An investigation into the measurement and management of frailty in surgical patients

Mr John Martin Trotter
MBChB MRCS

MD Thesis

The University of Hull &
The University of York

Hull York Medical School

November 2019

Abstract

Aims

This thesis intends to investigate a method of identification of frailty in the surgical population, CT defined sarcopenia, and a possible method to attenuate its effects in the preoperative period, prehabilitation.

Methods

845 patients that underwent emergency laparotomy in 4 acute hospitals were screened for sarcopenia by review of CT scans assessing sarcopenia by psoas density (PD) and area (PA). Primary outcomes were 30 day and 1 year mortality.
A pilot RCT was undertaken to assess the acceptability and achievability of walking-based prehabilitation monitored by wearable technology. Participants were randomised to either normal activity or a walking based exercise programme.

Results

Sarcopenia measured by PD was associated with increased mortality compared to non-sarcopenic patients at 30-days (23.2% vs. 9.6% p<0.0001 OR=2.84 (95% CI 1.88-4.30) and 1-year 37% vs. 19.2% p<0.0001 OR=2.46 (95% CI 1.75-3.47). Increased mortality was seen in sarcopenic patients measured by PA at 30-days (16.3% vs. 7.8% p=0.001 OR=2.31 (95% CI 1.38-3.88) and 1-year 32% vs. 18.7% p=<0.0001 OR=2.25 (95% CI 1.52-3.34)
For the RCT 45 patients were approached to recruit 40 participants. The median time in study was 12.5days (IQR 6-18). Mean compliance to the exercise programme was 58%. Mean distance change between initial and pre-operative assessment for the exercise and normal-activity groups was +16.4m and -13.6m respectively, p=0.013. Mean distance change between initial and 3-month postoperative assessment was -11.4m and -40m p=0.11.

Conclusion

Sarcopenia assessed by PD and PA on CT is associated with increased mortality following emergency laparotomy. The use of sarcopenia as a predictive tool may be useful to direct geriatric input and guide expectations in emergency surgery.
This pilot study confirms that acceptable compliance can be achieved using a user-friendly pedometer and that this is associated with measurable improvements in fitness. Further work is required to establish whether this translates into improved patient outcomes after surgery.

List of Contents

List Of Figures

List Of Tables

List of Abbreviations

ADP	Adenosine Triphosphate
ATP	Adenosine Biphosphate
ANOVA	Analysis of Variance
ASA	American Society of Anaesthesiologists
BMI	Body Mass Index
BPM	Beats Per Minute
CGA	Comprehensive Geriatric Assessment
CI	Confidence Interval
CPET	Cardio-Pulmonary Exercise Testing
CT	Computed Tomography
ERAS	Enhanced Recovery After Surgery
FFMI	Fat Free Mass Index
FMI	Fat Mass Index
GPS	Global Positioning System
HR	Hazard Ratio
ICD-10-CM	International Classification of Diseases, Tenth Revision, Clinical Modification
IMAC	Intra-Muscular Adipose Content
IQR	Interquartile Range
L3	3rd Lumbar Vertebrae
L4	4th Lumbar Vertebrae
MDT	Multidisciplinary Team
MRI	Magnetic Resonance Imaging
NBOCA	National Bowel Cancer Audit
NACRT	Neoadjuvant Chemo-Radiotherapy
NELA	National Emergency Laparotomy Audit
NIL	Non-Independent Living
OR	Odds Ratio
OS	Overall Survival
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QOL	Quality Of Life
RCT	Randomised Control Trial
RFS	Recurrence Free Survival
ROC	Receiver Operating Characteristic
RRF	Retro-Renal Fat
SSI	Surgical Site Infection
SD	Standard Deviation
T-Stage	Tumour Stage
T12	12^{th} Thorasic Vertebrae
TMA	Total Muscle Area
WHO	World Health Organisation
6MWT	6 Minute Walk Test

Acknowledgments

I would like to thank my supervisors, Ms Clare McNaught, Professor John MacFie and Mr John Hartley for the opportunity they have given me to complete this MD project and for all the supervision and guidance along the way. In particular the drive and impetus that Ms McNaught has provided, who has become a mentor and friend. I would like to thank Mr Marcel Gatt and Ms Karen Maude for their help with recruitment of patients and review of work that has been presented and published so far. I would also like to thank the Ms Judith Johnson, Mr Alvin Ng and Mr Philip Herrod who have collected data for me on various sites. I would like to thank my wife, Leanne, for not protesting and having patience whilst I have moved 100 miles away to complete this MD, and finally my dog, Branston, for sacrificing some of his walks.

Authors Declaration

I confirm that this work is original and that if any passage(s) or diagram(s) have been copied from academic papers, books, the internet or any other sources these are clearly identified by the use of quotation marks and the reference(s) is fully cited. I certify that, other than where indicated, this is my own work and does not breach the regulations of HYMS, the University of Hull or the University of York regarding plagiarism or academic conduct in examinations. I have read the HYMS Code of Practice on Academic Misconduct, and state that this piece of work is my own and does not contain any unacknowledged work from any other sources'. If applicable, the declaration should also include; 'I confirm that any patient information obtained to produce this piece of work has been appropriately anonymised'.

Presentations, Publications, Abstracts & Prizes from this Thesis

Publications

TROTTER, J., JOHNSTON, J., NG, A., GATT, M., MACFIE, J. & MCNAUGHT, C. 2018. Is sarcopenia a useful predictor of outcome in patients after emergency laparotomy? A study using the NELA database. *Annals of the Royal College of Surgeons of England,* 100**,** 377-381.

TROTTER, J. MCNAUGHT C. 2018. The Failty Epidemic. *Surgeons News*, March 2018 32-33

HERROD, J. BOYD CARSON, H. DOLEMAN, B. TROTTER, J. SCHLICHTEMEIER, S. SATHANAPALLY, G. SOMMERVILLE, J. WILLIAMS, J. LUND, J. 2019 Quick and simple; psoas density measurement is an independent predictor of anastomotic leak and other complications after colorectal resection. Tech Coloproctol, 23(2):129-134

Abstracts

2018 Narang, K. Trotter, M. Gatt, M. McNaught, C. Variability in Frailty Assessment. *BJS 2018; 105 (S5): 113–231*

2018 Trotter, M. Herrod, P. Johnston, J. Gatt, M. McNaught, C. Effect of Frailty Assessed by Sarcopenia in Emergency Laparotomy. *BJS 2018; 105 (S5): 12–66*

2017 Trotter, M. Ng, A. Johnston, J. Gatt, M. MacFie, J. McNaught, C. Outcomes after Emergency Laparotomy: Should Frailty be Considered Alongside P-POSSUM? *JACS 2017*; 225 (4, S1): 86

2017 Ng, A. Trotter, J. Johnston, J. Gatt, M. McNaught, C. Assessment of frailty by CT-determined sarcopenia in patients undergoing emergency laparotomy: Is one way better than the other? BJS 2017; 104 (S6): 83–243

2017 Trotter, J. Johnston, J. Ng, A. Gatt, M. McNaught, C. Outcomes following Emergency Laparotomy: Should Frailty beConsidered Alongside P-Possum? *BJS 2017*;104 (S6):8–9

Presentations

2018 Oral Presentation at Leeds Regional Surgical Club

2018 Oral Presentation at ASGBI International Surgical Conference Liverpool – Effect of Frailty Assessed by Sarcopenia in Emergency Lapaorotmy

2017 Oral Presentation at American Collage of Surgeons Clinical Congress San Diego – Frailty Assessment in Emergency Laparotomy

2017 NELA Prize Presentation at ASGBI International Surgical Congress Glasgow - Outcomes following Emergency Laparotomy: Should Frailty be Considered Alongside P-Possum?

2017 Poster Presentation at ASGBI International Surgical Congress, Glasgow - Assessment of frailty by CT-determined sarcopenia in patients undergoing emergency laparotomy: Is one way better than the other?

Prizes

2018 John Macfie Medal at Leeds Regional Surgical Club

2017 NELA Prize at Association of Surgeons Great Britain and Ireland International Surgical Congress - Glasgow

Thesis Introduction

Frailty is a widely used term that does not have an agreed clinical definition (Rodriguez-Manas et al., 2013) and, as such, does not appear in the WHO international statistical classification of diseases. Despite this “frailty”, however measured, has been used to predict adverse events and outcome (Fried et al., 2001, Rockwood et al., 2004, Bandeen-Roche et al., 2006), particularly in elderly patients. Examples of definitions of frailty include “a biologic syndrome of decreased reserve and resistance to stressors, resulting from cumulative declines across multiple physiologic systems, and causing vulnerability to adverse outcomes”(Fried et al., 2001), or “ a state of increased vulnerability to poor resolution of homoeostasis after a stressor event, which increases the risk of adverse outcomes, including falls, delirium, and disability”(Clegg et al., 2013).

Over the last 20 years much work has been done into improving surgical care by optimizing a patient’s recovery through multimodal care (McNaught and MacFie, 2002, Gatt et al., 2005, Bardram et al., 1995). Pre-operative investigations such as cardio-pulmonary exercise testing, is deployed to identify those less well adapted for surgery but can be specific and lacks the global assessment that frailty may give. If it is agreed that frailty of any definition is of interest in the surgical population then its identification and management is key. It has been shown (Theou et al., 2011, Jose Tarazona-Santabalbina et al., 2016) that it is possible to reverse some of the effects of frailty with exercise, but this is often far more difficult than with the non-frail population. It has also been demonstrated (McLennan et al., 2019) that despite the most modern multimodal care or “enhanced recovery” packages poor exercise capacity is one of the strongest determinates of poor outcome following major surgical intervention.

This thesis intends to investigate a method of identification of frailty in the surgical population, CT defined sarcopenia, and a possible method to attenuate its effects in the preoperative period, prehabilitation.

The Acceptability and Achievability of Walking Prehabilitation Prior to Resectional Bowel Surgery

1. Introduction

1.1 Current Practice

Traditionally patients undergoing major intestinal resection remained in hospital for up to two weeks during which time their oral intake was restricted, mobility was impaired and analgesia reliant on parenteral opioids. In the late 1990s a number of investigators, most notably Kehlet's group from Denmark, developed the concept of ‘fast track’ surgery (Kehlet and Wilmore, 2005, Kehlet and Wilmore, 2008). Their philosophy was to employ a combination of epidural or spinal anaesthesia with early mobilisation and oral feeding on the basis that these, and other interventions, would reduce the stress response to surgery and enhance recovery after surgery. The results of many largely observational studies seemed to confirm the benefits of such a ‘fast track’ approach (Kehlet and Wilmore, 2005, Kehlet and Wilmore, 2008). These principles were further developed by adopting the concept of ‘multimodal optimisation’ of perioperative care (Gatt et al., 2005, Anderson et al., 2003). In the last decade a number of studies have been reported all confirming the benefits of such a multimodal approach to perioperative care. The term ‘enhanced recovery after surgery’ (ERAS) is now most commonly used to describe this modern multimodal approach to surgical management (Fearon et al., 2005).

Modern surgery is becoming ever more complex which inevitably tests the physiological reserve of the patient. This includes the surgical procedure itself and adjuvant therapies. Twinned with this increase in physiological insult, is the trend for operating on ever more elderly and frail patients who have a larger co-morbid load and are less physically fit (Pearse et al., 2006). The 2017 National Bowel Cancer Audit (NBOCA) report found that 69% and 37% of patients undergoing a major resection for colonic or rectal cancer were over 65 and 75 respectively (NBOCA, 2017).

1.2 Exercise and its benefits

Exercise can be defined as "activity requiring physical effort, carried out to sustain or improve health and fitness". Exercise has been predominantly grouped into either aerobic or anaerobic. This refers to formation of the main unit of energy adenosine triphosphate (ATP) from adenosine biphosphate (ADP) and if oxygen is used for this reaction or not. Traditional thinking maintained that in times of high intensity exercise the demand for oxygen for this reaction would outstrip supply, and ATP would be formed without oxygen in an "anaerobic" reaction. This form of ATP synthesis creates lactic acid and is not sustainable for long periods of time. Logic was that this anaerobic reaction was used for short high intensity bouts of exercise for less intense sustained effort aerobic formation of ATP took place. However it has been demonstrated that this is over simplistic and separating exercise into anaerobic and aerobic is not helpful. There are a number of different pathways used for ATP synthesis that are employed dependant on the length of effort, its intensity or muscular effort. It is more useful to group exercise into; explosive efforts of <6s; high intensity efforts of >6s to 1 min; endurance intensive efforts >1min. This better categorises the ability of an individual to sustain a certain level of effort based on the underlying physiology and avoids

the misleading anaerobic versus aerobic division of exercise (Chamari, 2015).

1.3 Levels of physical fitness and methods to increase it

Advise on how much exercise or physical activity is needed to maintain health can be very varied. However the World Health Organisation (WHO) provides guidance on physical activity levels for adults in 4 key recommendations (WHO, 2010):

1. Adults aged 18–64 should do at least 150 minutes of moderate-intensity aerobic physical activity throughout the week or do at least 75 minutes of vigorous-intensity aerobic physical activity throughout the week or an equivalent combination of moderate and vigorous intensity activity.
2. Aerobic activity should be performed in bouts of at least 10 minutes duration.
3. For additional health benefits, adults should increase their moderate-intensity aerobic physical activity to 300 minutes per week, or engage in 150 minutes of vigorous-intensity aerobic physical activity per week, or an equivalent combination of moderate and vigorous intensity activity.
4. Muscle-strengthening activities should be done involving major muscle groups on 2 or more days a week

This guidance was written to provide basic information for campaigns or initiatives seeking to increase fitness levels. The activity levels set out above are not altered in the WHO guidance for the elderly or frail only adding that “When older adults cannot do the recommended amounts of physical activity due to health conditions, they should be

as physically active as their abilities allow". No specific exercise form or regime is suggested, just the act of being physically active. Examples of moderate exercise are given as walking, riding a bike, doubles tennis or anything that means you can still talk but cannot sing the words to a song. Examples of vigorous exercise are given as swimming fast, singles tennis, jogging, riding a bike uphill or not being able to say more than a few words without pausing for breath. Although one activity can be considered to be of differing intensity, varying with the underling fitness of the individual. As such no particular exercise is recommended. The recommendations state periods of physical activity, aiming to be less prescriptive and variable to each individual's abilities.

In the UK the Sport England Actives Lives Survey (Sport-England, 2017a) from 2016-17 sampled 198,000 adults and found those reaching the WHO targets to be 60%. However, this falls to 57% in the 55-74's, 37% in the 75-84's and 18% in the over 85's. Levels of inactivity (less than 30 minutes/week of moderate activity) were found to be 30% in the 55-74's, 48% in the 75-84's and 71% in the over 85's. Clearly these levels show activity significantly below that outlined in the WHO guidance, especially in the age groups who are most likely to have major abdominal surgery

1.4 Concept of prehabilitation

In response to this, increasing surgical complexity, age and co-morbidity of patients undergoing elective surgery, research into extending multimodal optimisation into the preoperative period is growing. This has been termed 'prehabilitation' and entails increasing a patient's fitness to create a larger physiological buffer to the upcoming surgical insult. It is logical that this physiological buffer that increased fitness creates, may enable patients to better withstand the surgical insult perhaps leading to improved outcome.

It has been demonstrated that the level of a person's physical fitness prior to surgery appears to be an important predictor of a patient's postoperative recovery (Levett and Grocott, 2015).

1.5 Published work on prehabilitation

The guidance on physical activity (WHO, 2010) recommends only durations of moderate and vigorous activity rather than specific types of activity. As such studies into prehabilitation have used varied interventions to increase fitness levels. In 2018 (Luther et al., 2018) a systematic review was published of studies where patients undergoing major abdominal surgery had been randomised to differing forms of prehabilitation. Following a search that revealed 2236 non-duplicate articles 16 studies from 2000 to 2018 were included in the review. 11 of these studies included a group randomised to differing types of exercise-based prehabilitation. In 2019 (Hughes et al., 2019) a meta-analysis was published that included trials where patients undergoing major abdominal surgery had been randomised to differing forms of prehabilitation. Following a search that revealed 490 non-duplicate articles from 1966 to 2017 15 studies were included in the meta-analysis. Of the 15 studies included in the meta-analysis 10 were not included in the 2018 systematic review. Table 1 displays the methodology and results of these 21 studies.

Study	Home or Hospital/Gym	Supervised or Unsupervised	Type of Exercise	Number in Trial	Results
(Dronkers et al., 2008)	Hospital/Gym	Supervised	Inspiratory muscle training	20	No effect
(Dronkers et al., 2010)	Hospital/Gym	Supervised	High intensity aerobic exercise	42	Improved respiratory function
(Minnella et al., 2017)	Hospital/Gym	Supervised	Aerobic, resistance and endurance	186	Exercise group returned to baseline fitness faster
(van Adrichem et al., 2014)	Hospital/Gym	Supervised	High intensity inspiratory muscle training	62	Reduced length of stay on pulmonary complication with exercise
(Inoue et al., 2013)	Hospital/Gym	Supervised	Muscle and aerobic training	100	Less pulmonary complication with exercise
(Dettling et al., 2013)	Hospital/Gym	Supervised	Inspiratory muscle training	83	No effect
(Cho et al., 2014)	Hospital/Gym	Supervised	Aerobic, resistance and stretching exercise	72	Less severe complication with exercise
(de Toledo Piza Soares et al., 2013)	Both	Supervised	Physical therapy sessions	32	Less pulmonary complication with exercise
(Carli et al., 2010)	Home	Unsupervised	Exercise bike and muscle training vs. Walking	112	Walking group demonstrated larger fitness gain
(Barakat et al., 2016)	Hospital/Gym	Supervised	Aerobic exercise programme	124	Lower complication with exercise
(Barberan-Garcia et al.,	Hospital/Gym	Supervised	High intensity endurance	125	Lower complication

2018)			training		with exercise
(Kim et al., 2009)	Home	Unsupervised	Exercise Bike	21	No effect
(Barbalho-Moulim et al., 2011)	Hospital/Gym	Supervised	Inspiratory muscle training	32	Increased inspiratory strength
(Dunne et al., 2016)	Hospital/Gym	Supervised	High intensity interval cycle	38	Improved CPET values and quality of life scores.
(Gillis et al., 2014)	Home	Unsupervised	Aerobic and resistance	77	Increased 6MWT distance
(Jensen et al., 2015)	Home	Unsupervised	Step trainer, muscle strength and endurance	107	No effect
(Kaibori et al., 2013)	Home	Unsupervised	Walking	51	Improved insulin resistance
(Kulkarni et al., 2010)	Home	Unsupervised	Inspiratory muscle training	80	Improved maximum inspiratory pressure
(Llorens et al., 2015)	Home	Unsupervised	Inspiratory muscle training	44	Improved postoperative oxygenation
(de Toledo Piza Soares et al., 2013)	Hospital/Gym	Supervised	Stretching and inspiratory muscle training	32	Lower postoperative pulmonary complications
(Llorens et al., 2015, Tew et al., 2017)	Hospital/Gym	Supervised	High intensity interval training (HIIT)	53	HIIT feasible in pre abdominal aortic aneurysm repair

Table 1: Methodology of prehabilitation studies including major abdominal surgery.

The meta-analysis by Hughes analysed common endpoints for the 457 prehabilitation and 450 control patients. Endpoints analysed

were length of stay, pulmonary morbidity and 6 minute walk test distance. The only significant difference observed was a reduction in pulmonary morbidity.

Each of these 21 studies used a different exercise programme that varied in intensity duration and type. Some were tailored to a specific endpoint such as inspiratory muscle training that was used as an attempt to affect an endpoint of pulmonary complications (Dronkers et al., 2008, van Adrichem et al., 2014, Llorens et al., 2015, Kaibori et al., 2013, Dettling et al., 2013). The rest of the studies used differing exercise interventions to attempt to achieve a general increase in fitness. 13 of these 21 studies used hospital or gym based exercise programmes with 7 home-based and 1 mixed. This heterogeneity highlights a complete lack of consensus on the best form of exercise to use in prehabilitation. Practically one would assume that high intensity, frequent and supervised programmes would have the best results. However it may not reasonable to ask a cohort of patients who are likely to be relatively inactive (Sport-England, 2017b) to train like professional athletes. This issue was approached by (Carli et al., 2010) where participants were randomised to lower intensity walking based exercise for 30 minutes a day or a bike and strengthening regime. The bike and strengthening regime were instructed to "exercise initially at 50 per cent of their maximal heart rate; this was increased by 10 per cent each week, if tolerable. Weight training was to be carried out three times a week, to avoid muscle soreness. Patients were instructed to do push-ups, sit-ups and standing strides (lunges) until volitional fatigue, increasing this number to reach 12 repetitions". Surprisingly it was the walking group that produced a larger fitness gain. This demonstrates the importance of the achievability of an exercise programme. An exercise programme may be perfectly designed to increase fitness but if its too intensive for the participants it may have less effect than a more achievable

programme, which in theory has less potential to deliver fitness gains.

This was observed by (Barakat et al., 2016) where, in the prehabilitation arm, primary outcomes of length of stay and complication rates were significantly reduced. The prehabilitation regime consisted of supervised one-hour exercise sessions three times per week at a hospital gym. Published results were broken down by attendance and the decrease in complication rate was only found in the 1/3 that attended the exercise programme for all sessions. The 2/3 that did not completely adhere to the programme saw no change in outcome compared with the control group. It would appear that designing a programme that is both achievable and acceptable to those who are to undertake it, is as important as the form of the exercise itself. (Ferreira et al., 2018) surveyed 52 cancer patients enrolled in a prehabilitation programme and found that their preferred mode of prehabilitation would be home based and the largest barrier to participation was transportation. Although this is not a randomised control trial it is a very important observation for prehabilitation. This highlights that the supervision of an exercise programme is an issue and where patients have to travel this can reduce compliance or recruitment into the programme. A trained individual can travel to the patient's home for supervision but this greatly increases cost and reduces the use of a programme outside a research setting.

The published work on prehabilitation shows mixed results on the benefits to patients from prehabilitation exercise. What does seem to be clear is that the choice of exercise should not be based entirely on its potential to deliver an increase in fitness. Patients are on the whole inactive and unwilling to travel and, as such, may not engage with a programme that is not designed with this in mind. Therefore designing a programme that is acceptable and achievable may be

the key to a successful prehabilitation programme that delivers the desired results.

Adherence to an exercise programme is an issue for those agreeing to take part, but a larger issue may be recruitment itself. Many studies do not publish their recruitment rate and those that do show varied results. Some that have published recruitment rates (Segal et al., 2001, Segal et al., 2003, Courneya et al., 2003a, Courneya et al., 2003b) to exercise regimes have quoted poor recruitment of around 40% of those approached. Other studies have demonstrated much higher recruitment (Jones et al., 2004, Mock et al., 2005), up to 94%. The exercise interventions for these studies did not include travel, active supervision or the introduction of an activity that participants were not already doing, i.e. walking. This was likely the reason for improved recruitment rates and is useful to note if prehabilitation is to be used across the healthcare setting outside the confines of a trial.

The systemic review of (Luther et al., 2018) and systemic review and meta-analysis of (Hughes et al., 2019) both comment on the lack of any consesus on any particular type of prehabilitation with huge hetrogeneity in the literature. Hughes recomends that prehabilitaion is of benefit and recomends its use, however Luther is more speculative concluding that there is not yet enough data to make a recommendation for prehabilitation in major abdominal procedures.

1.6 Measurement of fitness

Any programme that seeks to increase fitness must have a method to measure fitness in order to demonstrate results.

Heart rate

The simplest measure of fitness may be heart rate. Various heart rate indexes can be used as fitness measures. The most commonly used are resting heart rate, heart rate response to exercise and heart rate recovery post exercise. The change in heart rate with exertion is termed the chronotropic response. A normal chronotropic response to exercise is a rapid rise in heart rate mediated by a fall in parasympathetic tone and increase in sympathetic tone. The increase in heart rate is to compensate for the increased metabolic activity and oxygen demands. Oxygen uptake (VO_2) during maximal exercise increases by up to 4.4 times (VO_2max) and the largest contributor to this is heart rate increase. Ability to increase heart rate is vital to fitness. As each year passes from early adulthood the maximum heart rate reachable falls. The equation commonly used is 220-age however actual numbers are very individual and subject to variance. The inability for heart rate to respond to an increase in metabolic need is termed chronotropic incompetence.

Chronotropic incompetence is multifactorial. Changes in tone of the sympathetic and parasympathetic nervous supply, pathology of the intrinsic rate control system of the heart (in particular the sino-atrial node) and the effects of ischaemic heart disease all play a part (Brubaker and Kitzman, 2011). Chronotropic incompetence impairs an individual's ability to increase VO_2 and meet metabolic demands, thus lowering VO_2 max and fitness (Routledge and Townend, 2006). Chronotropic incompetence has been demonstrated to be a strong risk factor of early sudden cardiac death. As an individual becomes fitter the intensity and duration of exercise they are able undertake increases due to many changes. Lung capacity and efficiency of respiration, muscle volume and strength, and cardiac output in terms of stroke volume and ventricular function, amongst others, all improve. This makes the interpretation of heart rate as an indicator of fitness difficult unless an individual is being tested to VO_2 max. A certain heart rate response to walking may be low because of chronotropic incompetence or because the individual is fit and all

systems are very efficient, and thus only a small increase in heart rate is required. Heart rate recovery is a more reliable measure of fitness. As exercise finishes and an individual rests the heart rate will fall. The fittest athletes demonstrate a rapid fall in heart rate in the first 30 seconds post exercise before decline in heart rate falls at a steady rate. The least fit of individuals will have a lower maximum heart rate and will exhibit a slower fall back to resting rate. This makes rate of heart rate recovery a more reliable marker of fitness when an individual is not being pushed to VO_2 max.

Resting heart rate can also be a guide to fitness. An athlete will tend to have a similar metabolic need to less fit individuals at rest. As the general efficiency of their cardiovascular system to deliver oxygen is improved their heart rate can be lower and still deliver the oxygen requirements needed by various systems. When measuring fitness of an elderly population with various co-morbid conditions confounding factors may be met when measuring fitness with heart rate monitoring. Confounders such as osteoarthritis, medication that affects heart rate or vasoconstriction or lung pathology that effects ventilation may all vastly alter the heart rate recorded when exercising.

Cardiopulmonary exercise testing

Cardiopulmonary exercise testing (CPET) was originally used to measure fitness and performance of athletes but its use has become widespread in healthcare. In medical settings the derived parameters are used for investigation for cardiopulmonary disease (Balady et al., 2010) and for risk stratification before major surgical intervention (Older et al., 1993). CPET protocols aim to push the tested individual to their limit whilst connected to apparatus that is able to sense pulmonary performance, expired and inspired gasses as well as cardiac parameters and electrical activity. Exercise is most commonly done on a treadmill or bike. VO_2 max is calculated as well

as the point where metabolic demands outstrip oxygen supply and anaerobic pathways of glycolysis for formation of ATP are used. This is often termed the ventilatory threshold or anaerobic threshold and has been used to predict poor outcome in major surgical procedures. Changes in electrocardiogram (ECG) traces can also give information about exercise-induced ischaemia. CPET measures underlying cardiorespiratory fitness but its large amount of data can be further used to investigate and diagnose cardiorespiratory conditions.

