 Patients Undergoing Gastrointestinal Surgery for a Malignant Indication. *Journal of the American College of Surgeons*. 2016;222(4): 397-407.e2. https://doi.org/10.1016/j.jamcollsurg.2015.12.020.

17. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2005;173(5): 489–495. https://doi.org/10.1503/cmaj.050051.

18. Santhirapala R, Fleisher LA, Grocott MPW. Choosing Wisely: just because we can, does it mean we should? *British journal of anaesthesia*. 2019;122(3): 306–310. https://doi.org/10.1016/j.bja.2018.11.025.

19. Cruz-Jentoft AJ, Sayer AA. Sarcopenia. *Lancet (London, England)*. 2019;393(10191): 2636–2646. https://doi.org/10.1016/S0140-6736(19)31138-9.

20. Peng P, Hyder O, Firoozmand A, Kneuertz P, Schulick RD, Huang D, et al. Impact of sarcopenia on outcomes following resection of pancreatic adenocarcinoma. *Journal of gastrointestinal surgery: official journal of the Society for Surgery of the Alimentary Tract*. 2012;16(8): 1478–1486. https://doi.org/10.1007/s11605-012-1923-5.

21. Waduud MA, Wood B, Keleabetswe P, Manning J, Linton E, Drozd M, et al. Influence of psoas muscle area on mortality following elective abdominal aortic aneurysm repair. *The British Journal of Surgery*. 2019;106(4): 367–374. https://doi.org/10.1002/bjs.11074.

22. Cheng BT, Soult MC, Helenowski IB, Rodriguez HE, Eskandari MK, Hoel AW. Sarcopenia predicts mortality and adverse outcomes after endovascular aneurysm repair and can be used to risk stratify patients. *Journal of vascular surgery*. 2019;70(5): 1576–1584. https://doi.org/10.1016/j.jvs.2018.12.038.

23. Newton DH, Kim C, Lee N, Wolfe L, Pfeifer J, Amendola M. Sarcopenia predicts poor long-term survival in patients undergoing endovascular aortic aneurysm repair. *Journal of vascular surgery*. 2018;67(2): 453–459. https://doi.org/10.1016/j.jvs.2017.06.092.

24. Thurston B, Pena GN, Howell S, Cowled P, Fitridge R. Low total psoas area as scored in the clinic setting independently predicts midterm mortality after endovascular

aneurysm repair in male patients. *Journal of vascular surgery*. 2018;67(2): 460–467. https://doi.org/10.1016/j.jvs.2017.06.085.

25. Matsubara Y, Matsumoto T, Aoyagi Y, Tanaka S, Okadome J, Morisaki K, et al. Sarcopenia is a prognostic factor for overall survival in patients with critical limb ischemia. *Journal of vascular surgery*. 2015;61(4): 945–950. https://doi.org/10.1016/j.jvs.2014.10.094.

26. Morisaki K, Furuyama T, Matsubara Y, Inoue K, Kurose S, Yoshino S, et al. Thigh sarcopenia and hypoalbuminemia predict impaired overall survival after infrainguinal revascularization in patients with critical limb ischemia. *Vascular*. 2020;28(5): 542–547. https://doi.org/10.1177/1708538120913745.

27. Aasen DM, Wiesen BM, Singh AB, Piper C, Harnke B, Prochazka A v., et al. Systematic Review of Preoperative Risk Discussion in Practice. *Journal of Surgical Education*. 2020;77(4): 911–920. https://doi.org/10.1016/J.JSURG.2020.02.008.

28. Hammermeister KE, Henderson WG, Bronsert MR, Juarez-Colunga E, Meguid RA. Bringing quantitative risk assessment closer to the patient and surgeon a novel approach to improve outcomes. *Annals of Surgery*. 2016;263(6): 1039–1041. https://doi.org/10.1097/SLA.0000000000001668.

29. Jones K, Gordon-Weeks A, Coleman C, Silva M. Radiologically Determined Sarcopenia Predicts Morbidity and Mortality Following Abdominal Surgery: A Systematic Review and Meta-Analysis. *World journal of surgery*. 2017;41(9): 2266–2279. https://doi.org/10.1007/S00268-017-3999-2.

30. Moskven E, Bourassa-Moreau É, Charest-Morin R, Flexman A, Street J. The impact of frailty and sarcopenia on postoperative outcomes in adult spine surgery. A systematic review of the literature. *The spine journal : official journal of the North American Spine Society*. 2018;18(12): 2354–2369. https://doi.org/10.1016/J.SPINEE.2018.07.008.

31. Levolger S, van Vugt JLA, de Bruin RWF, IJzermans JNM. Systematic review of sarcopenia in patients operated on for gastrointestinal and hepatopancreatobiliary malignancies. *The British journal of surgery*. 2015;102(12): 1448–1458. https://doi.org/10.1002/BJS.9893.