Over the last 20 years the use of CPET in pre-operative assessment has grown. Various parameters, especially anaerobic threshold, have been used to predict poor outcome in terms of morbidity and mortality (Snowden et al., 2010, Prentis et al., 2012) and even 5-year survival (Colson et al., 2012). Concerns have grown however about possible over confidence in CPET results and the use of them for clinical decision making in the pre-operative environment. These concerns relate to evidence base for risk stratification in different surgical procedures, inter-operator variance, standardisation of method and cost analysis (Bramley and Brown, 2018). Simple fitness testing, where the data generated by CPET is not needed, may be better undertaken with simple tests, especially given the concordance for some simple tests such as 6MWT (Shulman et al., 2019) and activity questionnaires (Dumbrill et al., 2019) with CPET results.

Exercise testing without maximal effort

The determination of an individual's fitness need not necessarily entail pushing them to their VO_2 max or point of failure. In fact one could argue that in all but the healthy of subjects this may include risk. When testing the elderly or those with significant comorbidity this certainly requires presence of highly trained personnel in a situation were cardiorespiratory resuscitation can take place

immediately if needed. The most wildly used and validated sub-maximal exertion test in healthcare settings is the 6MWT. It has been found to be reproducible (Butland et al., 1982) and is used to assess function in the elderly (Harada et al., 1999), predict cardiopulmonary morbidity and mortality (Bittner et al., 1993, Lipkin et al., 1986), as well as negative outcomes from other conditions (McDonald et al., 2010). The 6MWT has also been demonstrated to be of use in monitoring results from training or rehabilitative programmes (Moalla et al., 2005, Rostagno and Gensini, 2008, Kim et al., 2009). Other tests such as the stair climb test (Holden et al., 1992) and incremental shuttle test (Singh et al., 1992) exist and fulfil a similar role but have been less widely used in the testing of rehabilitation or exercise programmes. Correlation between results obtained in the 6MWT and more in depth maximal exercise testing has been demonstrated by (Cahalin et al., 1995) where distance walked was found to be strongly predictive of VO_2max. Therefore the 6MWT is a reasonable alternative to maximal testing when equipment is not available or in-depth analysis is not required. (Solway et al., 2001) investigated a multitude of different non-maximal exercise tests (2MWT, 6MWT, 12MWT, shuttle walk test and self paced walk test). As well as being better researched, the 6MWT was found to be better tolerated and more reflective of activities of daily living than other tests.

A change in distance of 54m in the 6MWT was found to be clinically significant for patients with cardio-respiratory disease. Post colorectal resection Antonescu (Antonescu et al., 2014) found 19m to be the minimum clinically relevant distance change to effect outcome. The correlation of non-maximal testing results with maximal exertion testing such as VO_2max and CPET makes them attractive. If the detailed data provided by CPET is not required then there is no obvious flaw to these simpler non-maximal exercise tests. A significant upside is that they are much simpler and cheaper to perform (Palange et al., 2007). Lee echoed this sentiment where,

following multivariate analysis, a significant odds ration reduction to 0.995 was ($p<0.05$) demonstrated for every extra metre in the 6MWT for a group undergoing colorectal resection (Lee et al., 2013). The conclusion of this was that the 6MWT was a valid alternative to preoperative assessment where CPET was not available. Pecorelli (Pecorelli et al., 2016b)looked at post-operative recovery in patients undergoing colorectal resection and found that the 6MWT was a viable way to measure post operative recovery.

Subjective tests

If fitness is being tested with submaximal exertion then it is possible to ask participants for their subjective rating of fatigue, breathlessness or other indicator of exertion. Although these tests lack hard data gained from physiological readings they have been shown to be of use. Typically they are either a multi choice Likert scales with responses such as none, some, very much and maximal or a visual analogue scale scales where participants are asked to put a cross along a line that has the extremes of what is being tested at either end. Both types of scales have been found to be reproducible and are widely used (Guyatt et al., 1987). The Borg scale (Borg, 1982) is a specific version of these types of scale designed to measure perceived exertion on a scale of either 6-20 or 0-10 dependant on the version of the scale. It performs well against both visual analogue and Likert scales (Grant et al., 1999). The Borg shows good correlation with exercise induced lactate levels (Irving et al., 2006) and has also been shown to be of use for self-regulation of exercise intensity (Carvalho et al., 2009). It has also become widely used in exercise testing to assess impact of new medication (O'Donnell et al., 2004) and response to rehabilitation (Vaes et al., 2019).

Quality of life

Increased physical fitness has the effect of improving quality of life particularly in those who have significant co-morbid disease or in those who are undergoing arduous medical treatment (Latimer-Cheung et al., 2013, Thorsen et al., 2005). This is true of both physical and mental health related quality of life (Penedo and Dahn, 2005, Pretty et al., 2005, Stathopoulou et al., 2006). Multiple assessment tools exist and many studies use more than one tool. The most common assessment tool for mental health related quality of life assessment is the hospital acquired depression scale. This simple tool gives an anxiety and a depression score reading, is well validated (Spinhoven et al., 1997, Crawford et al., 2001) and widely used (Thomas et al., 2002, Mayo et al., 2011). General health-related quality of life assessment tools are numerous. The SF-36 tool (McHorney et al., 1993, Ware and Sherbourne, 1992), the EuroQuol EQ-5D (Ware and Sherbourne, 1992, Rabin and de Charro, 2001, Herdman et al., 2011), and the EORTC QLQ-C30 (King, 1996, Michelson et al., 2000) are all widely used and extensively validated in multiple populations and pathologies. The EuroQuol EQ-5D has been shown to be of use in assessing quality of life related to exercise intervention (Mayo et al., 2015) and quality of life pre-and post surgery (Werner et al., 2017) as has the EORTC QLQ-C30 (O'Neill et al., 2018) and hospital acquired depression scale (Lennon et al., 2008, Jolly et al., 2007).

1.7 Wearable activity Monitoring Technology

Advancing technology has a led to an upsurge in the use of wearable activity monitors with claims of increasing fitness levels by their usage. However it has not been demonstrated conclusively that simply wearing an activity monitor increases ones fitness levels

(Patel et al., 2015). Certain devices have advanced features including GPS and claim to be able to calculate calorie burn and other derived metrics. The use of GPS tracking is clearly not appropriate for the healthcare setting. Other metrics such as distance travelled and energy expenditure that can be derived from step counters have been shown to be unreliable (Evenson et al., 2017). It has been demonstrated that activity monitoring technology has the ability to accurately record steps taken by an individual giving them the ability to remote-supervise walking based exercise (Cadmus-Bertram et al., 2015, Takacs et al., 2014, Diaz et al., 2015). Step count error rates for Fitbit ™ devices tested on a treadmill were found to be 1.3%, with a very limited inter-device variability (intra-class correlation coefficient >0.9) (Takacs et al., 2014). Diaz investigated the reliability of Fitbit ™ step counters looking at overall reliability, and comparison of a hip and wrist based device. Both were found to be reliable but the hip placed devices were more so. Over a 6 minute walk on a treadmill the hip based Fitbit ™ was between -3.1 and -0.3 steps off the observer counted steps. This gave the hip based devices an accuracy of 97-99% compared with an accuracy of 77-85% for the wrist based FitBit ™. The technology behind these step counters is a three-dimensional accelerometer that detects motion in three directions as well as the intensity of this motion. Software algorithms them compute this raw data into a step count. The accuracy and data storing ability of these devices represent an attractive monitor of walking based exercise that may be an adjunct to prehabilitation.

1.8 Aims of the Research Study and Hypotheses

Aims: To assess acceptability and achievability of a home based prehabilitation programe in terms of recruitment and compliance and to assess the efficacy of the exercise intervention compared with standard practice of no intervention.

Primary Hypothesis

It is acceptable and achievable for patients to undertake a simple walking-based prehabilitation programe, monitored by wearable technology, in the waiting time before intestinal surgery.

Secondary Hypothesis

It is possible to increase fitness levels and quality of life scores of a simple walking-based prehabilitation programe, monitored by wearable technology, in the waiting time before intestinal surgery.

Rationale

It has been demonstrated that simpler, less intensive exercise programmes have improved recruitment rates and are possibly more effective than more complex and high intensity regimes. Patients who have undertaken prehabilitative programmes express a preference to low supervision home-based programmes. Activity monitoring technology has been proven to be accurate in measuring step count. Using such devices can remove the need for personal supervision and lessen the need for transportation that has been demonstrated to be a significant barrier to participants.

Approach

40 Participants listed for resectional bowel surgery were randomised to normal activity or walking based exercise monitored by a Fitbit

Zip™ advanced pedometer. Both groups underwent baseline, immediate preoperative and 3 month postoperative assessment with 6MWT and QOL questionnaires.

2.Methods

2.1 Design

This was a pilot randomised controlled trial (RCT).

Patients undergoing resectional bowel surgery were recruited from Scarborough General hospital. In total, 40 participants were recruited for this pilot study and randomly allocated into one of 2 groups:

1: A group who took part in a home-based walking programme
2: A group who continued with their usual level of physical activity.

The Leeds West ethics committee have granted ethical approval, 16/YH/0049, for this study.

Primary Hypothesis

It is acceptable and achievable for patients to undertake a simple walking-based prehabilitation programe, monitored by wearable technology, in the waiting time before colorectal surgery.

Acceptability was measured as the number of potential participants approached to recruit the 40 participants. Achievability of the pre-operative exercise programe was assessed by analysing the number of days each participant in the exercise group meets their daily exercise target, giving percentage compliance for each participant.

Secondary Hypothesis

It is possible to increase fitness levels and quality of life scores of a simple walking-based prehabilitation programe, monitored by wearable technology, in the waiting time before intestinal surgery. The secondary hypotheses will be measured using:

- The 6 Minute Walk Test (distance walked and change in heart rate)
- The Borg Scale
- The EQ-5D-5L
- The Hospital Anxiety and Depression Scale
- EORTC QLQ-C30

An RCT was chosen over a cohort study design. Although a cohort study would of allowed more participants to be tested for the primary hypothesis it does not allow testing of the secondary hypothesis, as a control arm is needed to detect any increase in fitness due to the intervention.

A decision was made not to restrict entry criteria to those with a malignant diagnosis in order to keep the study as similar as possible to normal practice. The expansion of entry criteria further to include all needing a general surgical operation would have increased numbers greatly. However patients who are listed for perhaps gallbladder or hernia surgery are a very different cohort to those that need resectional GI surgery, and therefore a decision to not include them was made.

2.2 Patient Selection

Inclusion Criteria

- Aged 18 years or older
- Requiring elective resectional bowel surgery
- Gives consent to participate in the study

Exclusion Criteria

- Younger than 18 years of age.
- A history of unstable angina/unstable coronary artery disease or a heart attack in the previous month.
- Any heart related disease including but not limited to aortic stenosis, pericarditis or any thromboembolic disease.
- Severe Infections and fever needing acute medical therapy.
- Uncontrolled or unstable metabolic diseases.
- Resting heart rate of more than 120 BPM.
- Systolic blood pressure of more than 180 mm Hg or diastolic blood pressure of more than 110 mm Hg.
- Cerebrovascular accident within the last month.
- Pregnancy
- Unwilling to allow their GP to be informed of their participation in the study. Not able or unwilling to consent to take part in the study

2.3 Identification and Recruitment

Potential participants were identified at the weekly colorectal Multidisciplinary (MDT) meeting at Scarborough Hospital. All patients discussed at the MDT meeting then attend an outpatient clinic within 48 hours as routine. At that outpatient clinic appointment, potential participants were approached about the study with verbal information and provided with an information sheet. After allowing potential participants to consider the information, the investigator met the people who have been approached at a further routine appointment (most commonly their preoperative assessment appointment). If willing to participate, consent was taken, and the initial assessment carried out. The time available for people to consider participation was restricted because there was limited time to instigate the exercise programme before the participant has their surgery.

2.4 Randomisation

Following initial assessment participants were randomised on a 1: 1 ratio into one of two groups

Group 1 - Exercise programme.
Participants in this group participated in a monitored, moderate intensity, home based and individualised walking programme.

Group 2 - Control group.
Participants in this group continued with their usual level of physical activity.

Block randomisation was used. The randomisation sequence was generated using a computer software programme from www.sealedenvelope.com. Allocations were concealed in sequentially numbered, opaque envelopes. Investigators were not involved in this process.

2.5 The Exercise Programme

Participants were given a personalised daily exercise target. This was three times the number of steps they completed in the initial 6-Minute Walk Test. This constitutes 18 minutes of moderate exercise that is above the minimum 10 minutes recommended in WHO guidance and per week almost reaches the recommended physical activity levels without considering other day-to-day activity undertaken. A 10 person patient advisory panel reviewed the protocol for the exercise intervention in an effort to make it as appropriate as possible. Participants were given an activity tracker (a Fitbit Zip™). This was chosen due to the simplicity of its use for participants and its ability to store data internally making the need for

data retrieval less frequent. The participants were asked to wear the device all day to monitor their activity. Once a day the participants were asked to complete the walking target they had been given at completion of their baseline assessment. They were asked to complete this target in one go. Participants were able to tell when the target number of steps has been reached by looking at the display of the device. The device stores data on the daily steps the participant completes with a minute-by-minute breakdown of this activity. The device has no buttons to alter its function and the display will show steps completed only, resetting to zero at midnight. Participants needed only to clip it onto an item of clothing close to the hip. No other interaction with the device is needed and no alteration of settings is possible. There is no recharging needed and a battery life was checked prior to hand over to participant to ensure it is adequate for the whole study period. The length of time in the walking programe depended on the length of time between recruitment and when the participant underwent their surgery.

2.6 Assessments

Assessments were carried out:

- On entry to the study
- Following the intervention / control period but prior to surgery
- 3 months after surgery

Participants were asked to wear comfortable clothing and flat shoes appropriate for walking exercise.

The assessment comprises:

- Documenting blood pressure and resting heart rate
- The 6 Minute Walk Test

- The Borg Scale
- The EQ-5D-5L
- The Hospital Anxiety and Depression Scale
- EORTC QLQ-C30

The participant's blood pressure (BP) and heart rate was measured by an automated BP cuff (GE Carescape V100 Dinamap Vital Signs Monitor). The BP cuff was placed on the upper arm and the BP and heart rate was measured after the participant had been resting for three minutes.

The Six Minute Walk Test is a marker of endurance capacity and requires a person to walk as far as they can during 6 minutes (Crapo et al., 2002). The test is self-paced, easy to administer and well tolerated. It has also been used for a number of clinical populations with good reliability and validity in cardiopulmonary patients in particular (Crapo et al., 2002).

A 30 metre flat walking surface is set out with cones marking each 3 metre interval with distinct markers at the start and end. Following a period of 10 minutes seated rest participants are required to walk as far as possible in the 6 minute period. They are allowed to stop and rest if required however the clock carries on ticking i.e. any rest periods are included in the six minutes.
A protocol is used to reduce variability within the test. Standardised words of encouragement are provided each minute as instructed by the American Thoracic Society. At the end of the 6 minutes, participants stop when instructed and the total distance walked provides the primary outcome measure.

The participants wore the activity tracker (Fitbit Zip) whilst undergoing this assessment. This provided a measure of the number of steps taken over this distance. This was used to inform the walking

programme that was set for participants allocated to the intervention group.
Heart rate was taken at the end of the Six Minute Walk test and after two minutes recovery time.

In addition to the total distance walked, a rating of perceived exertion was recorded each minute using the Borg scale (Borg, 1982). The Borg Scale is a fifteen point relative scale ranging from 0–10 where 0 is no noticeable exertion and 10 is absolute maximal exertion.

The EQ-5D is one of the most well known and commonly used generic measures of health status internationally (Devlin and Krabbe, 2013). The EQ-5D-5L descriptive system (Herdman et al., 2011) comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. In addition to the descriptive system, there is the EQ VAS that records the respondent's self-rated health on a 20 cm vertical, visual analogue scale where the endpoints are labelled 'Best imaginable health state' and 'Worst imaginable health state'.

The Hospital Anxiety and Depression Scale (HADS) is commonly used to determine the levels of anxiety and depression that a patient is experiencing. The HADS is a fourteen-item scale. Seven of the items relate to anxiety and seven relate to depression. This creates a tool for the detection of anxiety and depression in people with physical health problems (Spinhoven et al., 1997, Crawford et al., 2001). The HADS was chosen for use in this study as it is the most commonly used mental health related quality of life assessment (Thomas et al., 2002, Mayo et al., 2011).

The EORTC QLQ-C30 is a quality of life scale involves physical and mental assessment as well as the impact of medical therapy on quality of life. It comprises of 30 questions where the participant is asked to answer on a numeric scale(Aaronson et al., 1993).

2.7 Fitbit Zip™ Device

The Fit Bit zip is a pedometer that measures steps by using a tri-axial accelerometer. This enables it to detect steps when it is any orientation. By inputting user anthropomorphic data it can calculate distance travelled and calories burnt. This data is displayed on the devices screen and different modes can be brought up by tapping on the screen. The device stores minute by minute data. Each day at midnight the device resets itself to zero. The device comes with a rubber housing incorporating a clip to attach it to clothing. This renders it sweat and splash proof but not water proof. It can be worn on any item of clothing and still record steps accurately.

Downloading of the data stored on the fit bit is done via Bluetooth to a smart phone or to a small Bluetooth receiver supplied with the device that can be inserted into a computer via a USB port
To download the data an account with fit bit must first be created. A Fit Bit device can then be registered to this account and the downloaded data saved. Although it is possible to add multiple devices to a single account downloaded data from different devices may get mixed and as such render it invalid for research purposes. To avoid this, each separate device has its own account. To set up an account a valid email address is needed. Each Fit Bit Zip had a Google Gmail account set up for it. These are sghfitbit1@gmail.com,

sghfitbit2@gmail.com and so forth. These email addresses were then used to set up an account with Fit Bit to download data to. No participant data is stored within these accounts. We recorded which fit bit was given to which participant for the time they were in the study. This information was used to identify the data stored on the Fit Bit account and allocate it to the correct participant. The calculated data of calories burnt and distance travelled that the device generates is of no interest to this study. As such the anthropomorphic data that has to be imputed in to the Fit Bit was identical for each device (DOB 1/1/2000, height 170cm, weight 10 stone, male sex) and bore no resemblance to the participants. As this study uses the step count only to inform participants of the amount of walking exercise to do each all other modes were switched off. This means what is displayed on the screen of the device cannot be changed from step count completed that day.

2.8 Data Management and Statistical Analysis

All participants that consent to entry into this study were be included in the analysis whether they complete the study or not. There is no set minimum compliance and as such all will be included in an intention to treat analysis. Data were entered into a Microsoft Excel spreadsheet (Excel for Windows, Microsoft Corporation, Redmond, Washington, USA). For the analysis of the results, the IBM Statistical Package for the Social Sciences (SPSS) version 24 was used. A variety of statistical analysis techniques were used, depending on the distribution of the data. To assess for the distribution of the data, a histogram was drawn. This enabled easy identification of whether the data being analysed was normally distributed or skewed. Where it is unclear if a histogram is displaying normal distribution a Shapiro-Wilk test will be used with a value of >0.05 denoting a normal distribution.

Demographics of the two groups were compared for any significant difference with a p value set at 0.05. Normally distributed data were analysed with either a Chi-Squared test with Yates' continuity correction or an independent samples T test. Non-normally distributed data were analysed by a Mann-Whitney U test.

Demographics analysed were:

- Time between 1st and 2nd assessments (time in study)
- Time between 2nd and 3rd assessments (post operative recovery time before assessment)
- Completion of all 3 assessments

Acceptability of the proposed exercise programe to people scheduled for resection bowel surgery was calculated as a percentage of those who agreed to take part (not knowing which arm they would be randomised to) of the total approached. All those approached had the study explained to them and were given information leaflets to read. It was only after they had taken these away and had more than 24hrs to consider their decision that they were deemed to have declined or accepted the invitation.

Achievability of the pre-operative exercise programe was assessed by analysing the number of days each participant in the exercise group meet their daily exercise target, giving a percentage compliance for each participant.

Increase in fitness levels was defined by a change in the distance walked in meters between assessments and a change in heart rate. Change was analysed between the 1st and 2nd, 2nd and 3rd and 1st and 3rd assessments. Heart rate was analysed to give a change between; resting heart rate and immediate post 6MWT heart rate; immediate post 6MWT heart rate and 2 minutes post 6MWT heart rate. This was calculated for all assessments. These differences in

heart rate values were then compared between assessments for each participant and groups compared to analyse if the exercise intervention had altered how heart rate increased post 6MWT and recovered after it. Normally distributed data were analysed by an independent samples T-test and non-normally distributed data with a Mann-Whitney U test.

The three quality of life scoring assessments give vast amounts of data the majority of which is not relevant to this study. For example data output from the EQ-5D-5L includes information on bowel habit, which is then included in some of the composite scores. As bowel habit is not an outcome of this study we felt this makes interpretation of the scores generated where this is included difficult. Therefore certain outputs from each assessment have been chosen as below:

- HADS
    - Depression Score
    - Anxiety Score

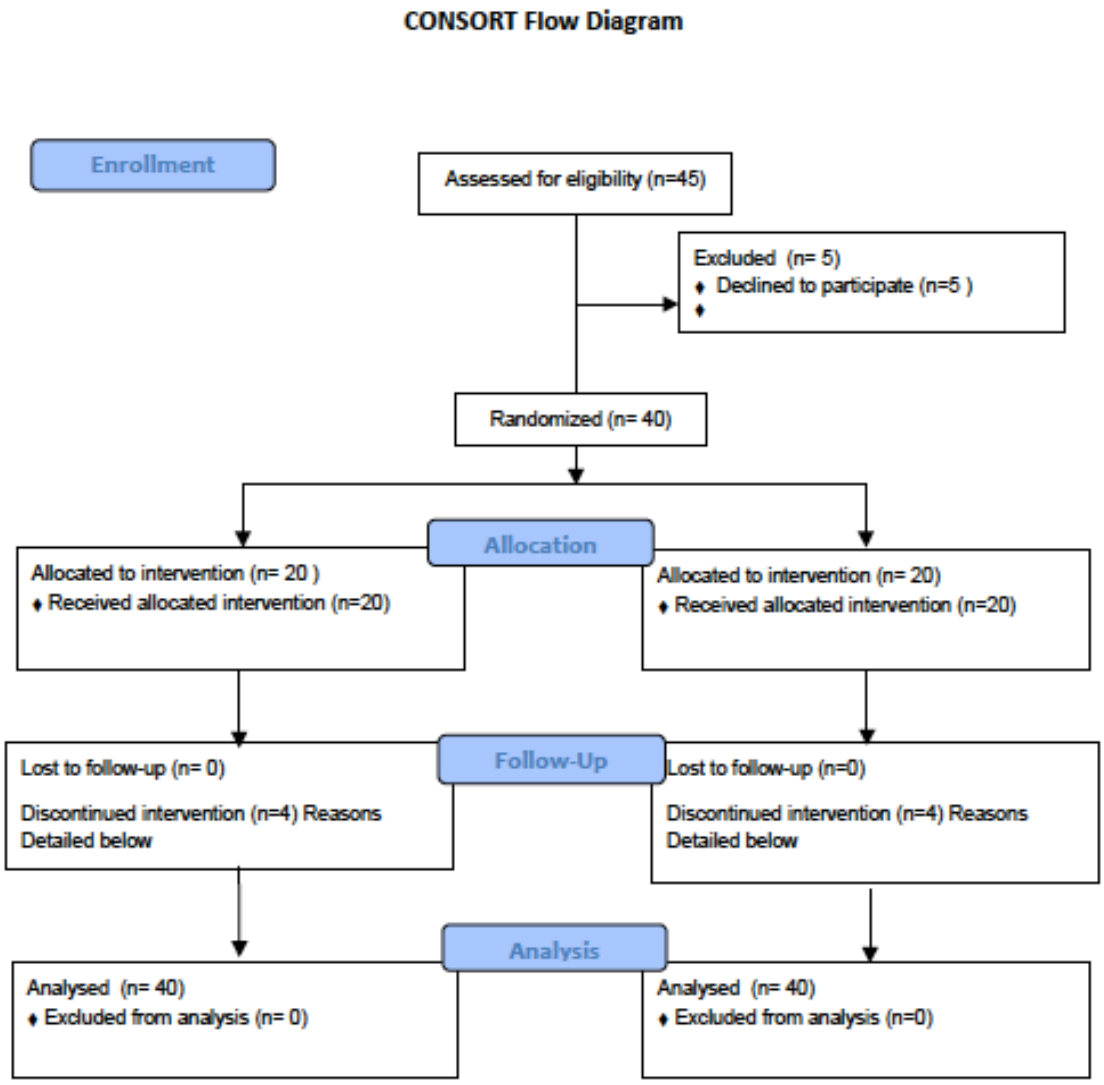
- EQ-5D-5L
    - Total Health Today

- EORTC QLQ-C30
    - Combined Quality of life and Heath in last week

The raw scores for each participant were calculated in line with the user manuals provided with each assessment tool. The change in each participant's score between assessments was then calculated and groups compared.

3. Results

3.1 Total Data Collected and CONSRT flow chart



40 participants were recruited in total with 20 in each arm of the trial. 40/40 (100%) of participants completed the 1st assessment in its entirety. 38/40 (95%) participants completed the 1st and 2nd assessments in their entirety. 32/40 (80%) participants completed all 3 assessments in their entirety. The 2 participants who did not complete the 2nd or 3rd assessments did neither the QOL assessments nor the 6minute walk. All of the 6 participants who only failed to complete the 3rd assessment did not attempt the 6-minute walk. 3/6 of these participants did complete the QOL assessments. A list of reasons given for not participating in the 2nd or 3rd assessment is shown below.

Not completing 2nd and 3rd assessment

- Too tired for 2nd and not interested in taking part any further for 3rd.
- Received news on pre-operative day that primary was gallbladder not colon cancer in hepatic flexure and therefore not resectable. Participant withdrew from study

Not completing 3rd assessment

- Painful peripheral neuropathy so not willing to complete 6MWT
- Inpatient still from surgery and not fit to undergo 3rd assessment
- Developed hip pain so not willing to undergo 6MWT
- Frequent visits to Christie hospital in Manchester for further treatment so unable to attend for 3rd assessment
- Felt too weak whilst on chemotherapy to undergo 6MWT
- Multiple admissions in the months postoperatively including delirium so deemed not fit to complete 3rd assessment after discussion with family.

3.2 Demographics

Demographics of the participants of the exercise and normal activity arms are displayed in table 2.

Demographic	Exercise	Normal Activity
Time between 1st and 2nd assessments in days (SD)	20.8 (24.2)	13.2 (9.9)
Time between 2nd and 3rd assessments in days (SD)	111.6 (28.6)	103.8 (18.8)
Completed all 3 assessments	16/20 (80%)	16/20 (80%)

Table 2: Demographics of Exercise and Normal Activity groups

Time in from 1st to 2nd assessment

Figure 3 is a stacked histogram displaying the non-normal distribution of time from 1st to second assessment (or time in study) for the exercise and normal activity arms. Figure 4 displays the mean time in study for exercise and normal activity arms with bars displaying the 95% CI. Data were analysed using the Mann-Whitney U test. The majority of participants had their operation

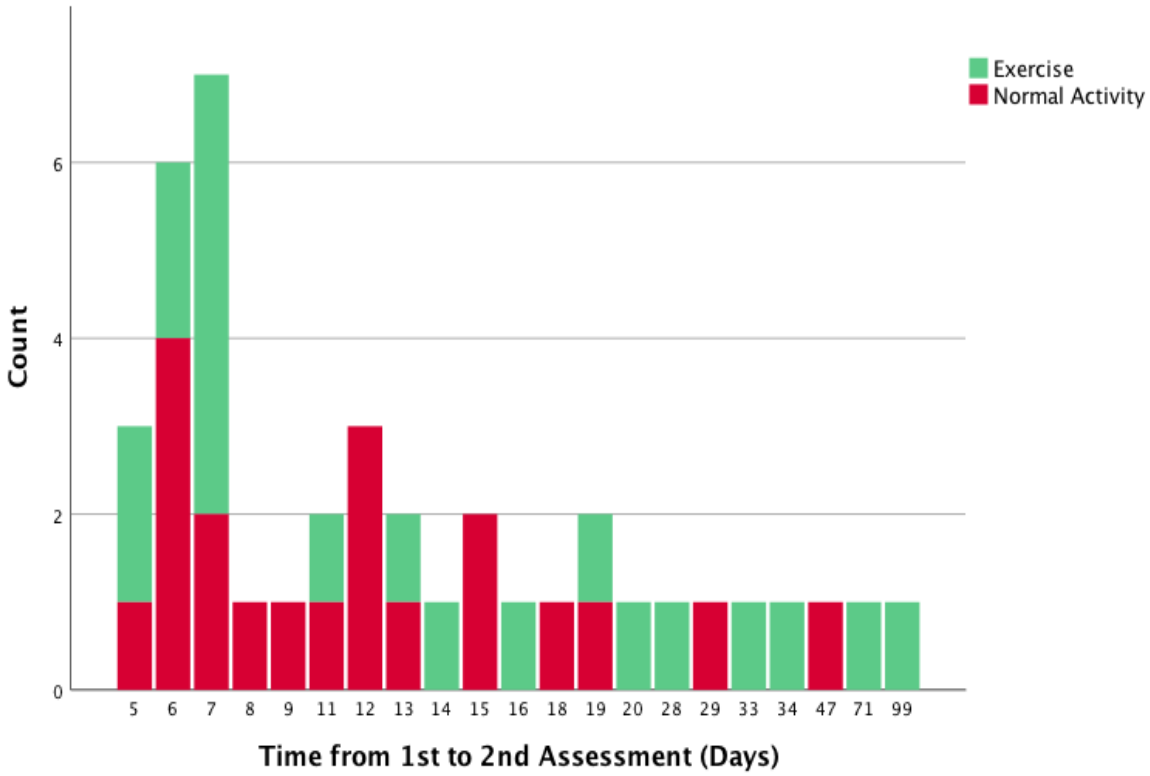


Figure 3: Stacked histogram time in study

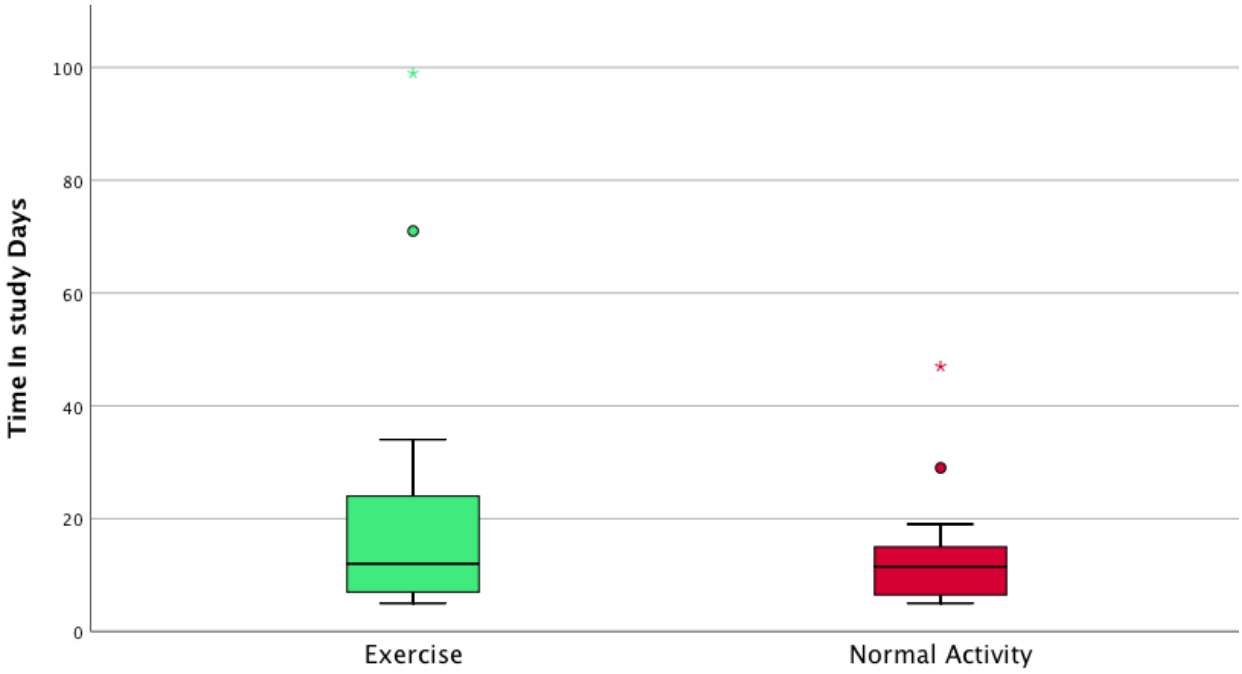


Figure 4: Mean time in study in days exercise vs. normal activity

Time from 2nd to 3rd Assessment

Figure 5 is a stacked histogram displaying the normal distribution of the time from 2nd to 3rd assessment (or time to follow up) for the exercise and normal activity arms. Figure 6 displays the mean time from 2nd to 3rd assessments for the exercise and normal activity groups with bars displaying the 95% CI. Assessments were undertaken when participants were visiting the hospital as much as possible. The 3rd assessment was to be undertaken at 3 months post operation as this is a usual time for a follow up clinic. However adjuvant chemotherapy, complication and further treatment at distant sites meant some participants were not willing or able to undertake their 3rd assessment at the 3 month post-op mark.

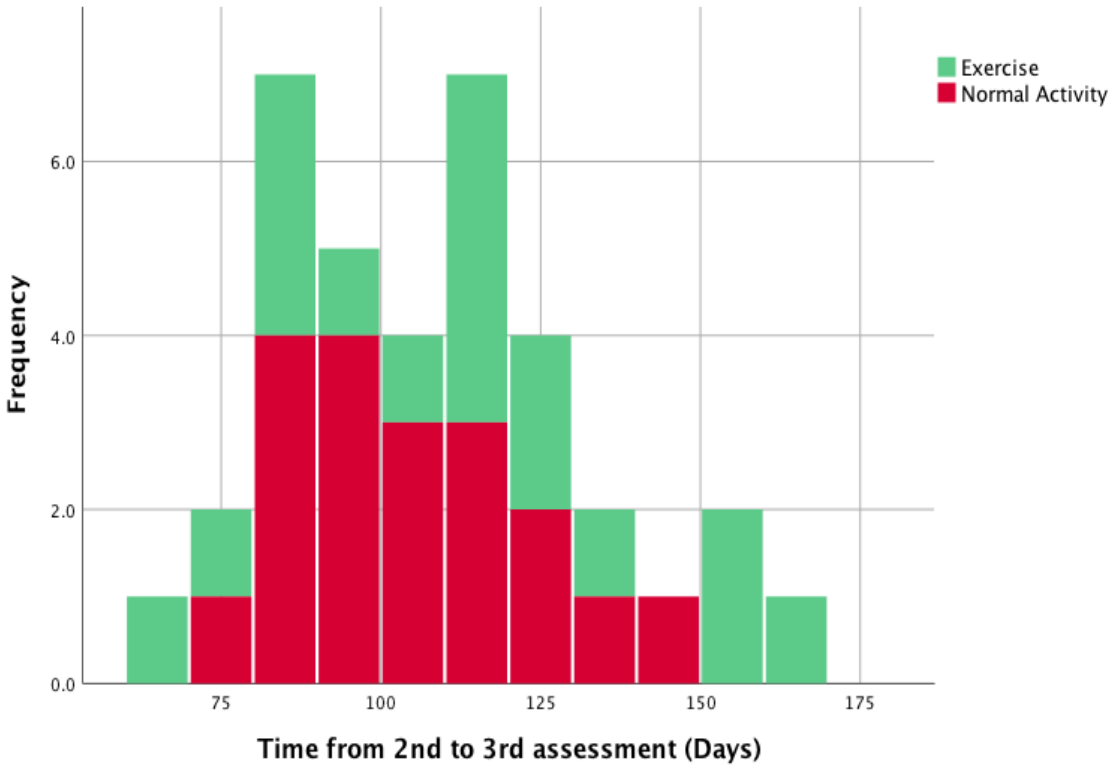


Figure 5: Stacked histogram time from 2nd to 3rd assessment

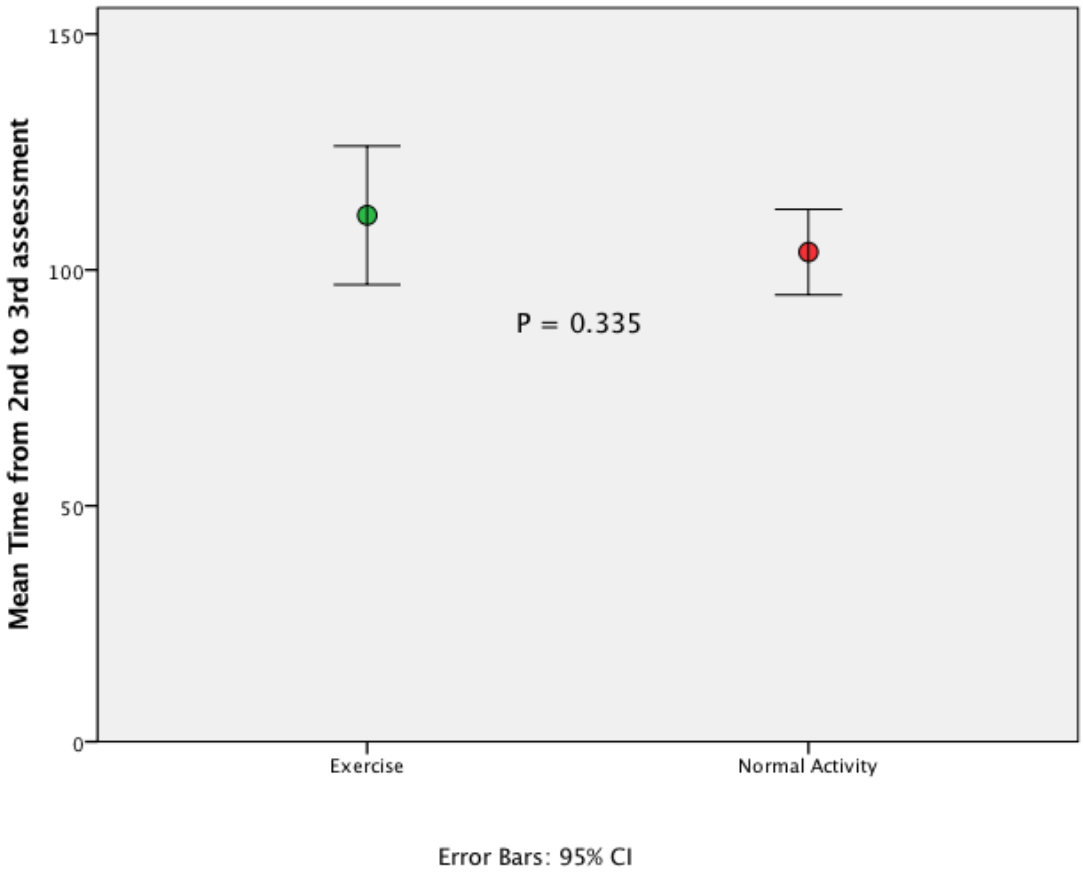


Figure 6: Mean time in days from 2nd to 3rd assessment exercise vs. normal activity

Completed All 3 Assessments

Figure 9 displays the percentage of the exercise and normal activity arms that completed all 3 assessments. A Chi-Square Test for independence (with Yates Continuity Correction) indicated no significant difference between groups, $\chi^2 = 0.0$, $p = 1$.

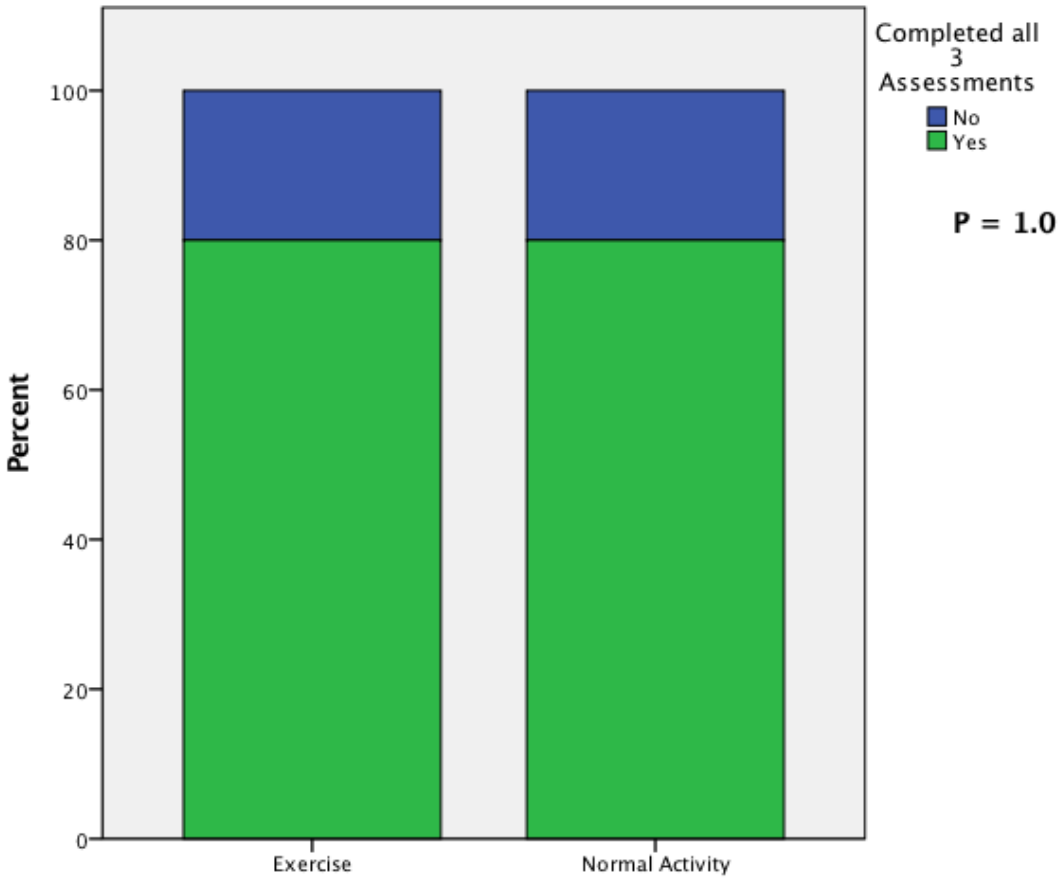


Figure 9: Completed all 3 assessments exercise vs. normal activity

3.3 Acceptability of home based walking exercise

45 potential participants were approached to recruit the 40 participants needed for the study. This gives an acceptability of this exercise programe of 89% in patients scheduled for resectional bowel surgery.

The Fitbit Zip™ device was worn for an average of 93% of the days available to the exercise group Figure 10 displays a scatter chart of this data.

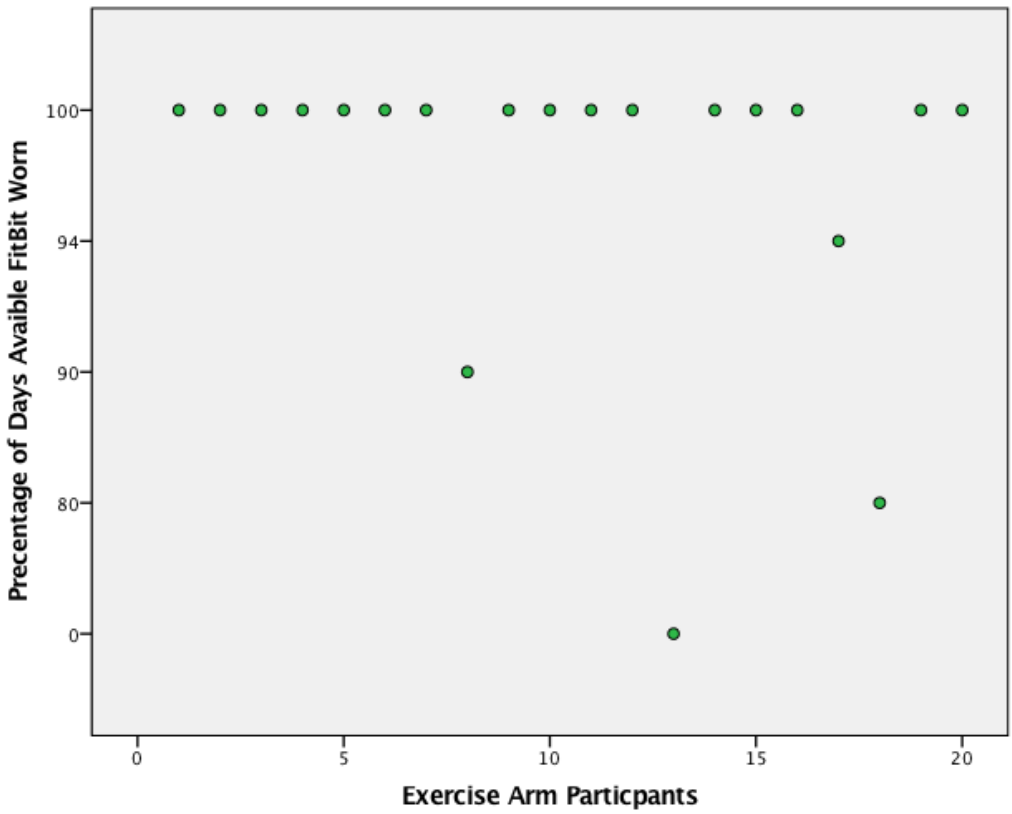


Figure 10: Percentage of days Fitbit worn

3.4 Achievability of home based walking exercise

Figure 11 displays a histogram displaying the non-normal distribution of percentage compliance with the exercise programme. Median compliance was 68%.

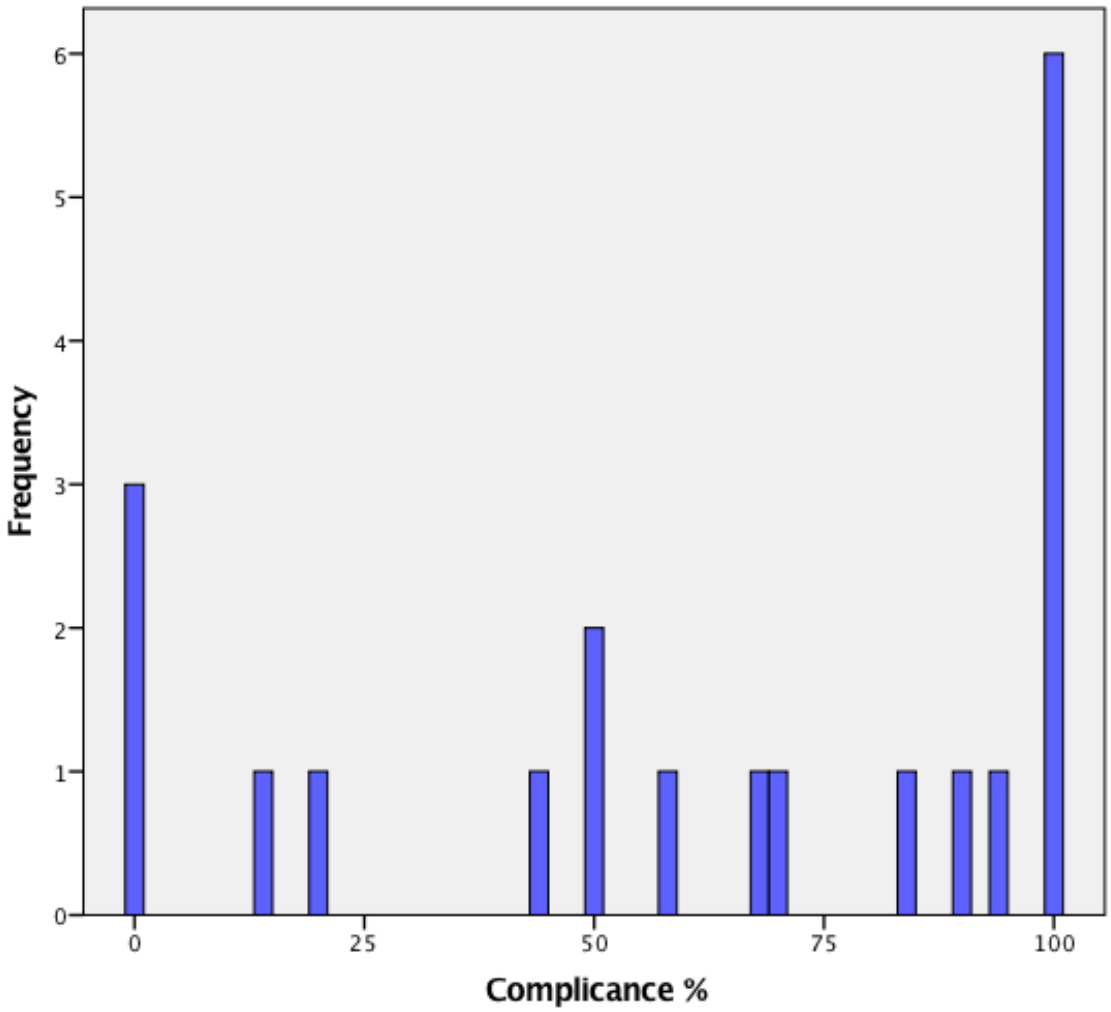


Figure 11: Percentage compliance with exercise programme

Figure 12 shows a scatter plot with best-fit line of distance walked at 1st assessment 6MWT and percentage compliance. As the least fit individuals (shortest distance in 6MWT) may stand to benefit the most from exercise it is of interest to see how they comply with exercise versus the most fit. The relationship between these variables was investigated using Spearman rho given the non-normal distribution of the data. There was a medium positive correlation between the two variables, rho = 0.455, p=0.044.

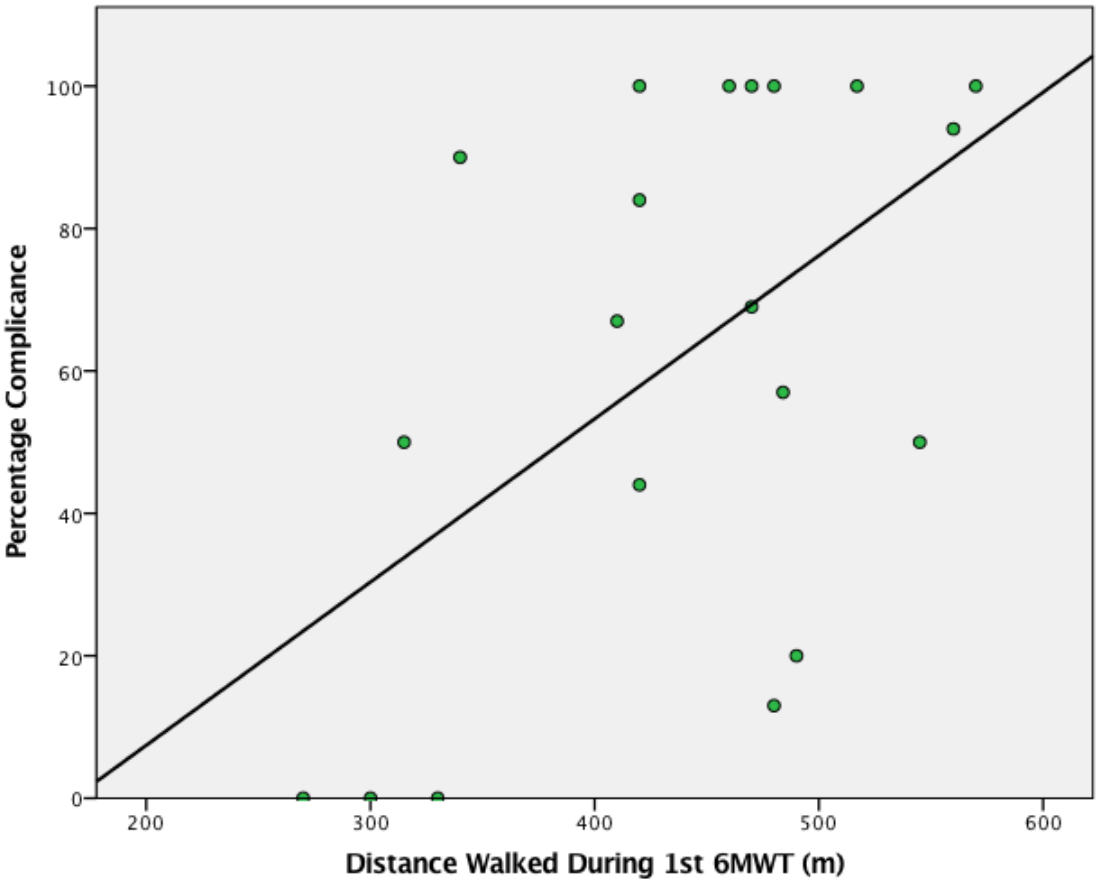


Figure 12: Scatter chart of distance walked in first 6MWT vs. percentage compliance

Figure 13 shows a scatter plot with best-fit line of distance walked at 1st assessment 6MWT and percentage change in distance walked pre and post exercise program. The relationship between these variables was investigated using Spearman rho given the non-normal distribution of the data. There was a medium negative correlation between the two variables, rho = 0.349, p=0.131.