32. Pigneur F, Di Palma M, Raynard B, Guibal A, Cohen F, Daidj N, et al. Psoas muscle index is not representative of skeletal muscle index for evaluating cancer sarcopenia. *Journal of Cachexia, Sarcopenia and Muscle*. 2023;14(4): 1613-1620. https://doi.org/10.1002/jcsm.13230.

33. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: Revised European consensus on definition and diagnosis. *Age and Ageing*. 2019;48(4):16–31. https://doi.org/10.1093/ageing/afy169.

34. Ishii S, Tanaka T, Shibasaki K, Ouchi Y, Kikutani T, Higashiguchi T, et al. Development of a simple screening test for sarcopenia in older adults. *Geriatrics and Gerontology International*. 2014;14(SUPPL.1):93–101. https://doi.org/10.1111/ggi.12197.

35. Pigneur F, Di Palma M, Raynard B, Guibal A, Cohen F, Daidj N, et al. Psoas muscle index is not representative of skeletal muscle index for evaluating cancer sarcopenia. *Journal of Cachexia, Sarcopenia and Muscle*. 2023;14(4):1613-1620. https://doi.org/10.1002/jcsm.13230.

36. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372. https://doi.org/10.1136/BMJ.N71.

37. *PROSPERO*. https://www.crd.york.ac.uk/prospero/ [Accessed 5th May 2022].

38. Veritas Health Innovation MA. *Covidence systematic review software*. Available at www.covidence.org.

39. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. *PLOS Medicine*. 2014;11(10): e1001744. https://doi.org/10.1371/JOURNAL.PMED.1001744.

40. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Annals of Internal Medicine*. 2013;158(4): 280–286. https://doi.org/10.7326/0003-4819-158-4-201302190-00009.

41. The Cochrane Collaboration. *Review Manager Web (RevMan Web)*. 2015.

42. Mark Mitchell, Baurzhan Muftakhidinov, Tobias Winchen. *Engauge Digitizer Software.* http://markummitchell.github.io/engauge-digitizer/ [Accessed 12th May 2022].

43. Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4 (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook

44. Bang JY, Jun IG, Lee JB, Ko Y, Kim KW, Jeong J, et al. Impact of Sarcopenia on Acute Kidney Injury after Infrarenal Abdominal Aortic Aneurysm Surgery: A Propensity Matching Analysis. *Nutrients*. 2021;13(7): 2212. https://doi.org/10.3390/NU13072212.

45. Cheng BT, Soult MC, Helenowski IB, Rodriguez HE, Eskandari MK, Hoel AW. Sarcopenia predicts mortality and adverse outcomes after endovascular aneurysm repair and can be used to risk stratify patients. *Journal of Vascular Surgery*. 2019;70(5): 1576–1584. https://doi.org/10.1016/J.JVS.2018.12.038.

46. Kim YK, Lee HS, Ryu JJ, In Lee H, Seo SG. Sarcopenia increases the risk for mortality in patients who undergo amputation for diabetic foot. *Journal of Foot and Ankle Research*. 2018;11(1): 1–8. https://doi.org/10.1186/S13047-018-0274-1/FIGURES/3.

47. Koter S, Cohnert TU, Hindermayr KB, Lindenmann J, Brückner M, Oswald WK, et al. Increased hospital costs are associated with low skeletal muscle mass in patients undergoing elective open aortic surgery. *Journal of Vascular Surgery*. 2019;69(4): 1227–1232. https://doi.org/10.1016/J.JVS.2018.06.224.

48. Matsubara Y, Matsumoto T, Inoue K, Matsuda D, Yoshiga R, Yoshiya K, et al. Sarcopenia is a risk factor for cardiovascular events experienced by patients with

critical limb ischemia. *Journal of Vascular Surgery*. 2017;65(5): 1390–1397. https://doi.org/10.1016/J.JVS.2016.09.030.

49. Selçuk N, Albeyoğlu Ş, Bastopcu M, Selçuk İ, Barutca H, Şahan H. Sarcopenia is a risk factor for major adverse cardiac events after surgical revascularization for critical limb ischemia. *Vascular*. 2023;31(1):64-71. https://doi.org/10.1177/17085381211059383.

50. Vedder IR, Levolger S, Dierckx RAJO, Zeebregts CJ, de Vries JPPM, Viddeleer AR, et al. Effect of muscle depletion on survival in peripheral arterial occlusive disease: Quality over quantity. *Journal of Vascular Surgery*. 2020;72(6): 2006-2016.e1. https://doi.org/10.1016/J.JVS.2020.03.050.