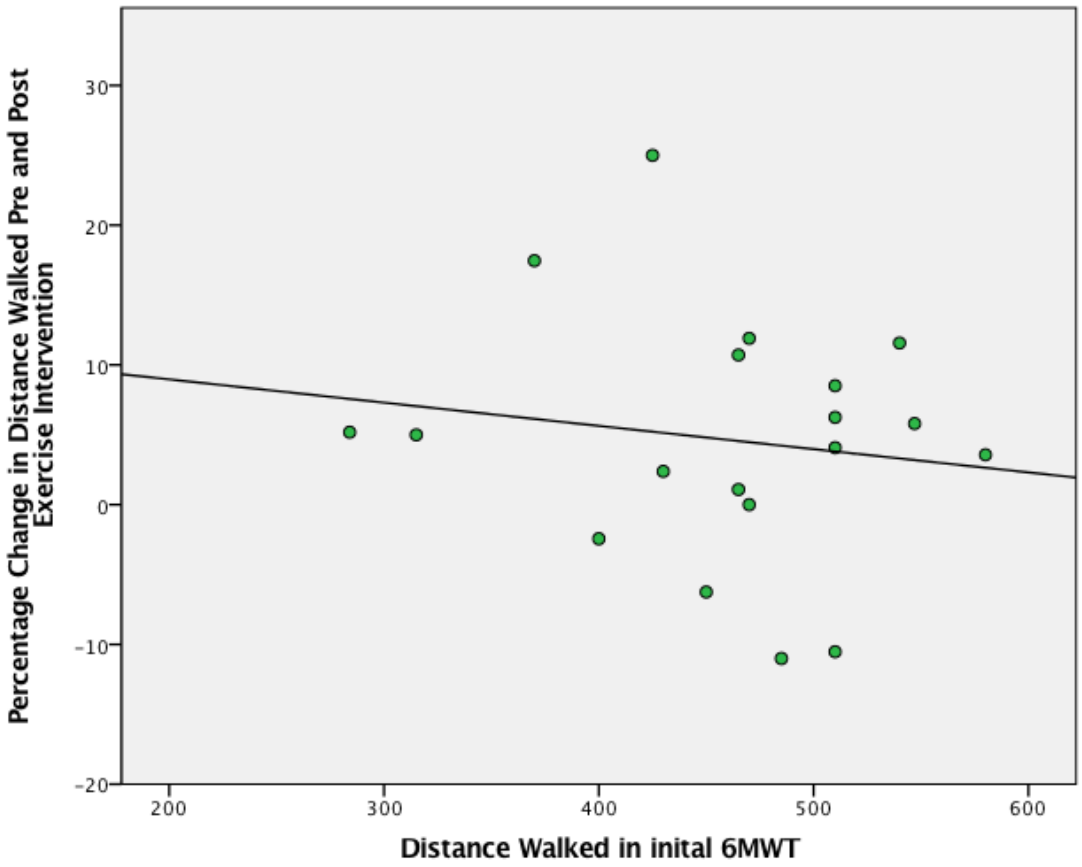


Figure 13: Scatter chart of percentage change in distance walked pre-and post exercise vs. distance walked in 1st assessment

3.5 Change in Distance Walked in 6-Minute Walk Test between Assessments

1st to 2nd assessment

Figure 14 is a histogram displaying the normal distribution of the change in distance walked in the 6MWT between the 1st and 2nd assessment. Mean change in distance walked between the exercise and the normal activity groups was +16.8m (SD 37.7) and -12.6m (SD 38.1) respectively. Data were analysed with independent samples T-Test and the difference of 29.2m between groups was significant, p = 0.023. Figure 15 displays the mean change in distance walked for exercise and normal activity arms with bars displaying the 95% CI.

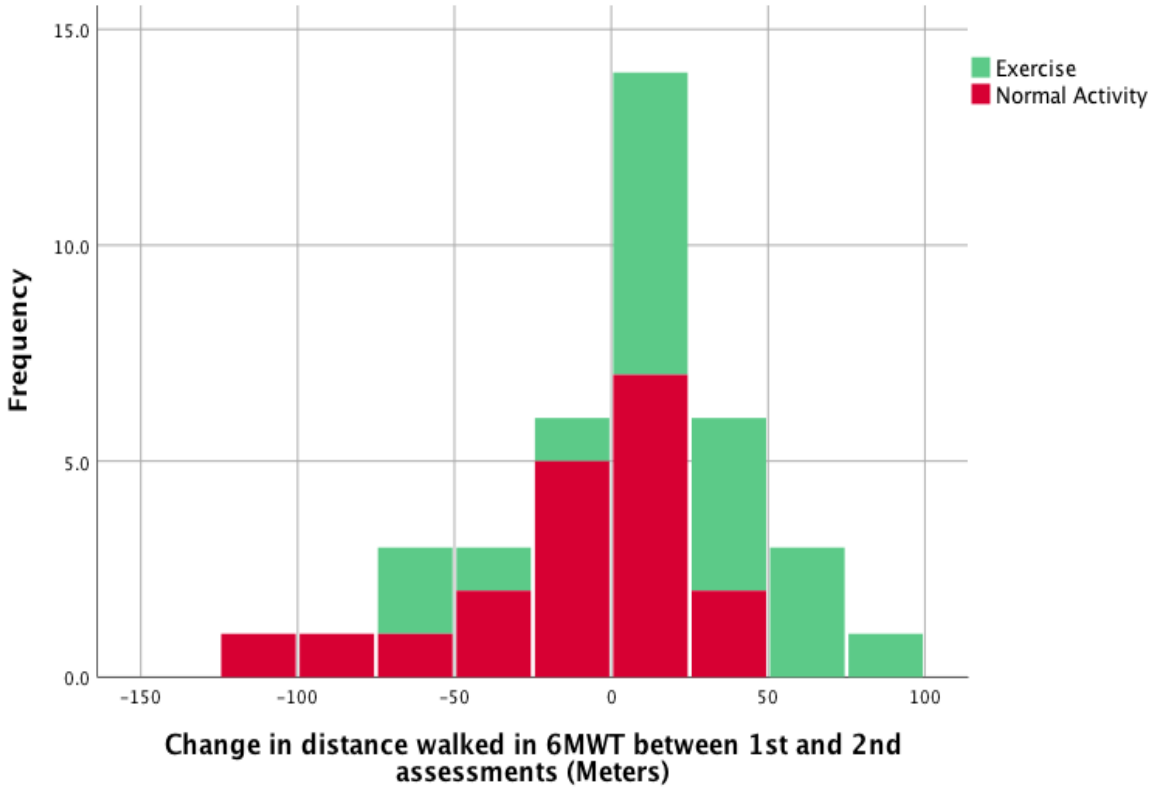


Figure 14: Stacked histogram of change in distance walked in 6MWT between 1st and 2nd assessments

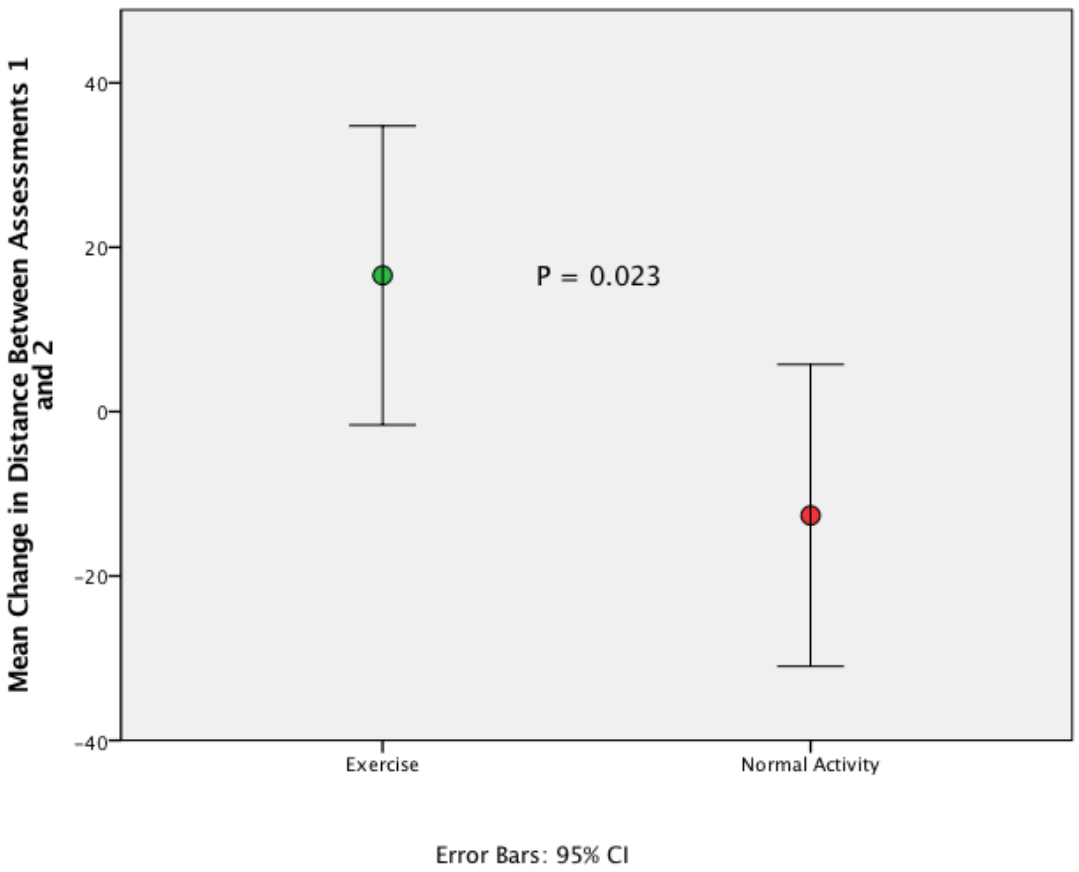


Figure 15: Mean change in distance walked (meters) in 6MWT between 1st and 2nd assessments, exercise vs. normal activity

2nd to 3rd assessment

Figure 16 is a stacked histogram displaying the normal distribution of the change in distance walked in the 6MWT between the 2nd and 3rd assessment. Mean change in distance walked between the exercise and the normal activity groups was -26.3m (SD 47.2) and -27.3m (SD 54.4) respectively. Data were analysed with independent samples T-Test and the difference of 1m between groups was not significant, p = 0.957. Figure 17 displays the mean distance walked for exercise and normal activity arms with bars displaying the 95% CI.

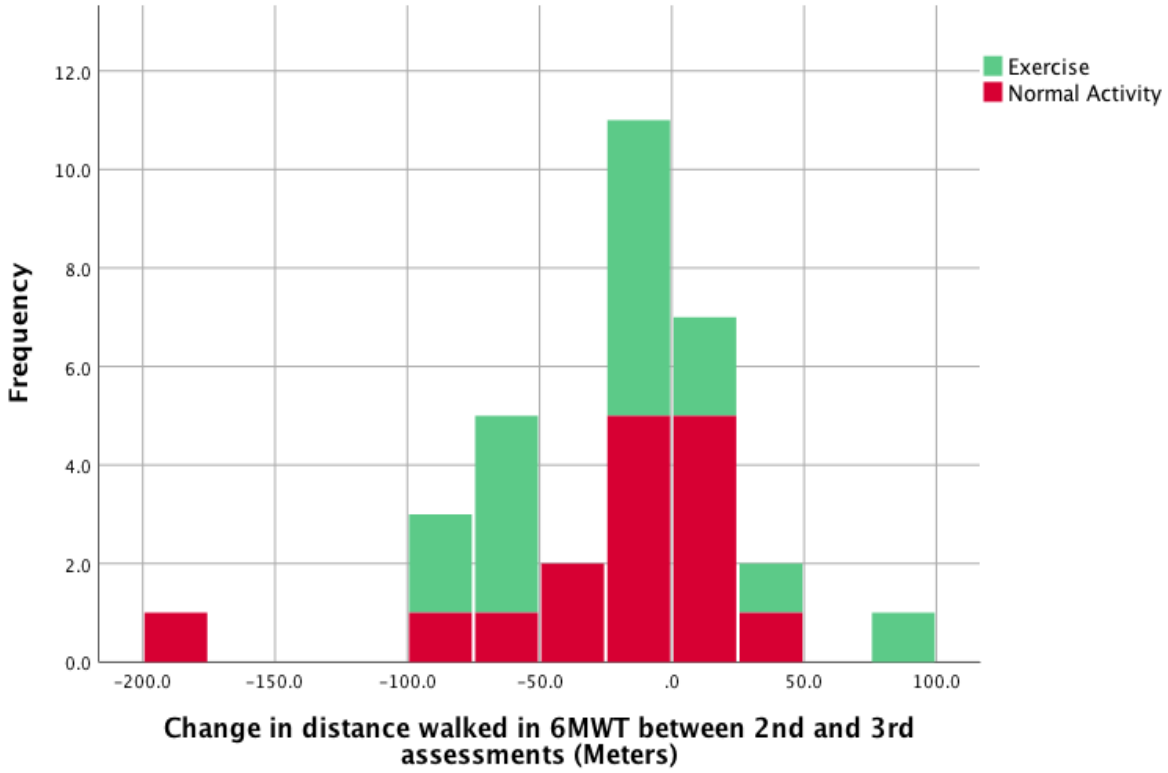


Figure 16: Stacked histogram of change in distance walked in 6MWT between 2nd and 3rd assessments

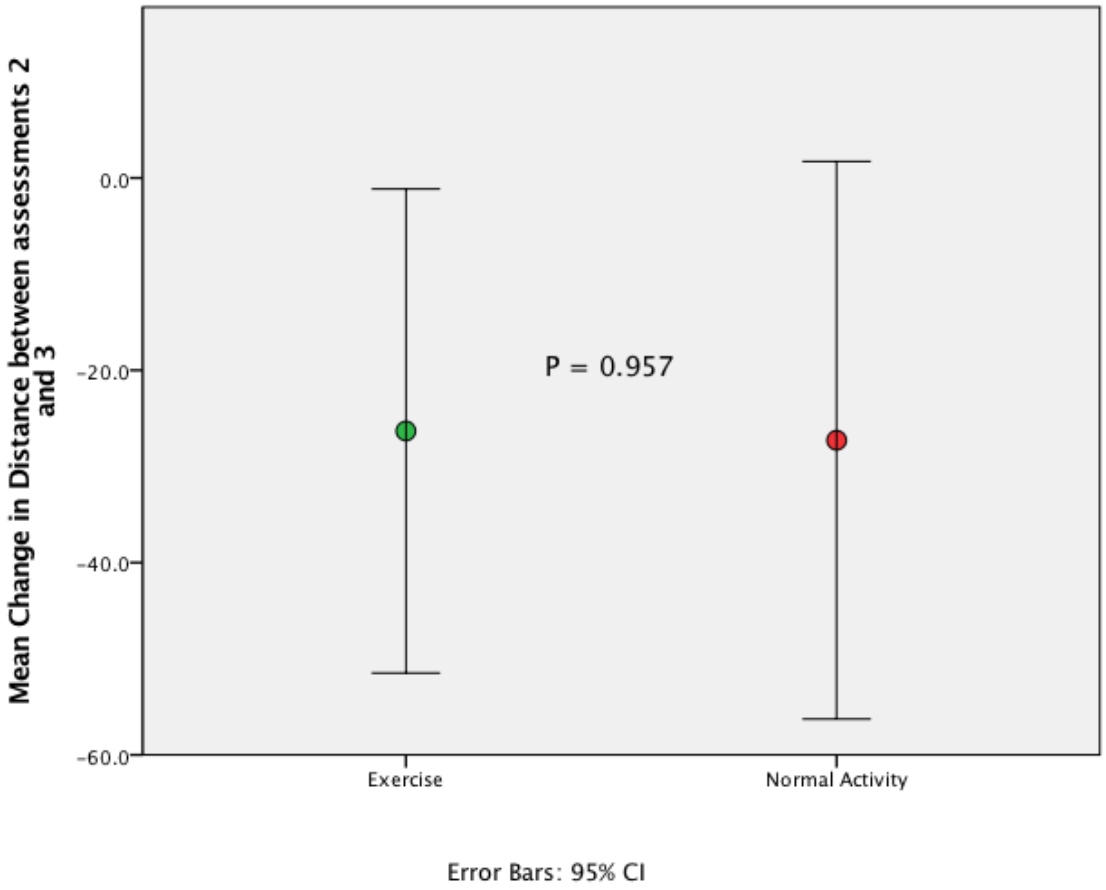


Figure 17: Mean change in distance walked (meters) in 6MWT between 2nd and 3rd assessments, exercise vs. normal activity

1st to 3rd Assessment

Figure 18 is a stacked histogram displaying the normal distribution of the change in distance walked in the 6MWT between the 1st and 3rd assessment. Mean change in distance walked between the exercise and the normal activity groups was -10.8m (SD 43.5) and -38.8m (SD 59.3) respectively. Data were analysed with independent samples T-Test and the difference of 28.8m between groups was not significant, p = 0.137. Figure 19 displays the mean distance walked for exercise and normal activity arms with bars displaying the 95% CI.

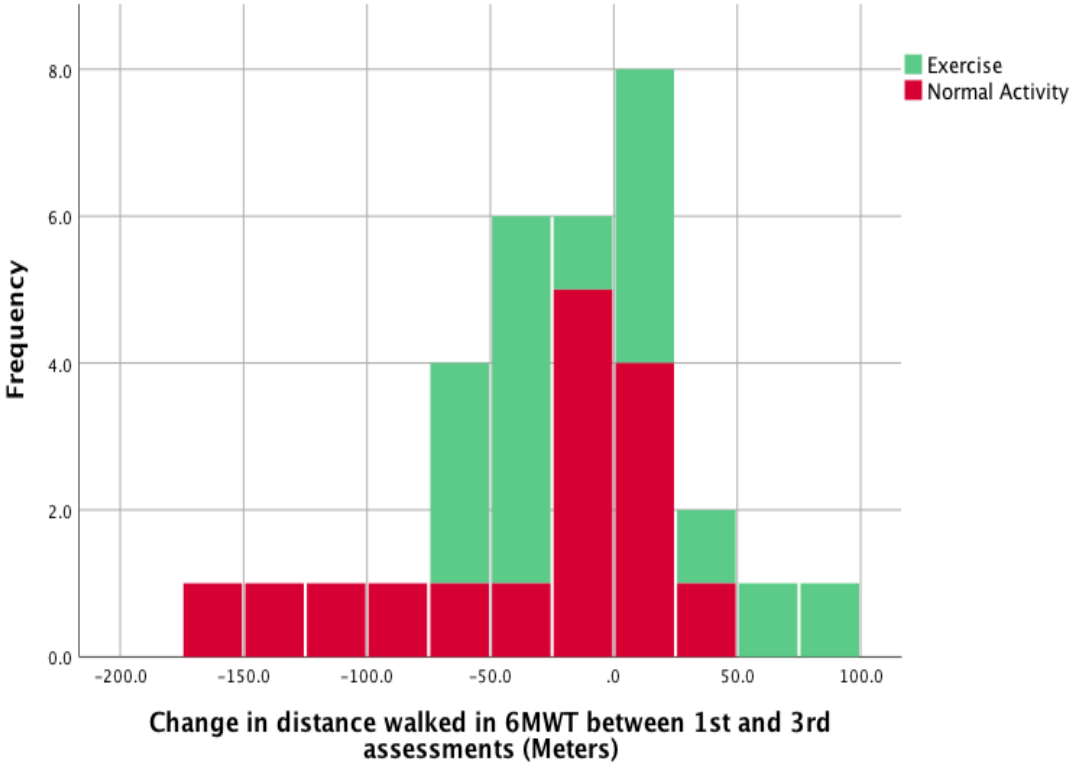


Figure 18: Stacked histogram of change in distance walked in 6MWT between 1st and 3rd assessments

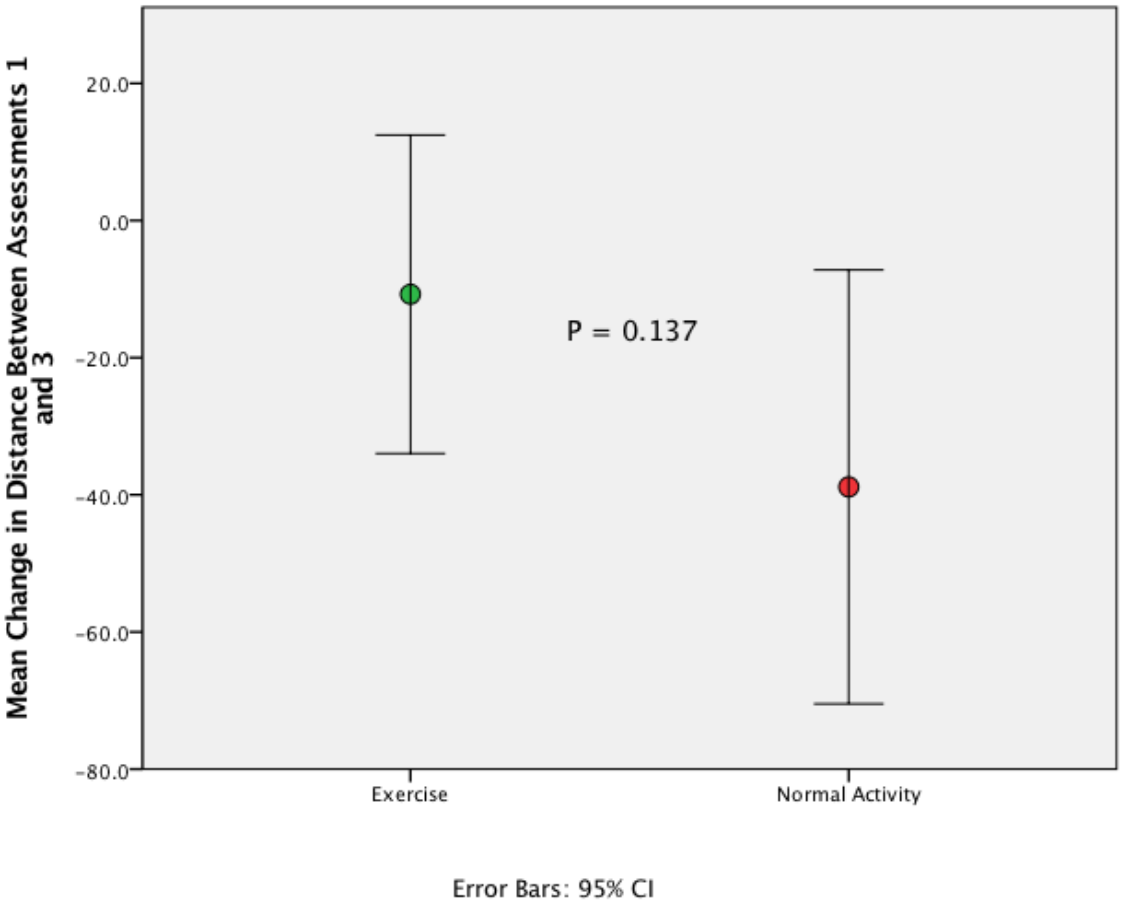


Figure 19: Mean change in distance walked (meters) in 6MWT between 1st and 3rd assessments, exercise vs. normal activity

3.6 Change in heart rate monitoring between Assessments

Heart rate was measured in BPM at rest, post exercise and after a further 1 minute of seated recovery. Results shown are changes in resting heart rate, increase in heart rate post exercise and heart rate recovery between assessments.

Change in resting heart rate between assessments 1 and 2

Figure 20 is a stacked histogram displaying the normal distribution of the change in the resting heart rate between the 1st and 2nd assessment. Mean change in BPM between the exercise and the normal activity groups was -4.7 BPM (SD 10.2) and -0.7 BPM (SD 11.0) respectively. Data were analysed with independent samples T-Test and the difference of 4 BPM between groups was not significant, p = 0.121. Figure 21 displays the mean change in the resting heart rate between the 1st and 2nd assessment for exercise and normal activity arms with bars displaying the 95% CI.

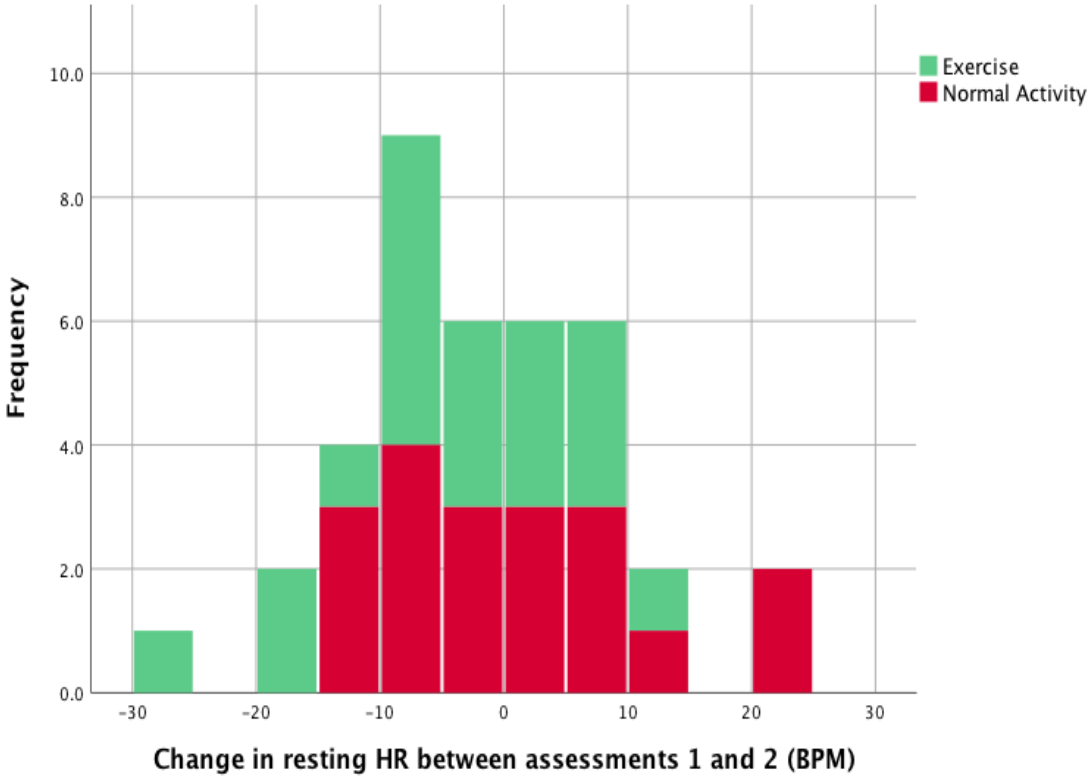


Figure 20: Stacked histogram of change in resting HR between 1st and 2nd assessments

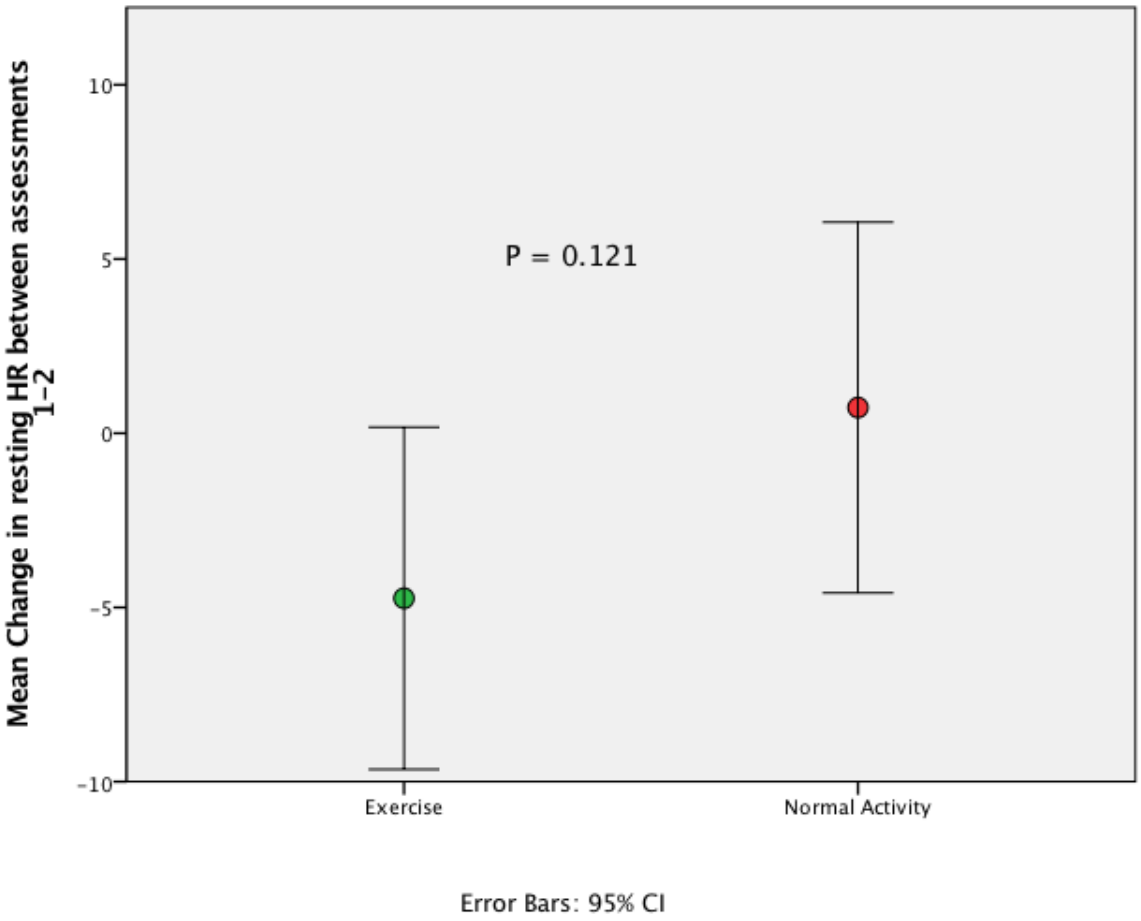


Figure 21: Mean change in resting HR (BPM) between 1st and 2nd assessments, exercise vs. normal activity

Change in heart rate increase post exercise between assessments 1 and 2

Figure 22 is a stacked histogram displaying the normal distribution of the change in the heart rate increase post exercise between the 1st and 2nd assessment. Mean change in BPM between the exercise and the normal activity groups was 2.42 BPM (SD 13.5) and 0.11 BPM (SD 13.7) respectively. Data were analysed with independent samples T-Test and the difference of 2.31 BPM between groups was not significant, p = 0.897. Figure 23 displays the mean heart rate increase post exercise between the 1st and 2nd assessment for exercise and normal activity arms with bars displaying the 95% CI.

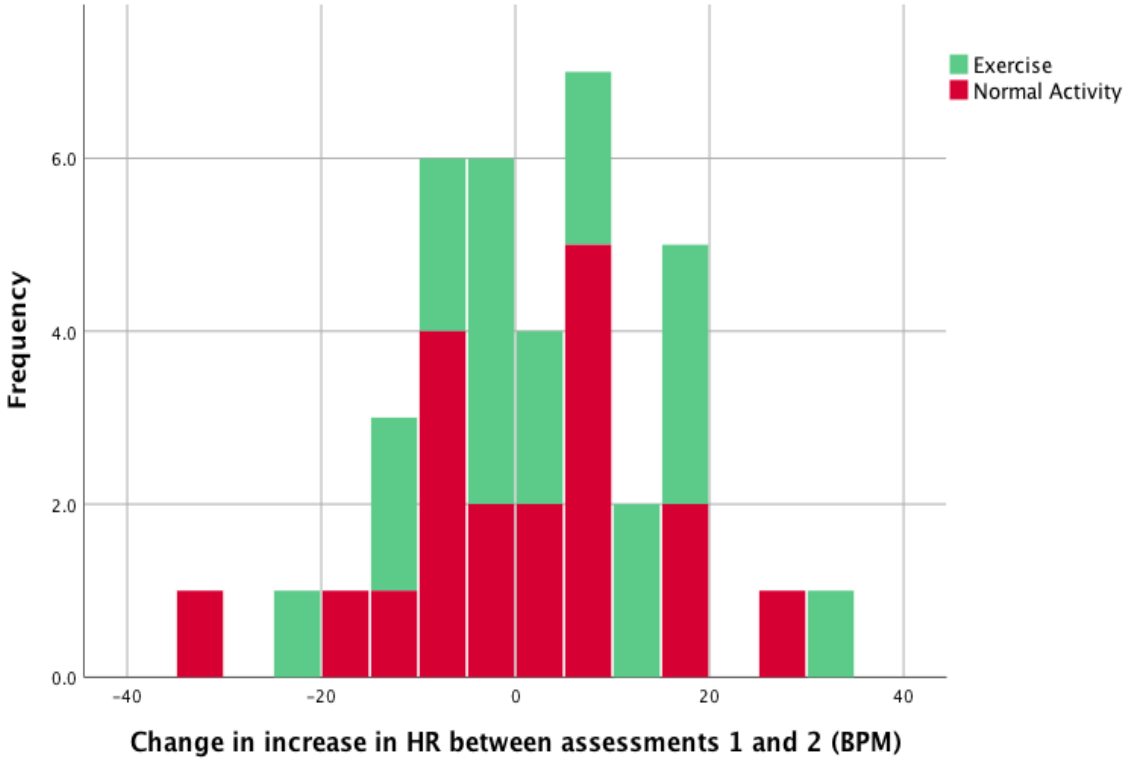


Figure 22: Stacked histogram of change in post exercise HR between 1st and 2nd assessments

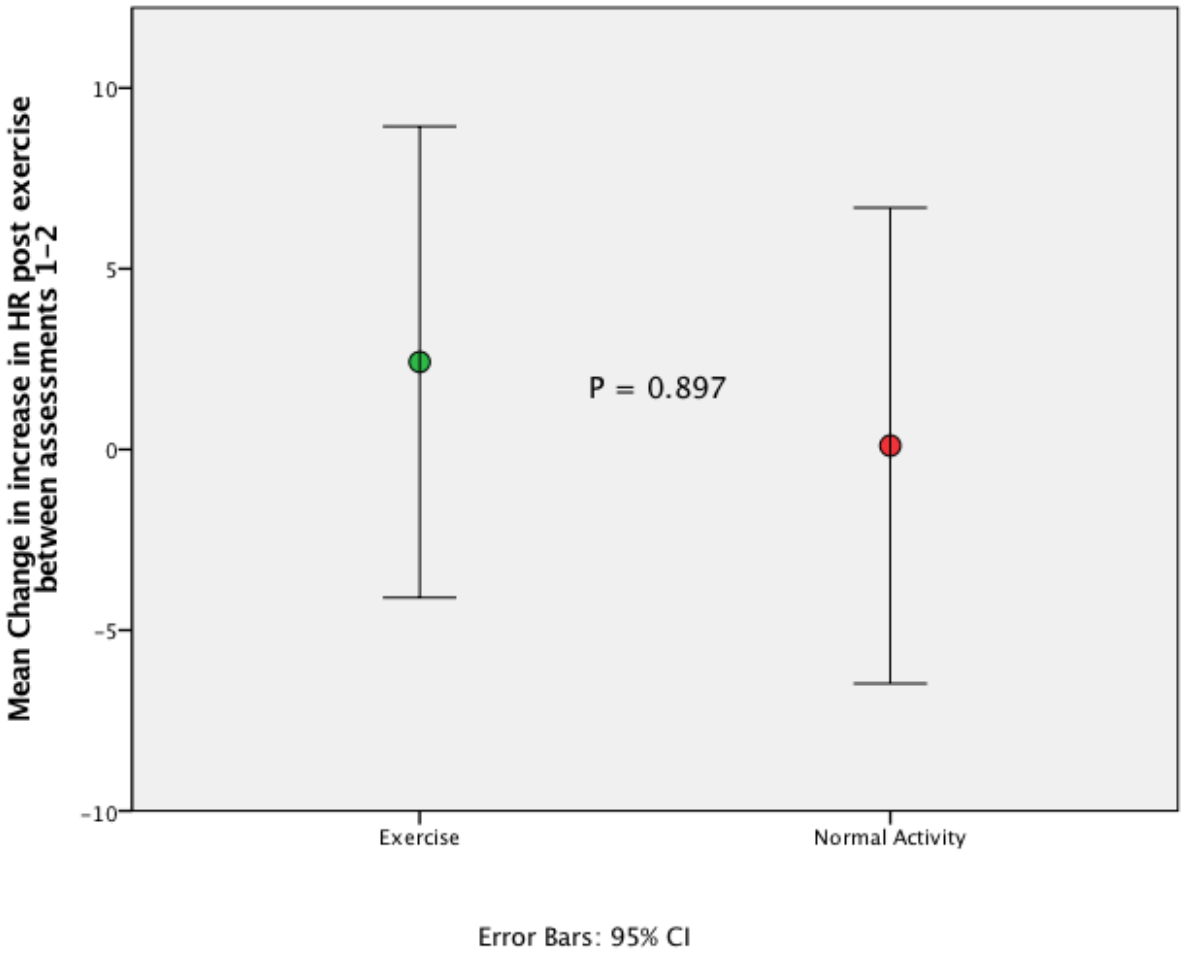


Figure 23: Mean change in post exercise HR (BPM) between 1st and 2nd assessments, exercise vs. normal activity

Change in heart rate recovery post rest between assessments 1 and 2

Figure 24 is a stacked histogram displaying the normal distribution of the change in the heart rate recovery post 1-minute rest between the 1st and 2nd assessment. Mean change in BPM between the exercise and the normal activity groups was 3.11 BPM (SD 13.7) and -1.95 BPM (SD 12.2) respectively. Data were analysed with independent samples T-Test and the difference of 5.06 BPM between groups was not significant, p = 0.237. Figure 25 displays the mean change in heart rate recovery post 1-minute rest between the 1st and 2nd assessment walked for exercise and normal activity arms with bars displaying the 95% CI.

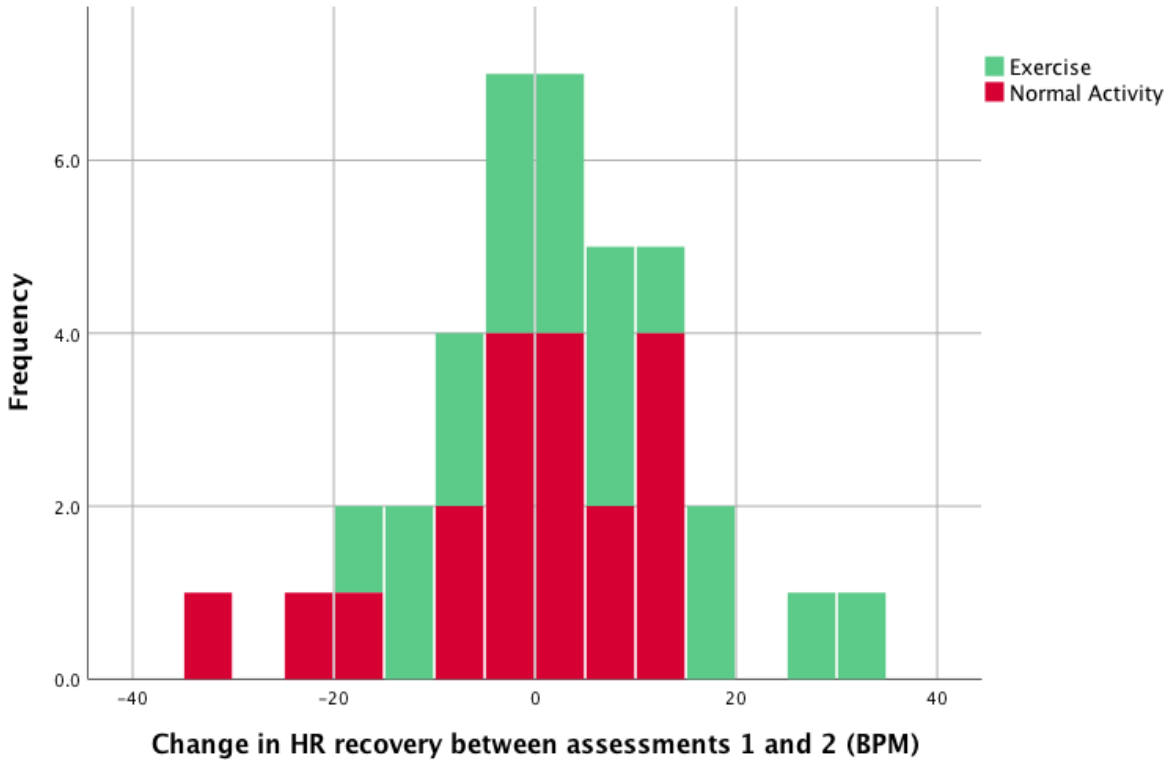


Figure 24: Stacked histogram change in HR recovery between 1st and 2nd assessments

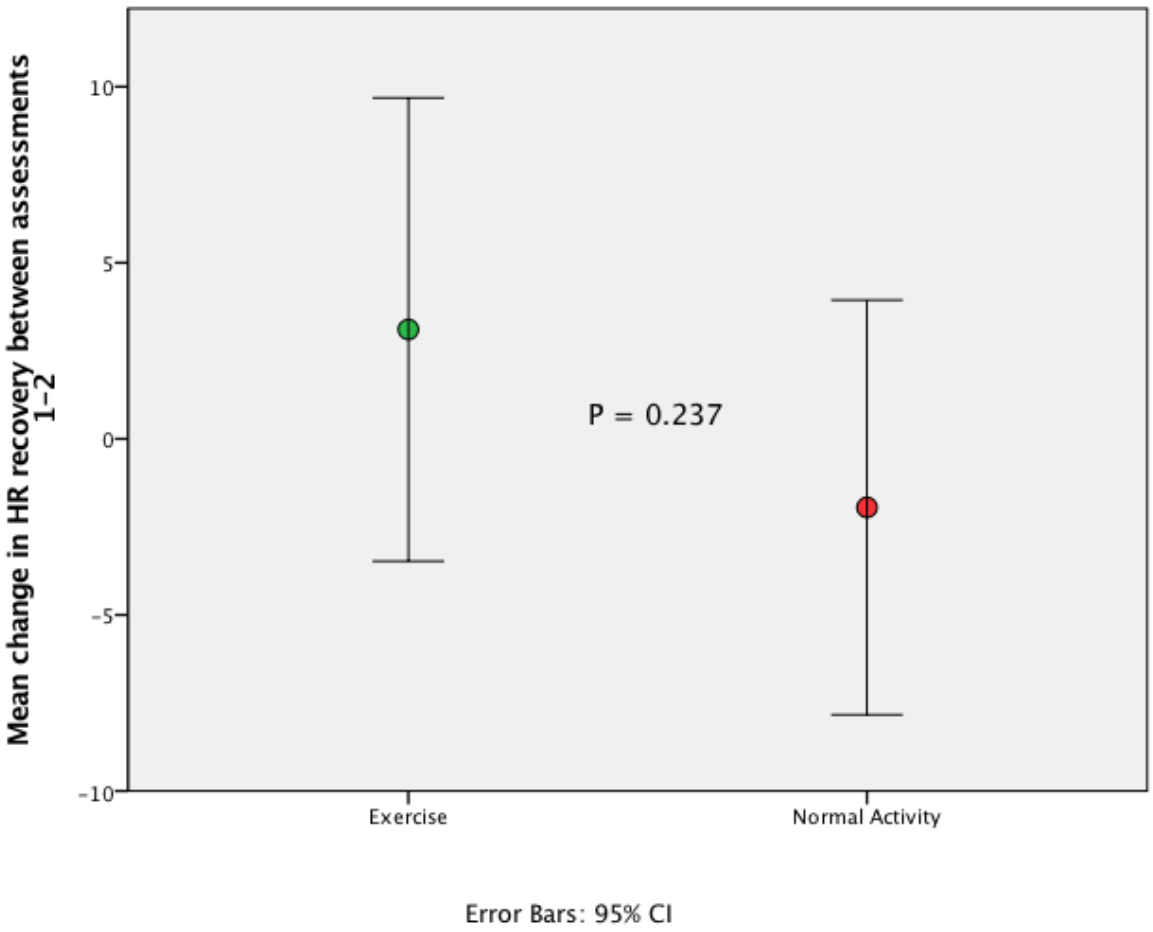


Figure: 25 Mean change in HR recovery (BPM) between 1st and 2nd assessments, exercise vs. normal activity

Change in resting heart rate between assessments 2 and 3

Figure 26 is a stacked histogram displaying the normal distribution of the change in the resting heart rate between the 2nd and 3rd assessment. Mean change in BPM between the exercise and the normal activity groups was 0.38 BPM (SD 8.6) and 0.75 BPM (SD 10.8) respectively. Data were analysed with independent samples T-Test and the difference of 0.37 BPM between groups was not significant, p = 0.914. Figure 27 displays the mean change in resting HR between the 2nd and 3rd assessments for exercise and normal activity arms with bars displaying the 95% CI.

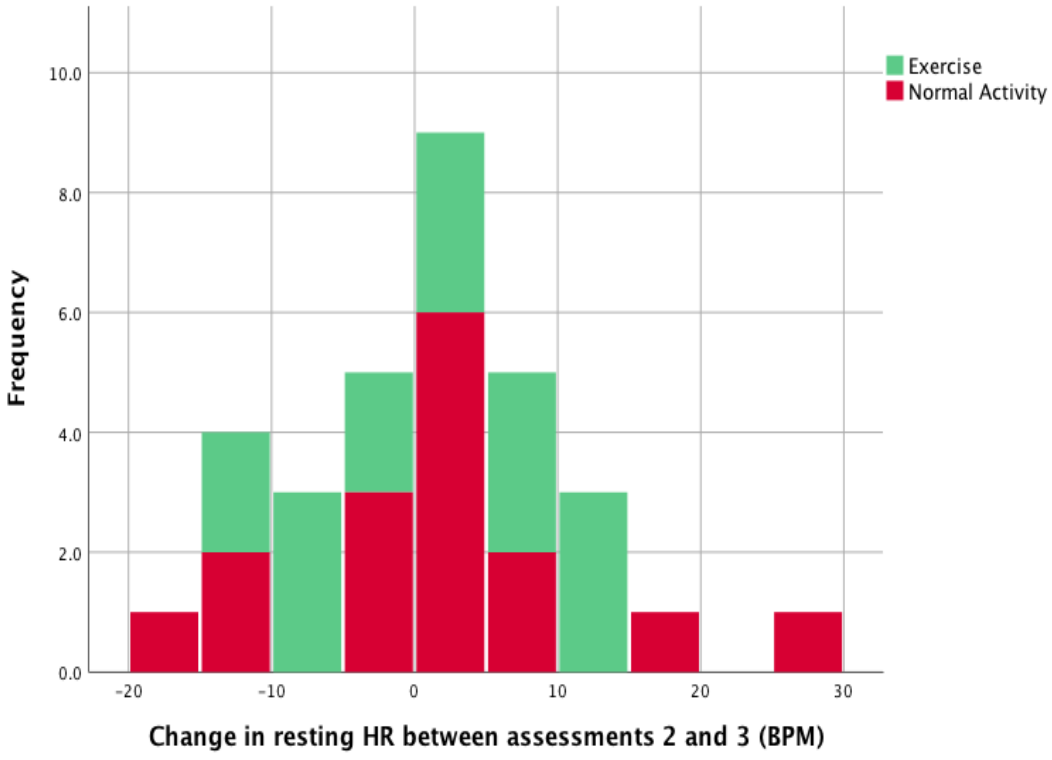


Figure 26: Stacked histogram of change in resting HR between 2nd and 3rd assessments

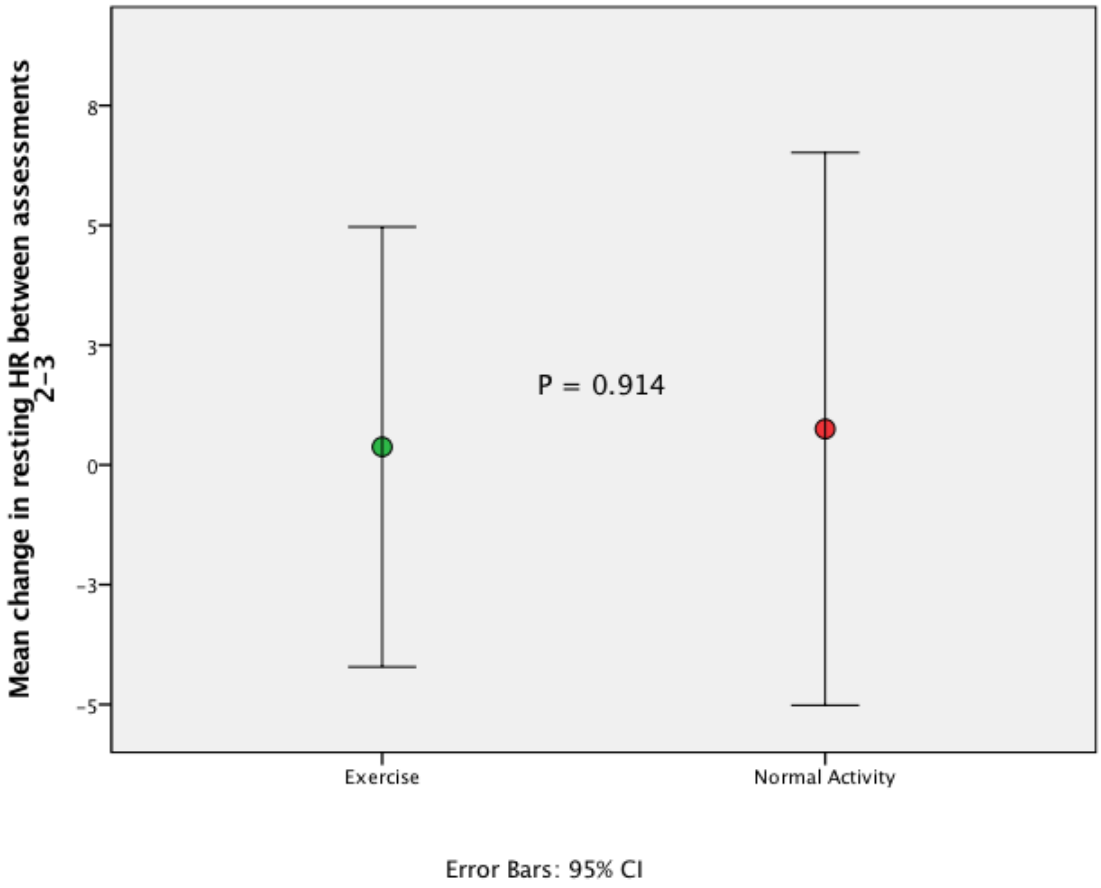


Figure 27: Mean change in resting HR (BPM) between 2nd and 3rd assessments exercise vs. normal activity

Change in heart rate increase post exercise between assessments 2 and 3

Figure 28 is a stacked histogram displaying the normal distribution of the change in the heart rate increase post exercise between the 2nd and 3rd assessment. Mean change in BPM between the exercise and the normal activity groups was -0.38 BPM (SD 12.1) and -2.69 BPM (SD 9.9) respectively. Data were analysed with dependent samples T-Test and the difference of 3.07 BPM between groups was not significant, p = 0.557. Figure 29 displays the mean change in post exercise HR between the 2nd and 3rd assessments for exercise and normal activity arms with bars displaying the 95% CI.