51. Indrakusuma R, Zijlmans JL, Jalalzadeh H, Planken RN, Balm R, Koelemay MJW. Psoas Muscle Area as a Prognostic Factor for Survival in Patients with an Asymptomatic Infrarenal Abdominal Aortic Aneurysm: A Retrospective Cohort Study. *European Journal of Vascular and Endovascular Surgery*. 2018;55(1): 83–91. https://doi.org/10.1016/J.EJVS.2017.10.007/ATTACHMENT/52E473D1-1ACE-4D4C-A649-AF0AE7EAD41C/MMC1.DOC.

52. Jones KI, Doleman B, Scott S, Lund JN, Williams JP. Simple psoas cross-sectional area measurement is a quick and easy method to assess sarcopenia and predicts major

surgical complications. *Colorectal Disease*. 2015;17(1):20–26. https://doi.org/10.1111/CODI.12805.

53. Antoniou GA, Rojoa D, Antoniou SA, Alfahad A, Torella F, Juszczak MT. Effect of Low Skeletal Muscle Mass on Post-operative Survival of Patients With Abdominal Aortic Aneurysm: A Prognostic Factor Review and Meta-Analysis of Time-to-Event Data. *European Journal of Vascular and Endovascular Surgery*. 2019;58(2): 190–198. https://doi.org/10.1016/J.EJVS.2019.03.020/ATTACHMENT/0C556752-DC18-4E13-9BDB-6A37EE6DAD78/MMC4.DOCX.

54. Sharpe M, Okoye E, Antoniou GA. Prognostic review and time-to-event data meta-analysis of low skeletal muscle mass in patients with peripheral arterial disease of the lower limbs undergoing revascularization. *International Angiology*. 2020;39(1): 50–59. https://doi.org/10.23736/S0392-9590.19.04248-2.

55. Islam S, Kanavati F, Arain Z, Da Costa OF, Crum W, Aboagye EO, et al. Fully automated deep-learning section-based muscle segmentation from CT images for sarcopenia assessment. *Clinical Radiology*. 2022;77(5): e363–e371. https://doi.org/10.1016/j.crad.2022.01.036.

56. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for

providing translational research informatics support. *Journal of biomedical informatics*. 2009;42(2): 377–381. https://doi.org/10.1016/j.jbi.2008.08.010.

57. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*. 2019;95: 103208. https://doi.org/10.1016/j.jbi.2019.103208.

58. Dean Bidgood W, Horii SC, Prior FW, Van Syckle DE, Bidgood WD. Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging. *Journal of the American Medical Informatics Association*. 1997;4(3): 199-212. https://doi.org/ 10.1136/jamia.1997.0040199.

59. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*. 2006;31(3): 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015.

60. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiology. Artificial intelligence*. 2020;2(2): e200029. https://doi.org/10.1148/ryai.2020200029.

61. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. *Surgery*. 1992;111(5): 518–526.

62. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Annals of surgery*. 2004;240(2): 205–213. https://doi.org/10.1097/01.sla.0000133083.54934.ae.

63. Strasberg SM, Linehan DC, Hawkins WG. The accordion severity grading system of surgical complications. *Annals of Surgery*. 2009;250(2): 177–186. https://doi.org/10.1097/SLA.0b013e3181afde41.

64. Waton S, Johal A, Birmpili P, Li Q, Atkins E, Cromwell DA, et al. *National Vascular Registry: 2022 Annual Report*. 2022 Nov.

65. Takeji Y, Yamaji K, Tomoi Y, Okazaki J, Tanaka K, Nagae A, et al. Impact of Frailty on Clinical Outcomes in Patients With Critical Limb Ischemia. *Circulation. Cardiovascular interventions*. 2018;11(7): e006778. https://doi.org/10.1161/CIRCINTERVENTIONS.118.006778.

66. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*. 2016;15(2): 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

APPENDIX 1 – ETHICS APPROVAL

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Professor Alun Davies
Professor of Vascular Surgery and Consultant Surgeon
Imperial College London
Room 3, 4th Floor East Wing
Charing Cross Hospital
London
W68RF

Email: approvals@hra.nhs.uk

06 May 2021

Dear Professor Davies

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**The use of an automated Sarcopenia measurement tool as a Predictor of outcome in patients with Abdominal aortic aneuRysms (AAA)**
IRAS project ID:	**285472**
Protocol number:	**21SM6746**
Sponsor	**Imperial College London**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?
The *"After HRA Approval – guidance for sponsors and investigators"* document on the HRA website gives detailed guidance on reporting expectations for studies with HRA and HCRW Approval, including:

- Registration of Research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **285472**. Please quote this on all correspondence.

Yours sincerely,

Sarah Prothero
Approvals Specialist

Email: approvals@hra.nhs.uk

Copy to: *Mr Keith Boland*