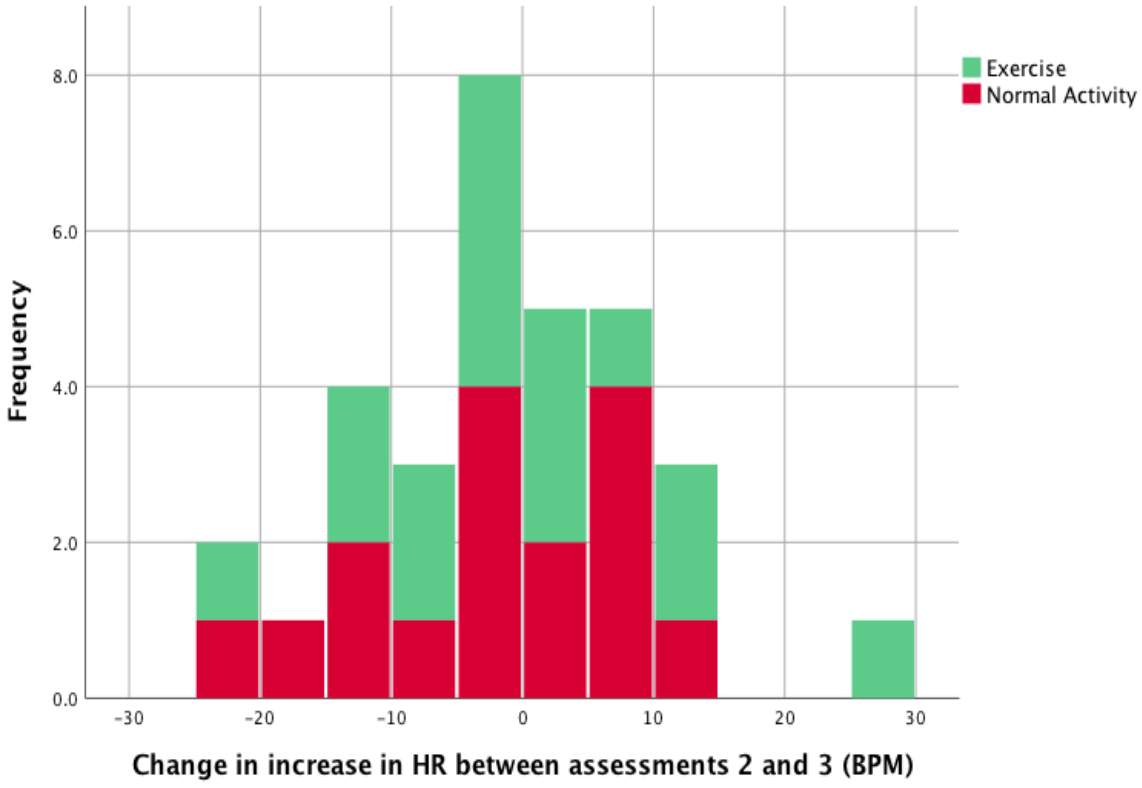


Figure 28: Stacked histogram of change in post exercise HR between 2nd and 3rd assessments

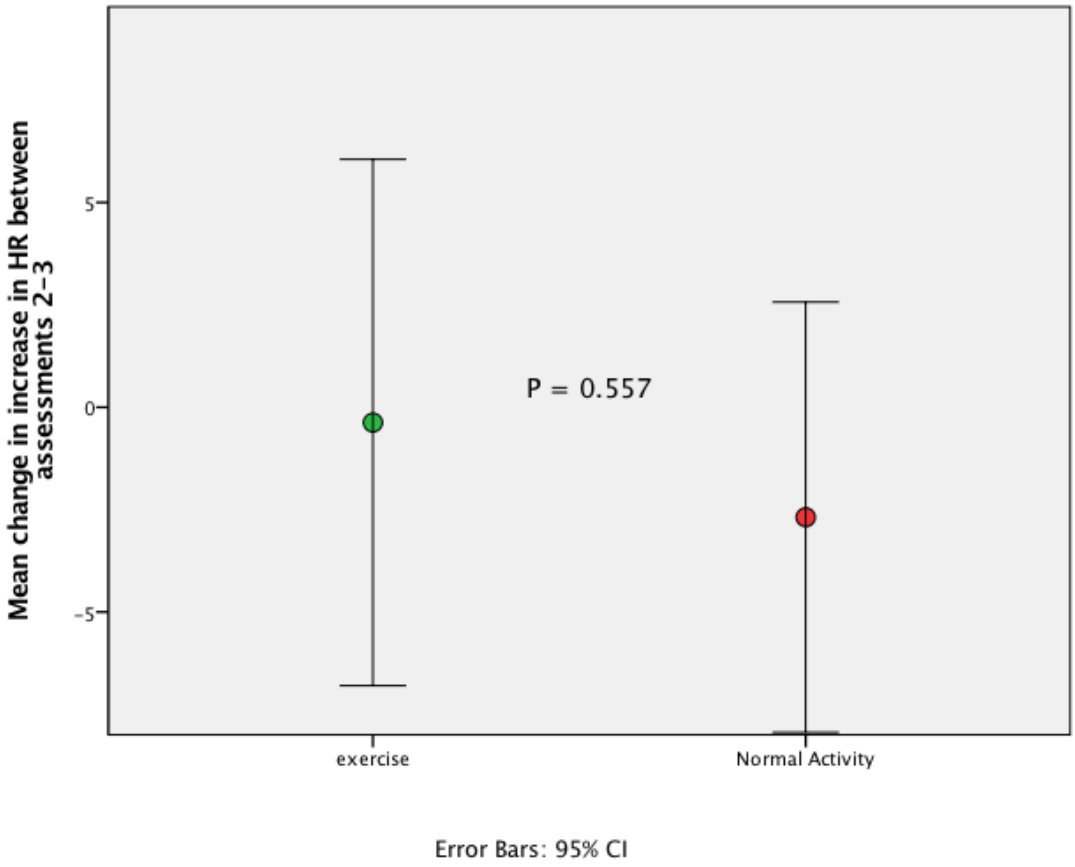


Figure 29: Mean change in post exercise HR (BPM) between 2nd and 3rd assessments, exercise vs. normal activity

Change in heart rate recovery post rest between assessments 2 and 3

Figure 30 is a stacked histogram displaying the normal distribution of the change in the heart rate recovery post 1-minute rest between the 2nd and 3rd assessment. Mean change in BPM between the exercise and the normal activity groups was -0.56 BPM (SD 10.0) and -0.44 BPM (SD 10.0) respectively. Data were analysed with independent samples T-Test and the difference of 0.12 BPM between groups was not significant, p = 0.972. Figure 31 displays the mean change in HR recovery between 2nd and 3rd assessments for exercise and normal activity arms with bars displaying the 95% CI.

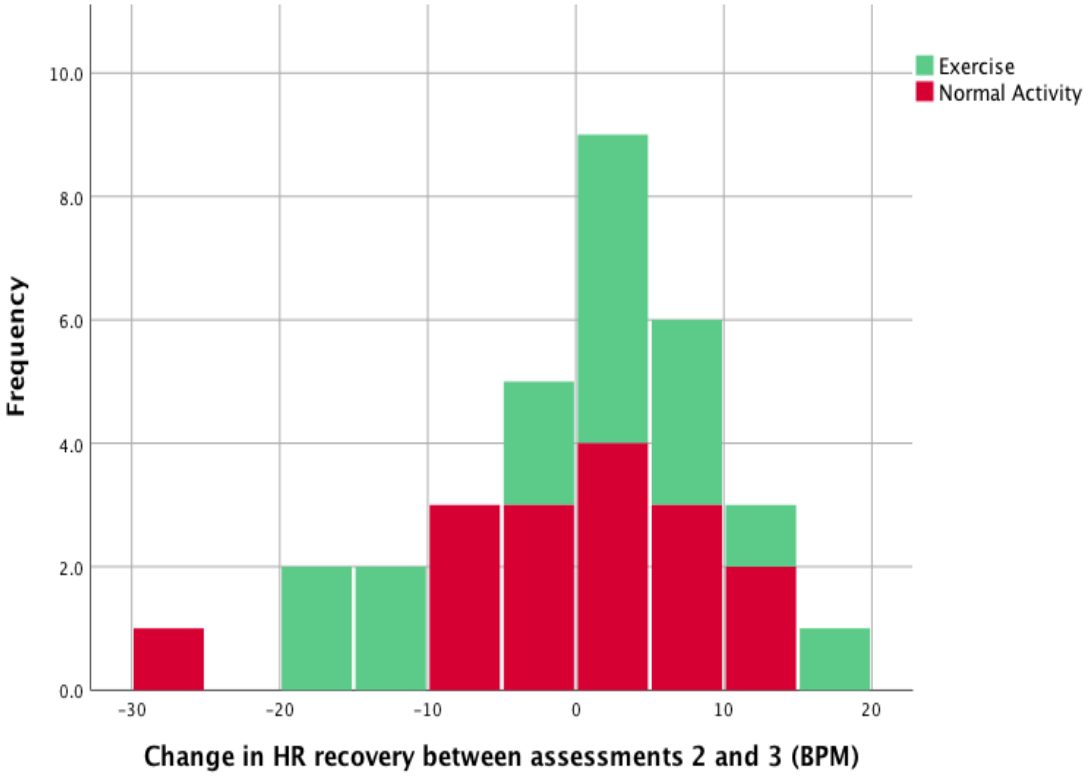


Figure 30: Stacked histogram of change in post HR recovery between 2nd and 3rd assessments

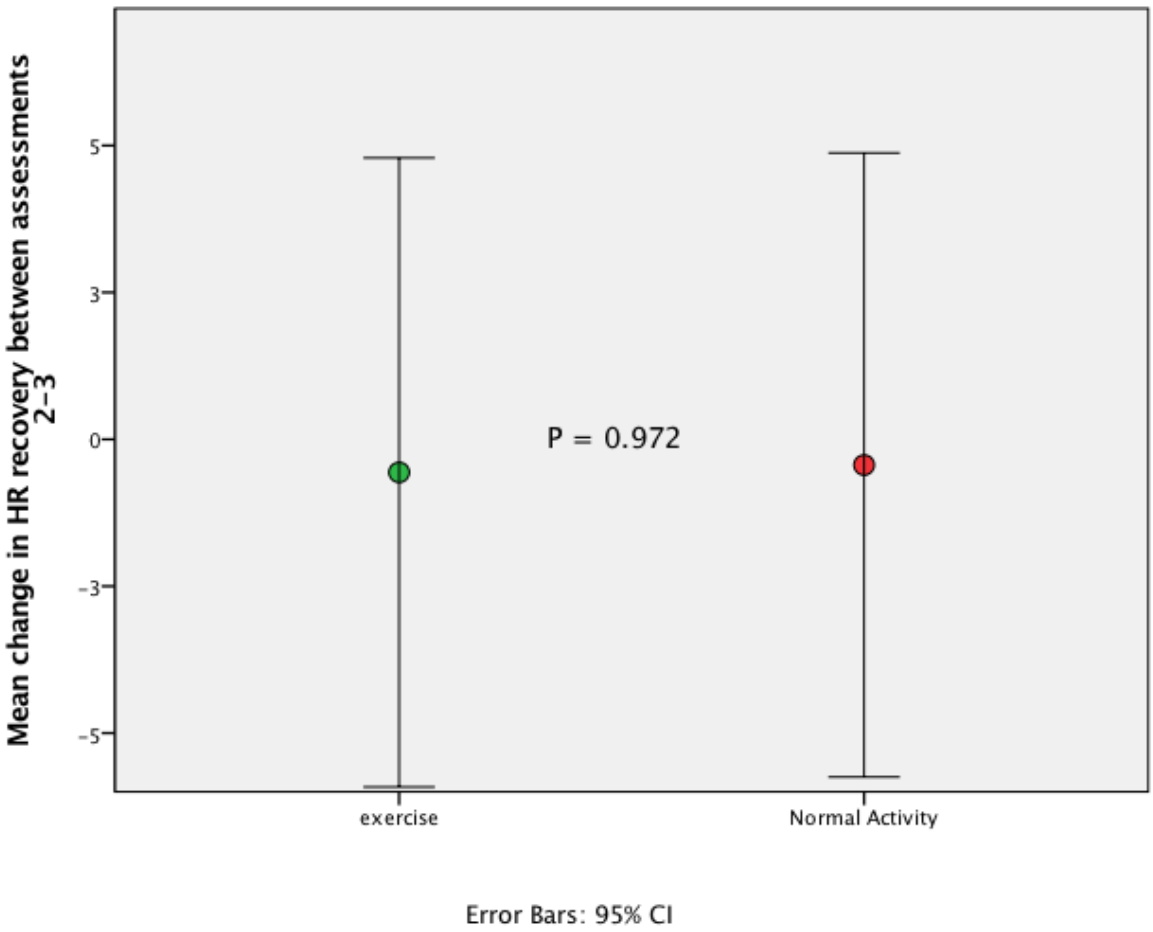


Figure 31: Mean change in HR recovery (BPM) between 2nd and 3rd assessments, exercise vs. normal activity

Change in resting heart rate between assessments 1 and 3

Figure 32 is a stacked histogram displaying the normal distribution of the change in the resting heart rate between the 1st and 3rd assessment. Mean change in BPM between the exercise and the normal activity groups was -2.94 BPM (SD 12.1) and 2.19 BPM (SD 13.8) respectively. Data were analysed with independent samples T-Test and the difference of 5.13 BPM between groups was not significant, p = 0.272. Figure 33 displays the mean change in resting HR between 1st and 3rd assessments for exercise and normal activity arms with bars displaying the 95% CI.

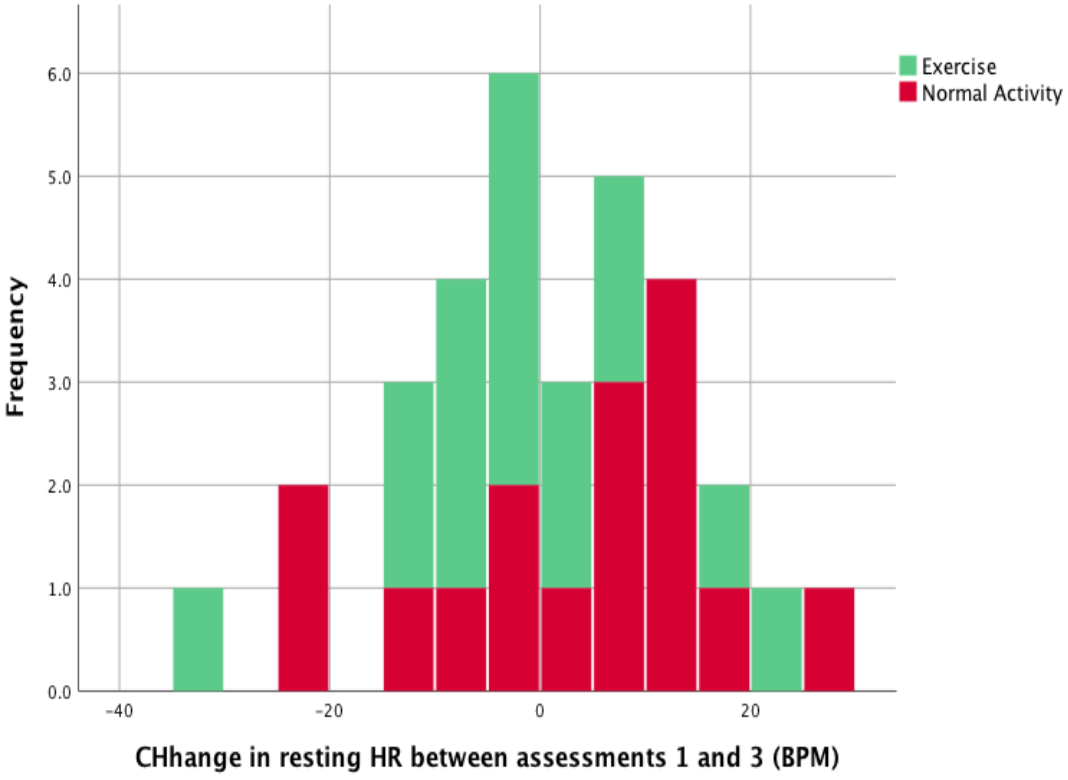


Figure 32: Stacked histogram of change in resting HR between 1st and 3rd assessments

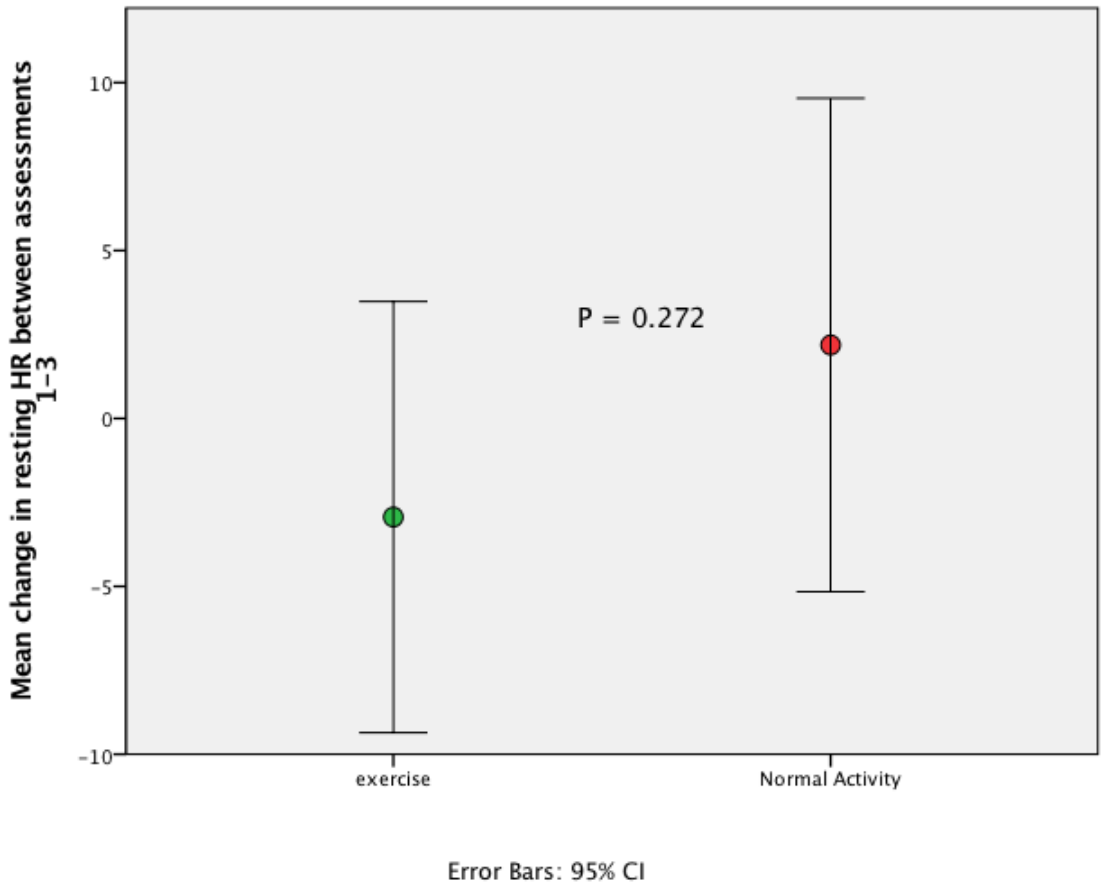


Figure 33: Mean change in resting HR (BPM) between 1st and 3rd assessments, exercise vs. normal activity

Change in heart rate increase post exercise between assessments 1 and 3

Figure 34 is a stacked histogram displaying the non-normal distribution of the change in the heart rate increase post exercise between the 1st and 3rd assessment. Median change in BPM between the exercise and the normal activity groups was -0.44 BPM (SD 10.2) and -1.44 BPM (SD 11.5) respectively. Data were analysed with the Mann-Whitney U Test and the difference between groups was not significant, p = 0.748. Figure 35 displays the median change in the heart rate increase post exercise between the 1st and 3rd assessment for exercise and normal activity arms with bars displaying the 95% CI.

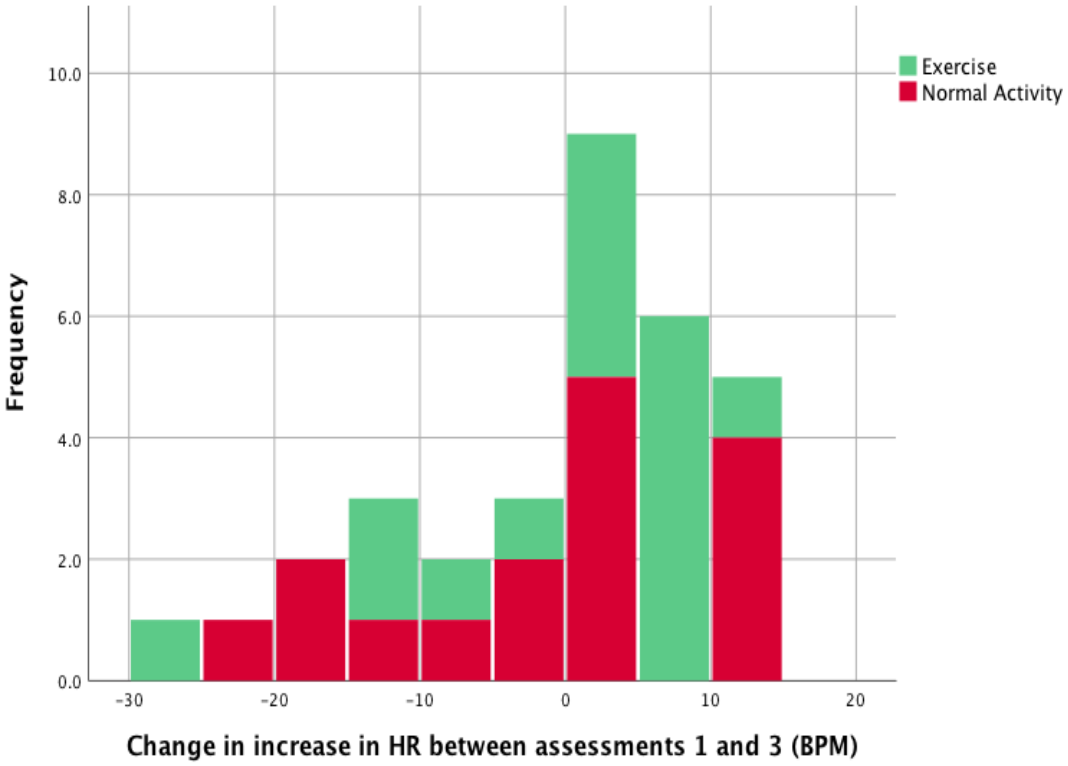


Figure 34: Stacked histogram of change in post exercise HR between 1st and 3rd assessments

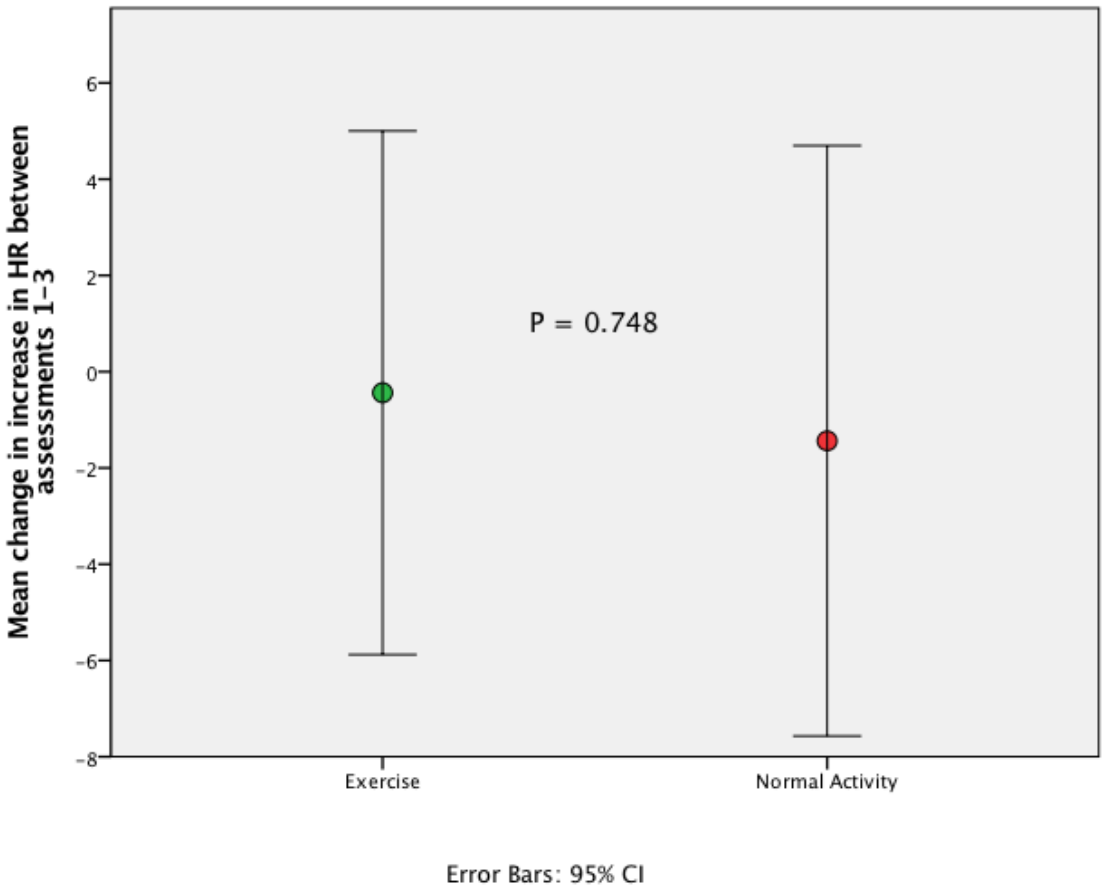


Figure 35: Mean change in post exercise HR (BPM) between 1st and 3rd assessments, exercise vs. normal activity

Change in heart rate recovery post rest between assessments 1 and 3

Figure 36 is a stacked histogram displaying the normal distribution of the change in the heart rate recovery post 1-minute rest between the 1st and 3rd assessment. Mean change in BPM between the exercise and the normal activity groups was -0.13 BPM (SD 8.4) and -2 BPM (SD 9.7) respectively. Data were analysed with independent samples T-Test and the difference of 2.13 BPM between groups was not significant, p = 0.584. Figure 37 displays the mean change in the heart rate recovery post 1-minute rest between the 1st and 3rd assessment for exercise and normal activity arms with bars displaying the 95% CI.

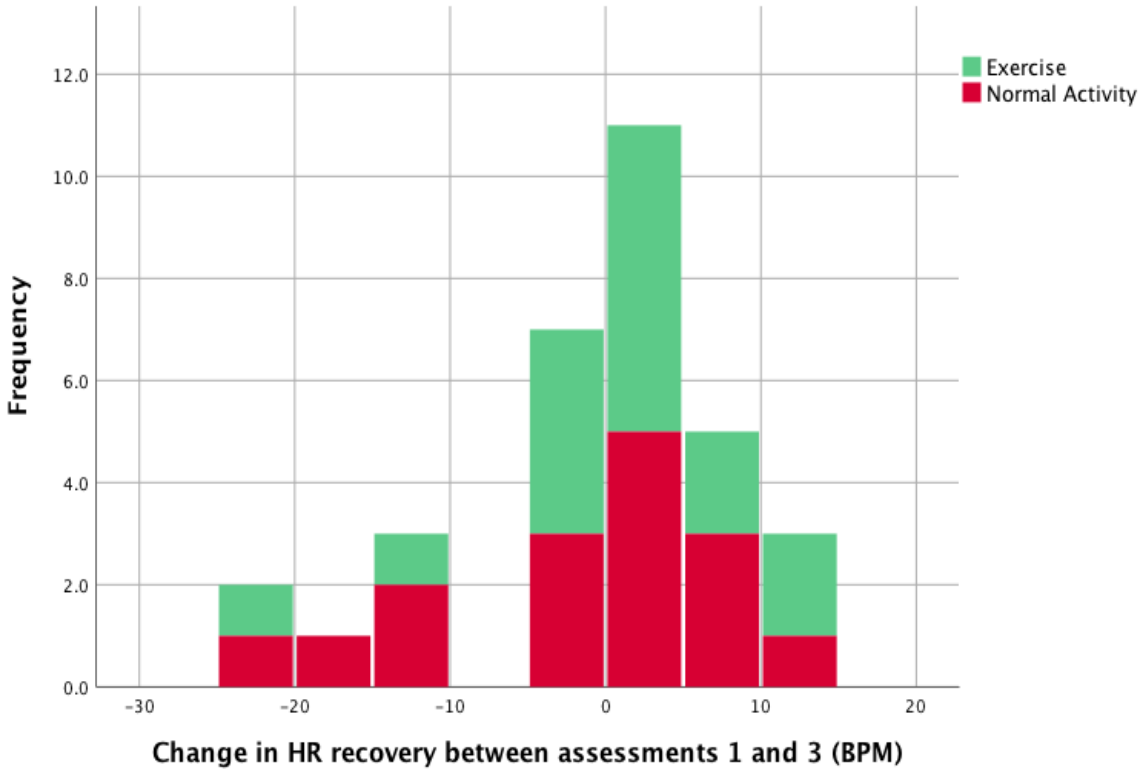


Figure 36: Stacked histogram of change in HR recovery between 1st and 3rd assessments

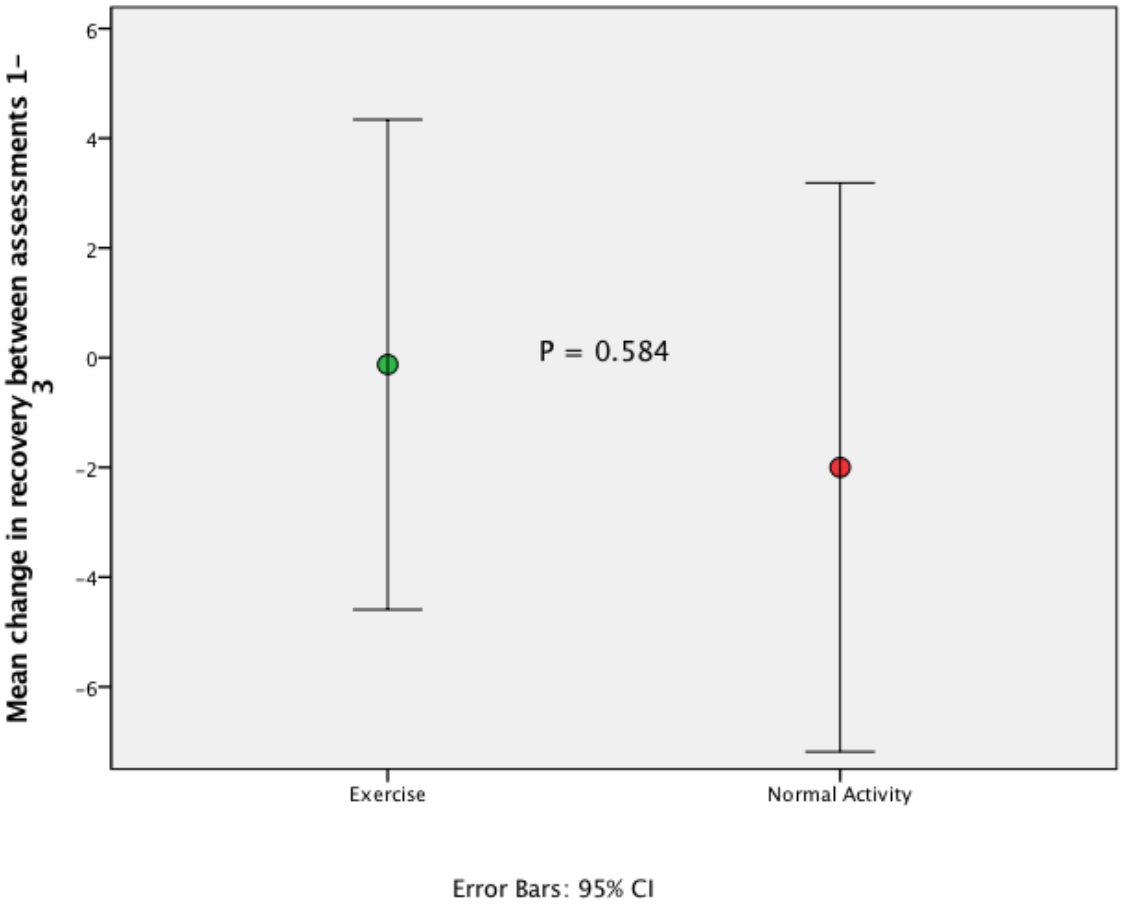


Figure 37: Mean change in post exercise HR (BPM) between 1st and 3rd assessments, exercise vs. normal activity

3.7 Change in BORG score between assessments

Participants were asked to rate their exertion at the end of the 6MWT on the 0-10 BORG scale at each assessment. Figures 38, 39 and 40 show the absolute values of the participants at each assessment.

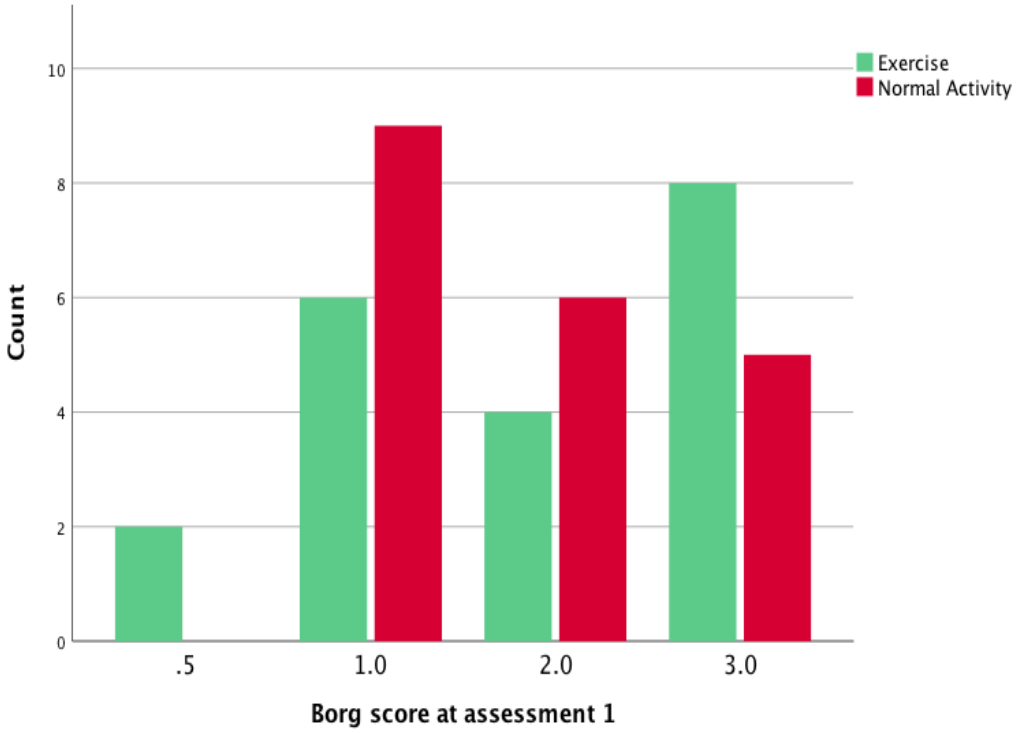


Figure 38: Clustered bar chart of Borg scores at assessment 1

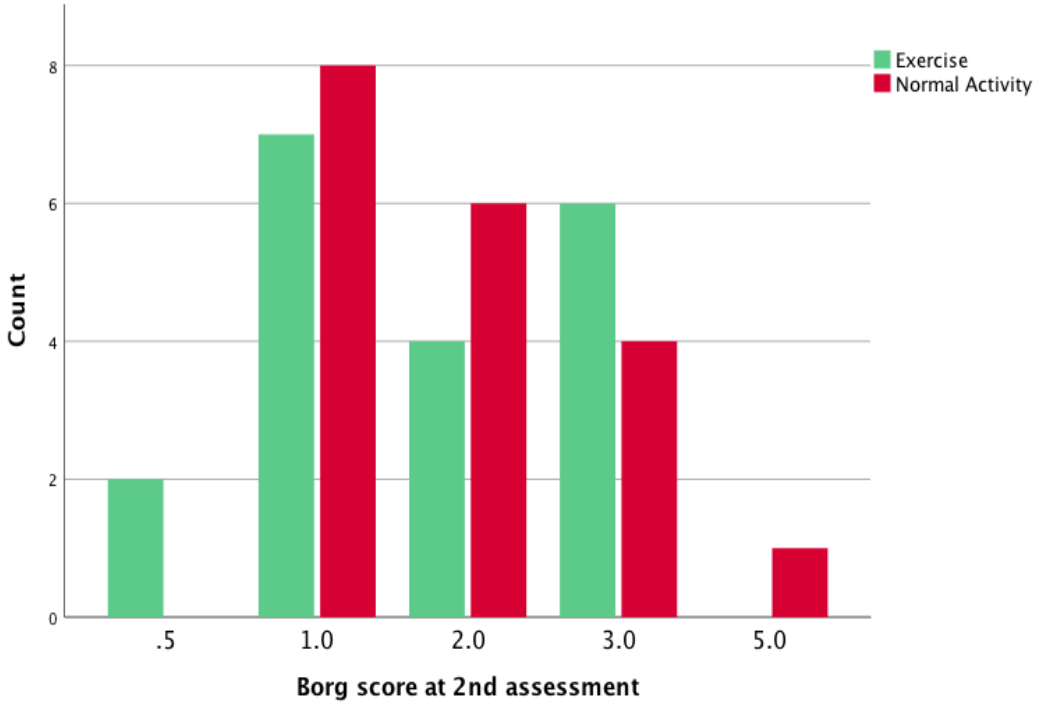


Figure 39: Clustered bar chart of Borg scores at assessment 2

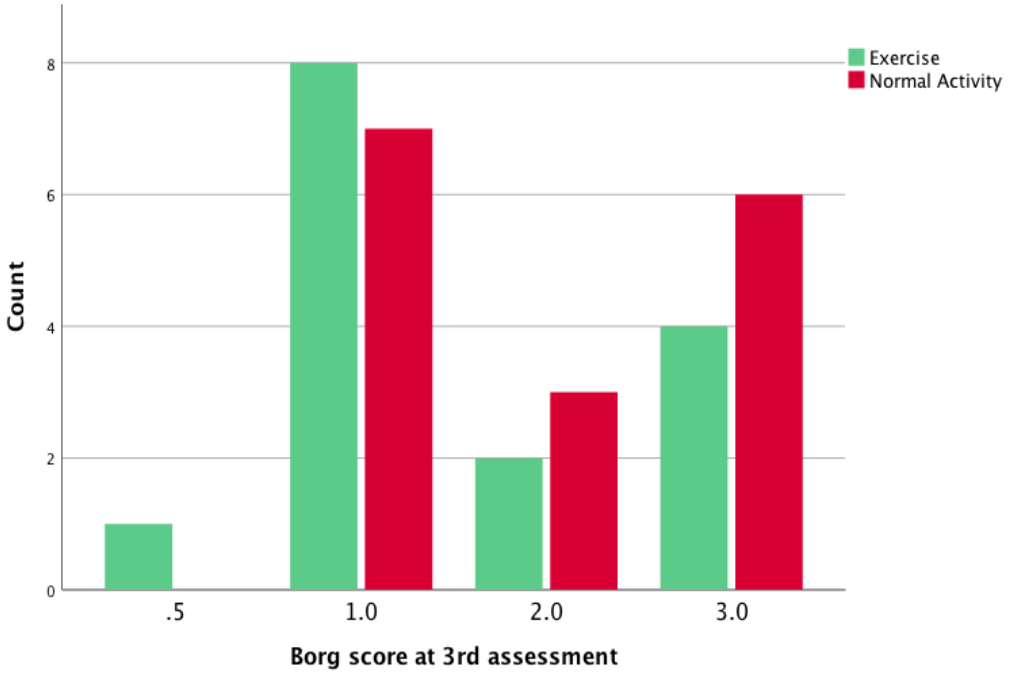


Figure 40: Clustered bar chart of Borg scores at assessment 3

Change In BORG scale between assessments 1-2

Figure 41 is a stacked histogram of the change in BORG score between assessments 1 and 2 and displays a non-normal distribution. Median values for both groups were 0.0. Mann-Whitney U test revealed a p value of 0.027. Figure 42 displays a Box and whisker plot of this data.

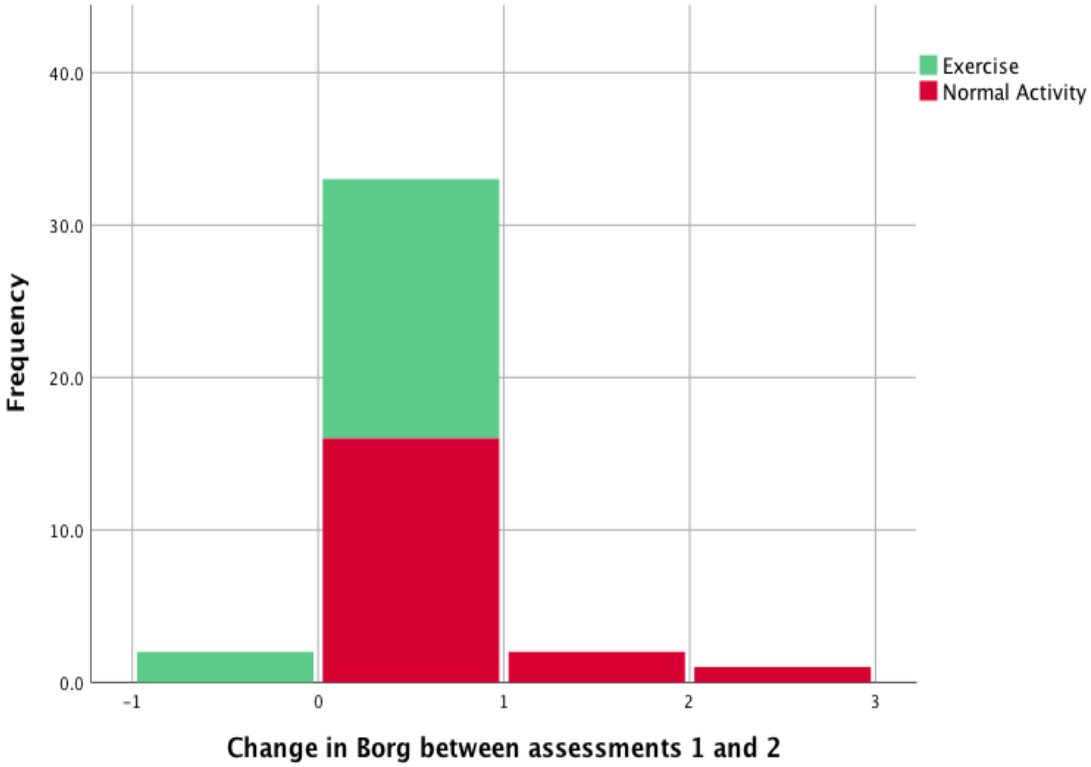


Figure 41: Stacked histogram of change in BORG score between 1st and 2nd assessments

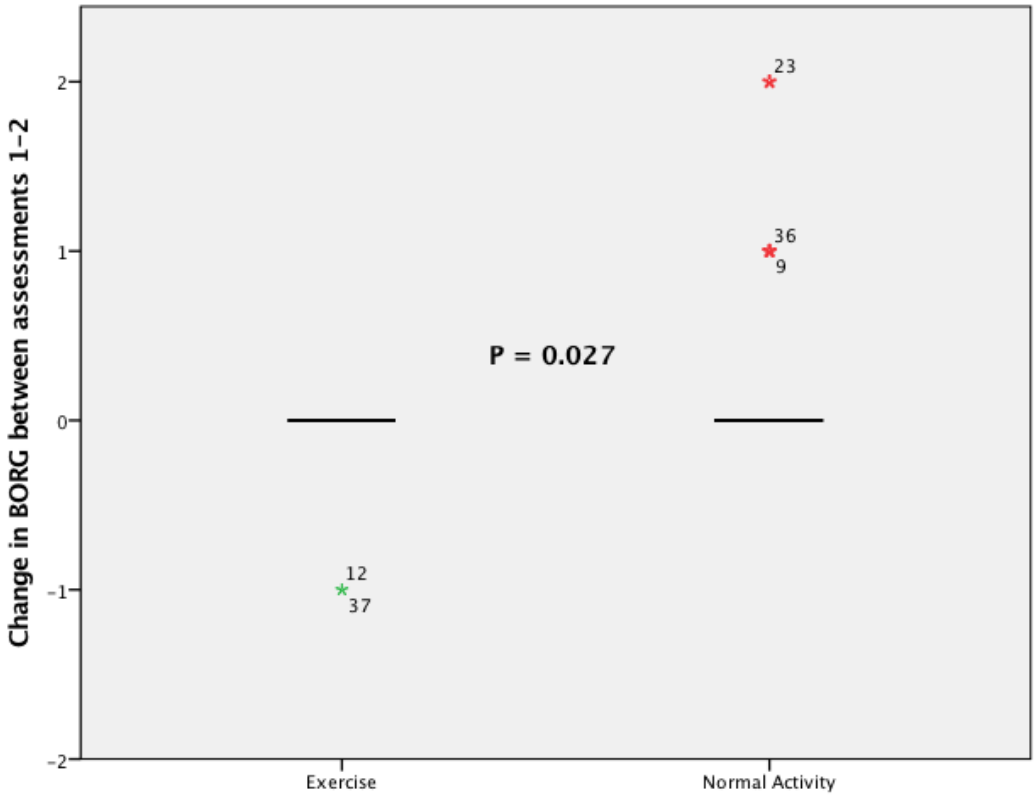


Figure 42: Box and whisker plot of change in BORG score between 1st and 2nd assessments, exercise vs. normal activity

Change In BORG scale between assessments 2-3

Figure 43 is a stacked histogram of the change in BORG scale between assessments 2 and 3 and displays a non-normal distribution. Median values for both groups were 0.0. Mann-Whitney U test revealed a p value of 0.556. Figure 44 displays a Box and whisker plot of this data.

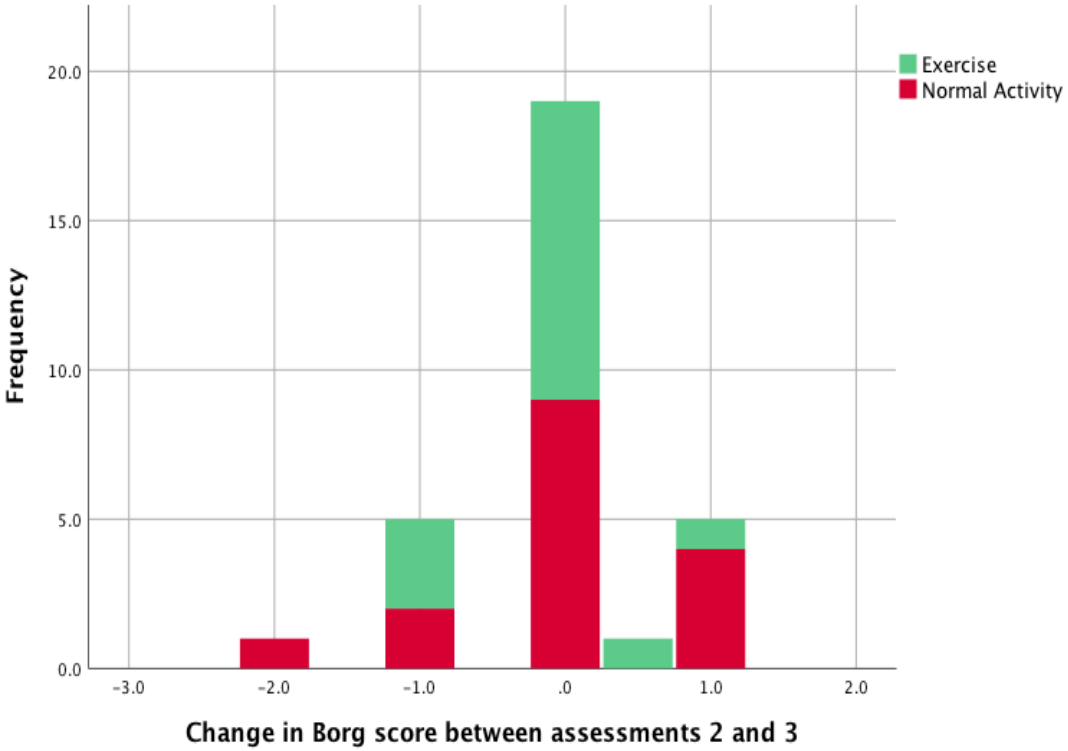


Figure 43: Stacked histogram of change in BORG score between 2nd and 3rd assessments

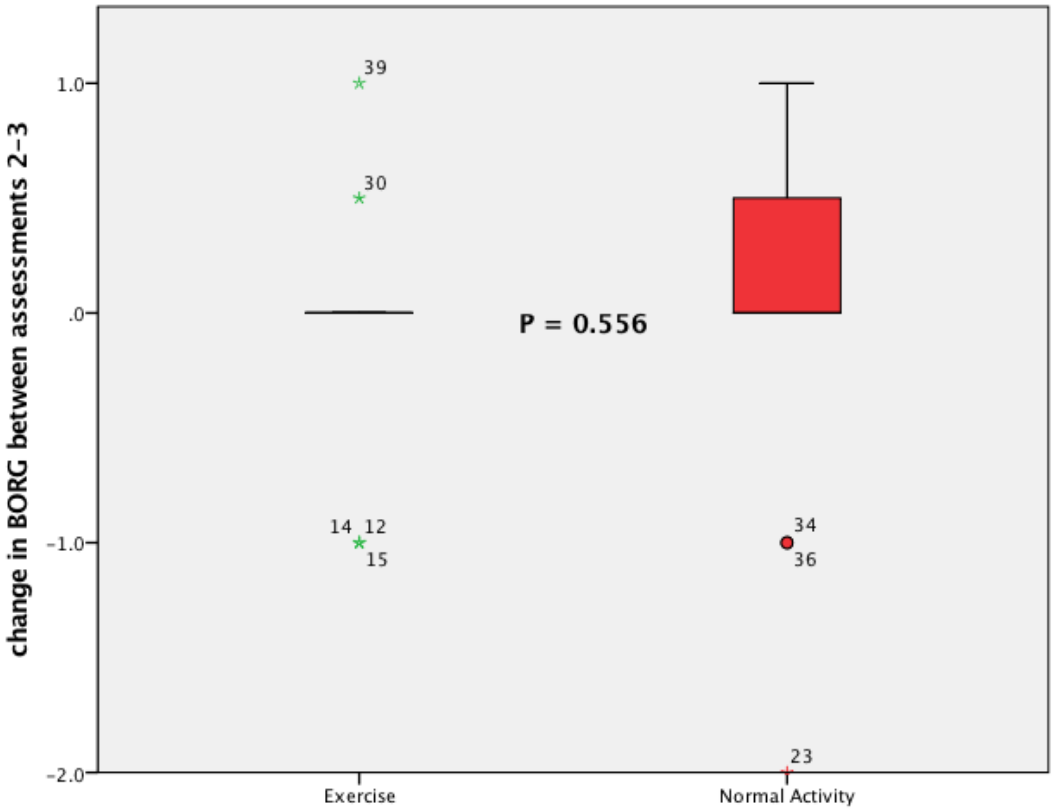


Figure 44: Box and whisker plot of change in BORG score between 2nd and 3rd assessments, exercise vs. normal activity

Change In BORG scale between assessments 1-3

Figure 45 is a stacked histogram of the change in BORG scale between assessments 1 and 3 and displays a non-normal distribution. Median values for both groups were 0.0. Mann-Whitney U test revealed a p value of 0.115. Figure 46 displays a Box and whisker plot of this data.

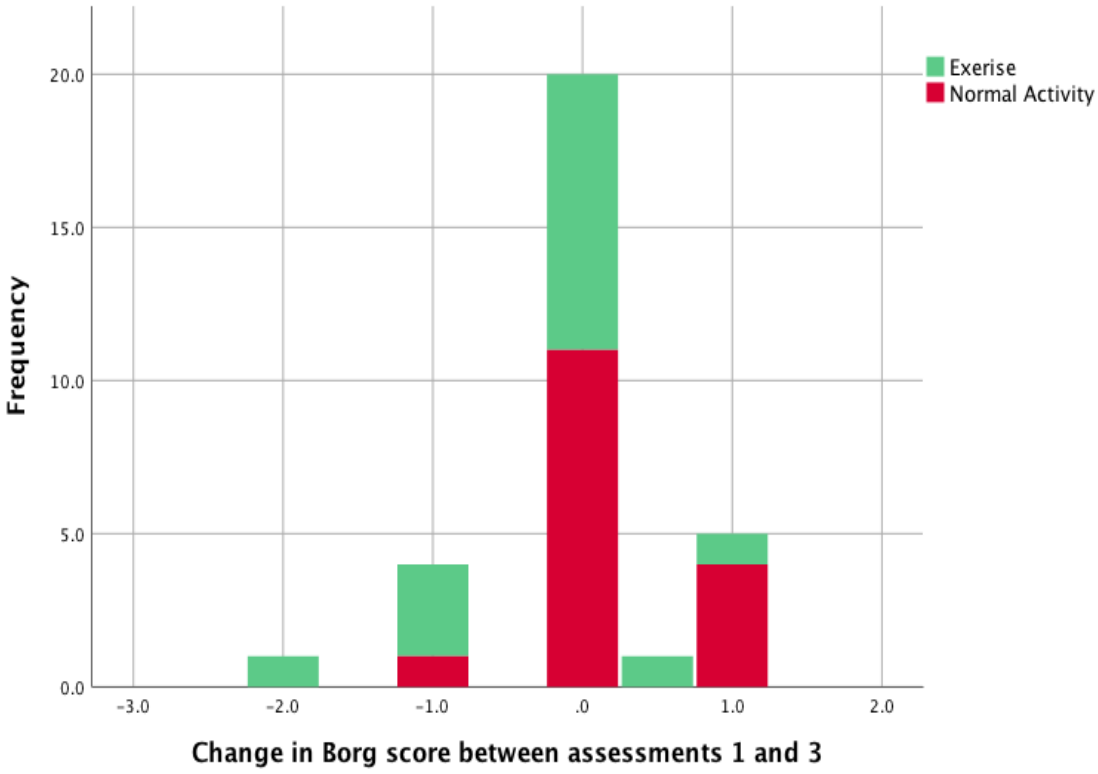


Figure 45: Staked histogram of change in BORG score between 1st and 3rd assessments

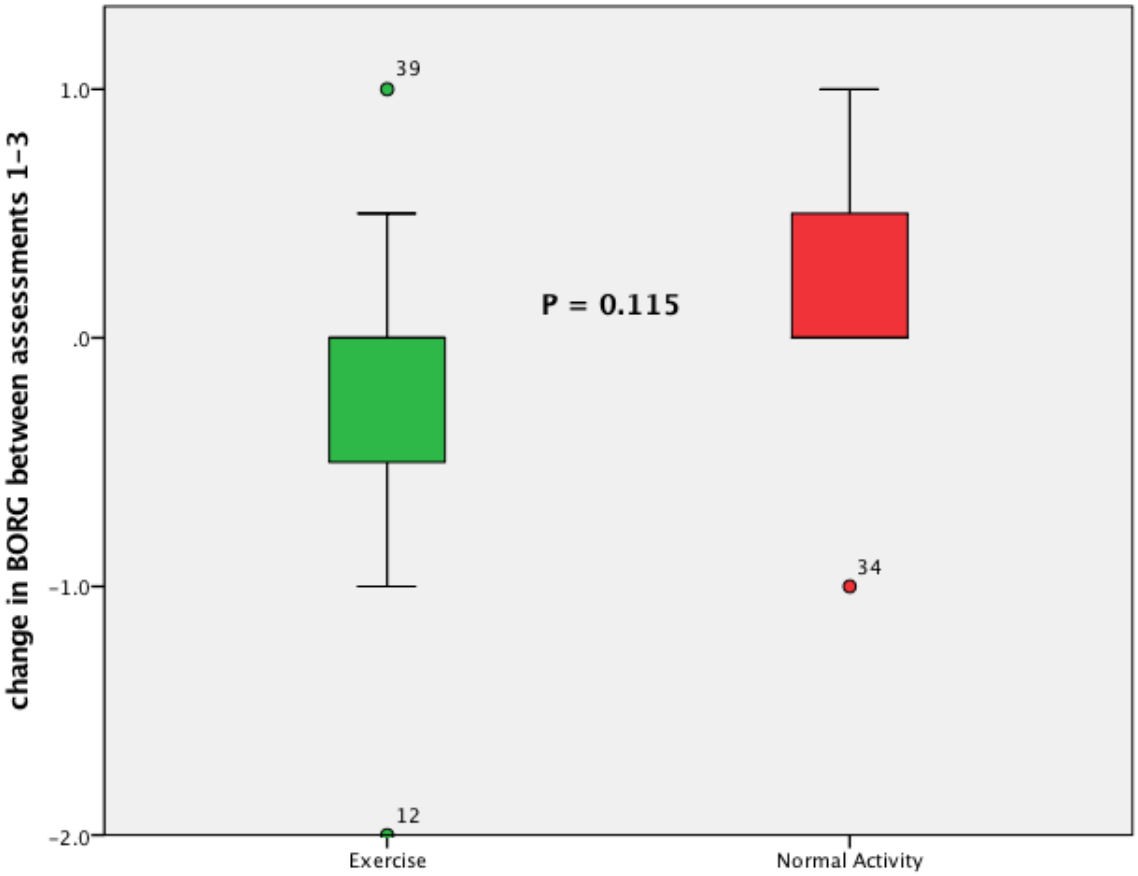


Figure 46: Box and whisker plot of change in BORG score between 1st and 3rd assessments, exercise vs. normal activity

3.8 Change in Quality of Life Scores Between Assessments

As per the methods three QOL scoring systems were used and data analysed from specific outputs from them as listed below.

- HADS
    - Depression Score
    - Anxiety Score

- EQ-5D-5L
    - Total Health Today

- EORTC QLQ-C30
    - Combined Quality of life and Heath in last week

Change in HADS Depression Score

Total depression score ranges from 0-21 with a higher score equating to a higher level of depression. Test for normal distribution of data were a Shapiro-Wilk Test where distribution was considered normal at >0.05. Normally distributed data were analysed with independent samples T-Test and non-normally distributed data with Mann-Whitney U Test. Table 3 Displays results and Figure 49 displays the mean change between assessments with error bars showing the 95% CI. Absolute values for HADS depression are shown in figures 47, 48 & 49.

	Exercise	Normal Activity	Test	P-Value
1st to 2nd Assessment (SD)	-0.53 (2.2)	-0.05 (1.3)	M-W U* Test	0.16
2nd to 3rd Assessment (SD)	1.29 (3.3)	0.11 (2.6)	Independent Samples T**	0.24
1st to 3rd Assessment (SD)	0.41 (1.7)	- 0.56 (2.5)	Independent Samples T	0.87

* Mann-Whitney U Test ** Independent Samples T-Test

Table 3: Change in HADS Depression score

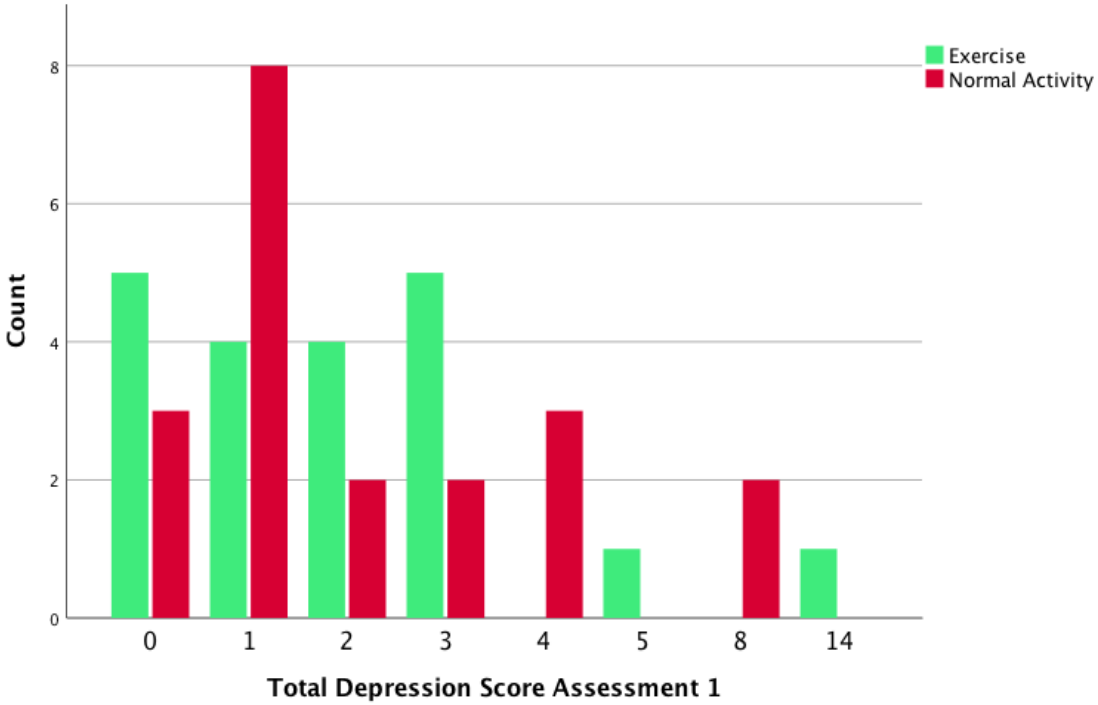


Figure 47: Clustered bar chart of depression scores at assessment 1

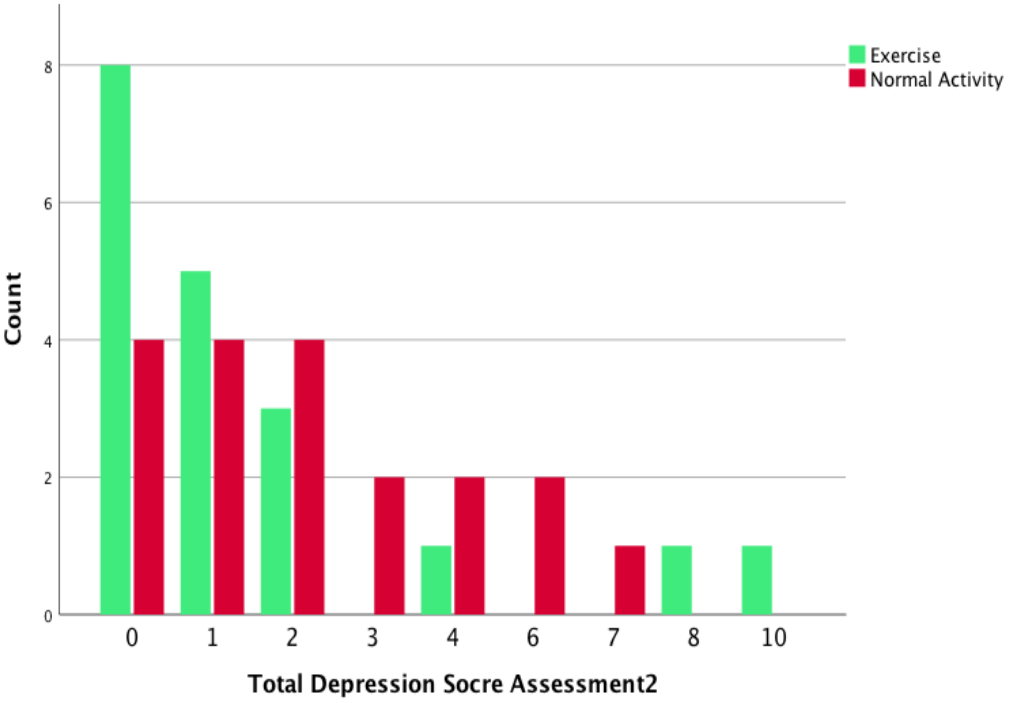


Figure 48: Clustered bar chart of depression scores at assessment 2

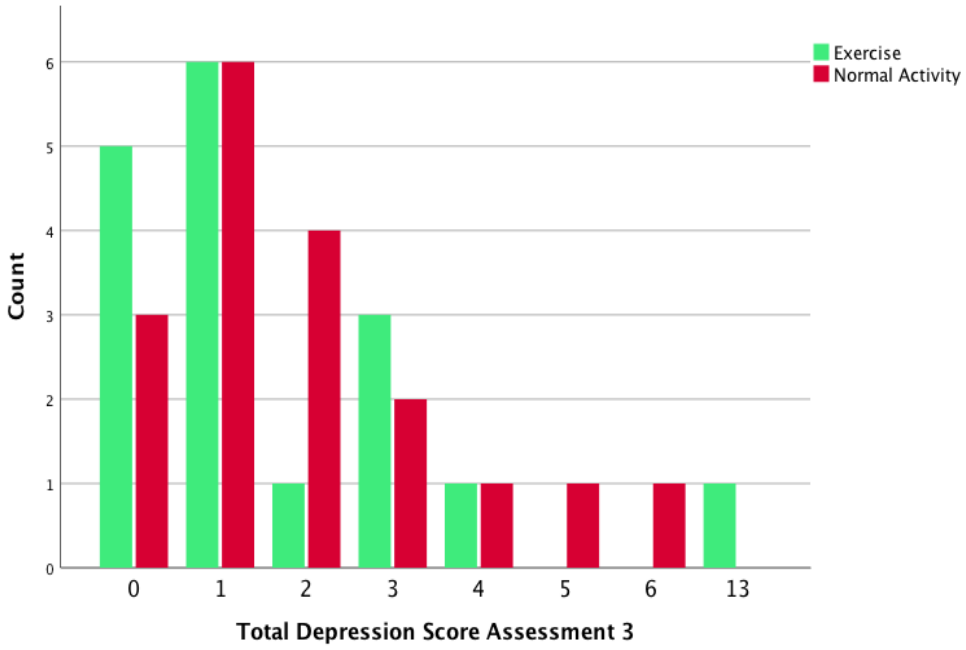


Figure 49: Clustered bar chart of depression scores at assessment 3

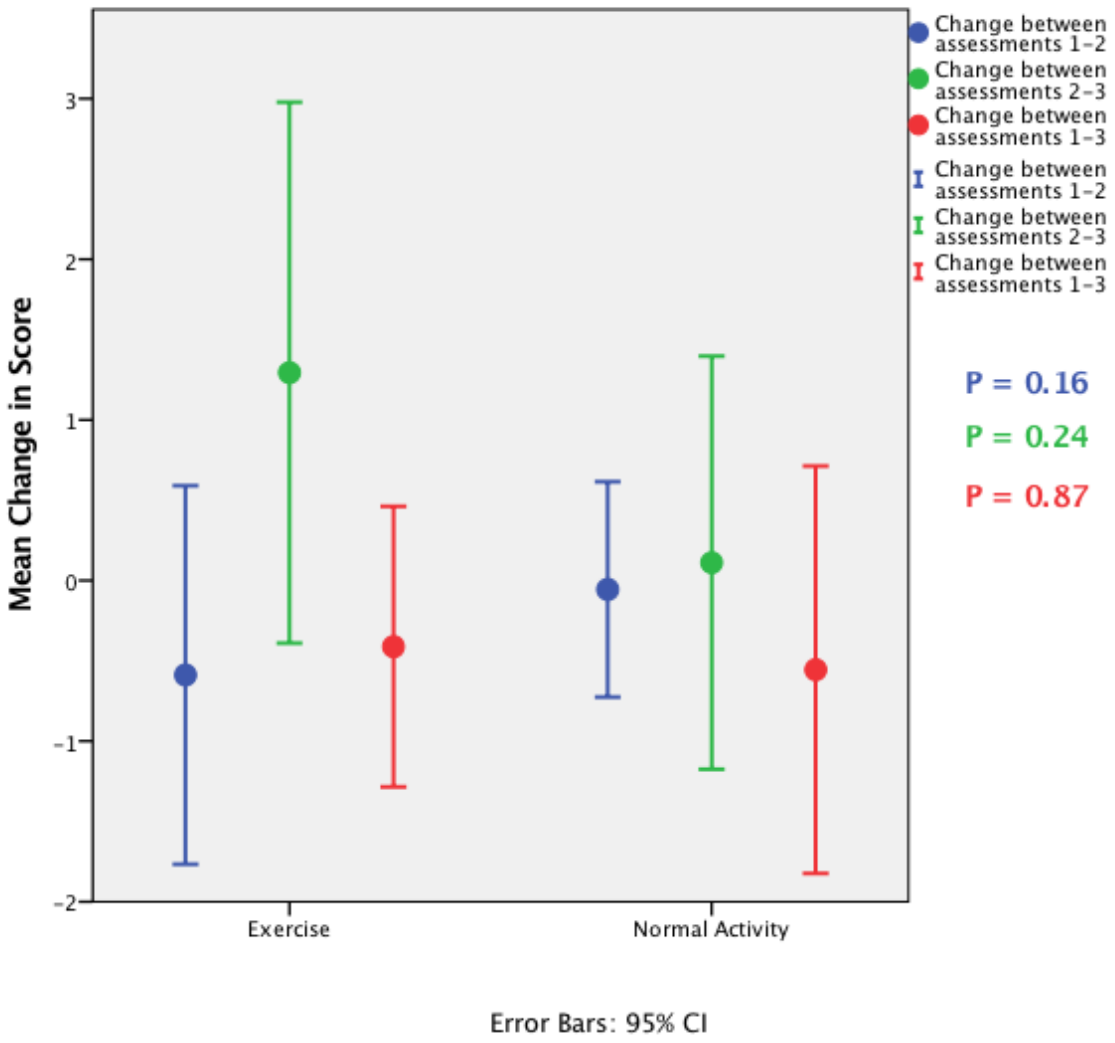


Figure 50: Mean change in HADS depression score, exercise vs. normal activity

Change in HADS Anxiety Score

Total anxiety score ranges from 0-21 with a higher score equating to a higher level of anxiety. Test for normal distribution of data were a Shapiro-Wilk Test where distribution was considered normal at >0.05. Normally distributed data were analysed with independent samples T-Test and non-normally distributed data with Mann-Whitney U Test. Table 4 Displays results and Figure 54 displays the mean change between assessments with error bars showing the 95% CI. Absolute values for HADS anxiety are shown in figures 51, 52 & 53

	Exercise	Normal Activity	Test	P-Value
1st to 2nd Assessment (SD)	-0.53 (2.4)	-0.47 (2.5)	M-W U* Test	0.77
2nd to 3rd Assessment (SD)	-1.71 (3.5)	-1.72 (3.1)	Independent Samples T**	0.99
1st to 3rd Assessment (SD)	-2.24 (3.2)	- 2.28 (3.8)	Independent Samples T	0.97

* Mann-Whitney U Test ** Independent Samples T-Test

Table 4: Change in HADS anxiety score

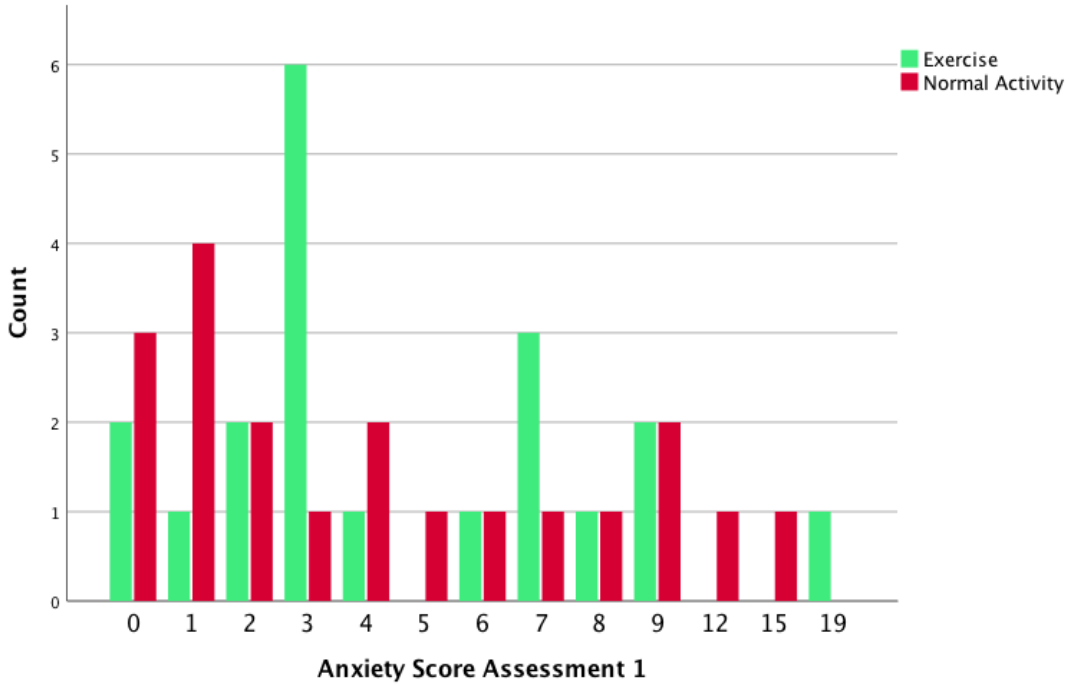


Figure 51: Clustered bar chart of anxiety scores at assessment 1

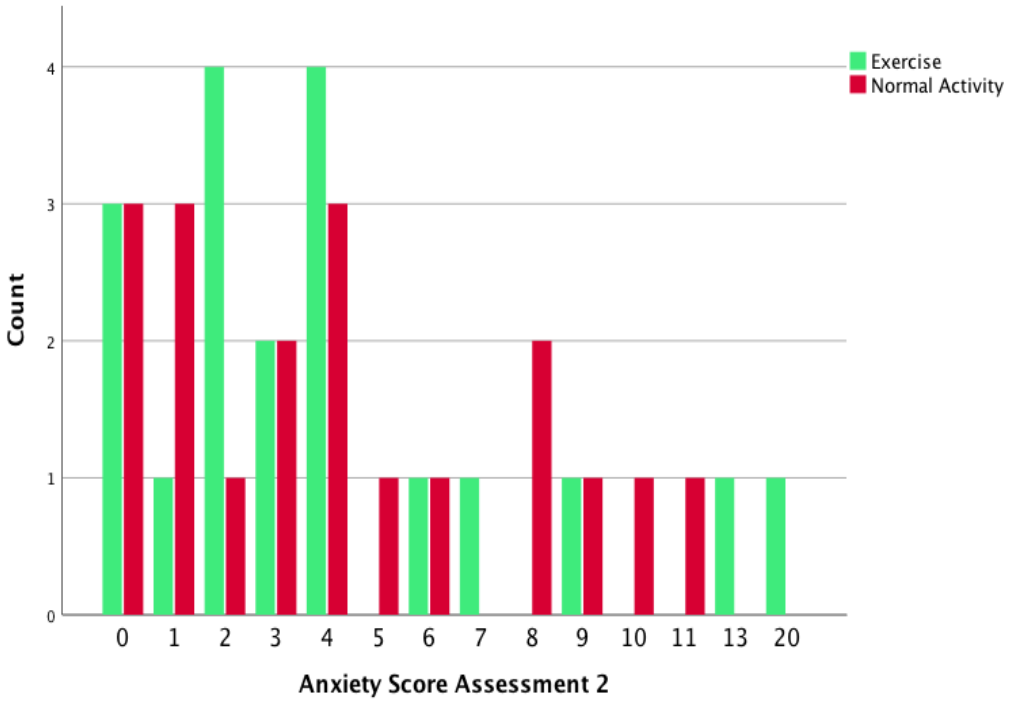


Figure 52: Clustered bar chart of anxiety scores at assessment 2

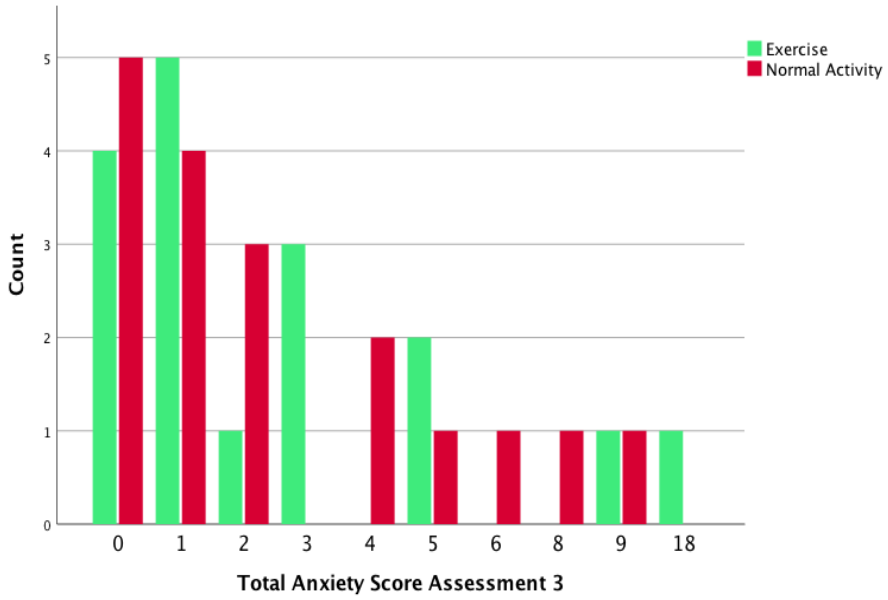


Figure 53: Clustered bar chart of anxiety scores at assessment 3

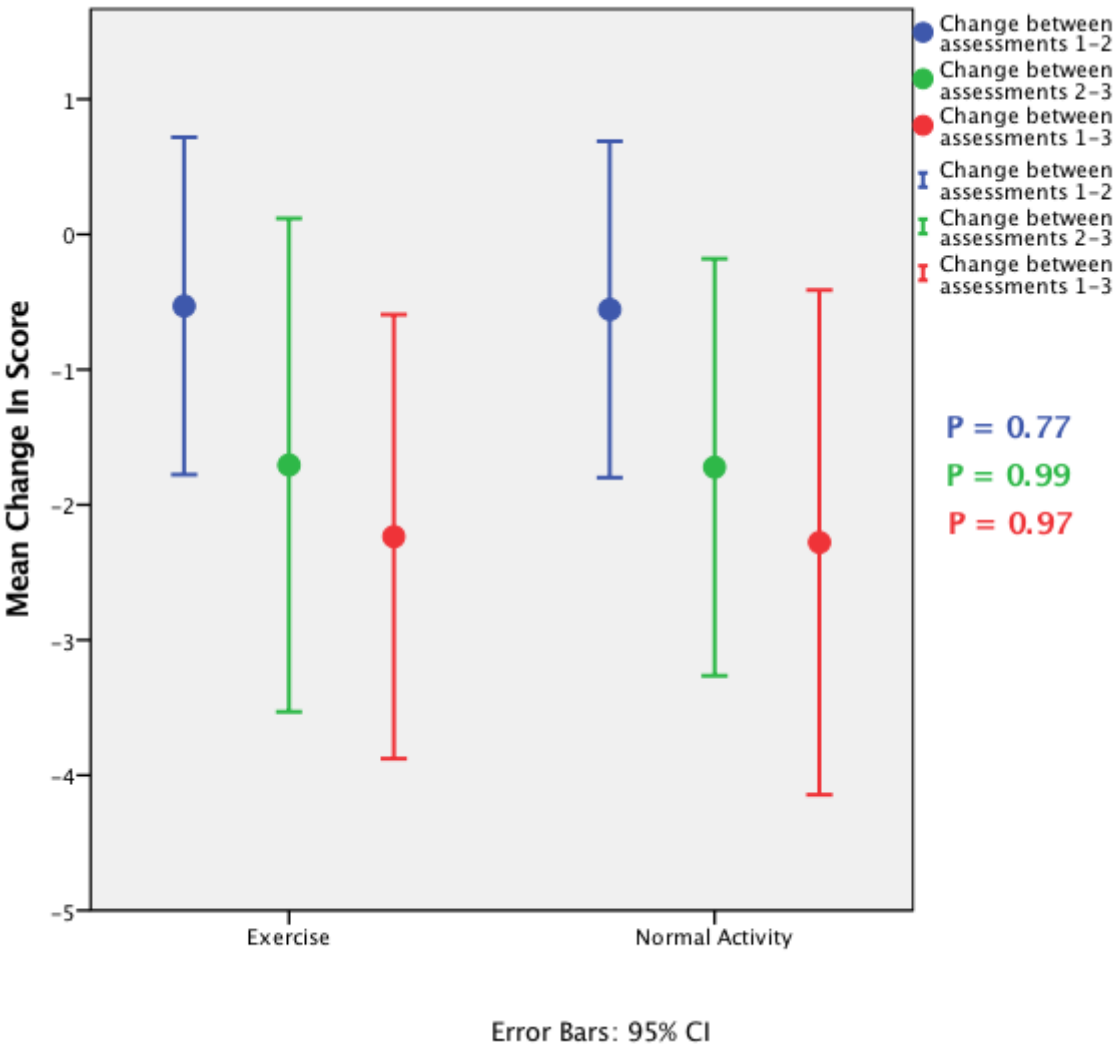


Figure 54: Mean change in HADs anxiety score, exercise vs. normal activity

Change in EQ-5D-5L Total Health Today Score

Perceived total health today score ranges from 0-100 with a higher score equating to a feeling of better health. Test for normal distribution of data were a Shapiro-Wilk Test where distribution was considered normal at >0.05. Normally distributed data were analysed with independent samples T-Test and non-normally distributed data with Mann-Whitney U Test. Table 5 Displays results and Figure 55 displays the mean change between assessments with error bars showing the 95% CI.

	Exercise	Normal Activity	Test	P-Value
1st to 2nd Assessment (SD)	-0.34 (10.3)	6.60 (19.9)	M-W U* Test	0.77
2nd to 3rd Assessment (SD)	-0.32 (20.0)	-2.61 (12.6)	M-W U* Test	0.46
1st to 3rd Assessment (SD)	-0.29 (19.9)	4.98 (24.4)	M-W U* Test	0.91

* Mann-Whitney U Test

Table 5: Change in EQ-5D-5L score

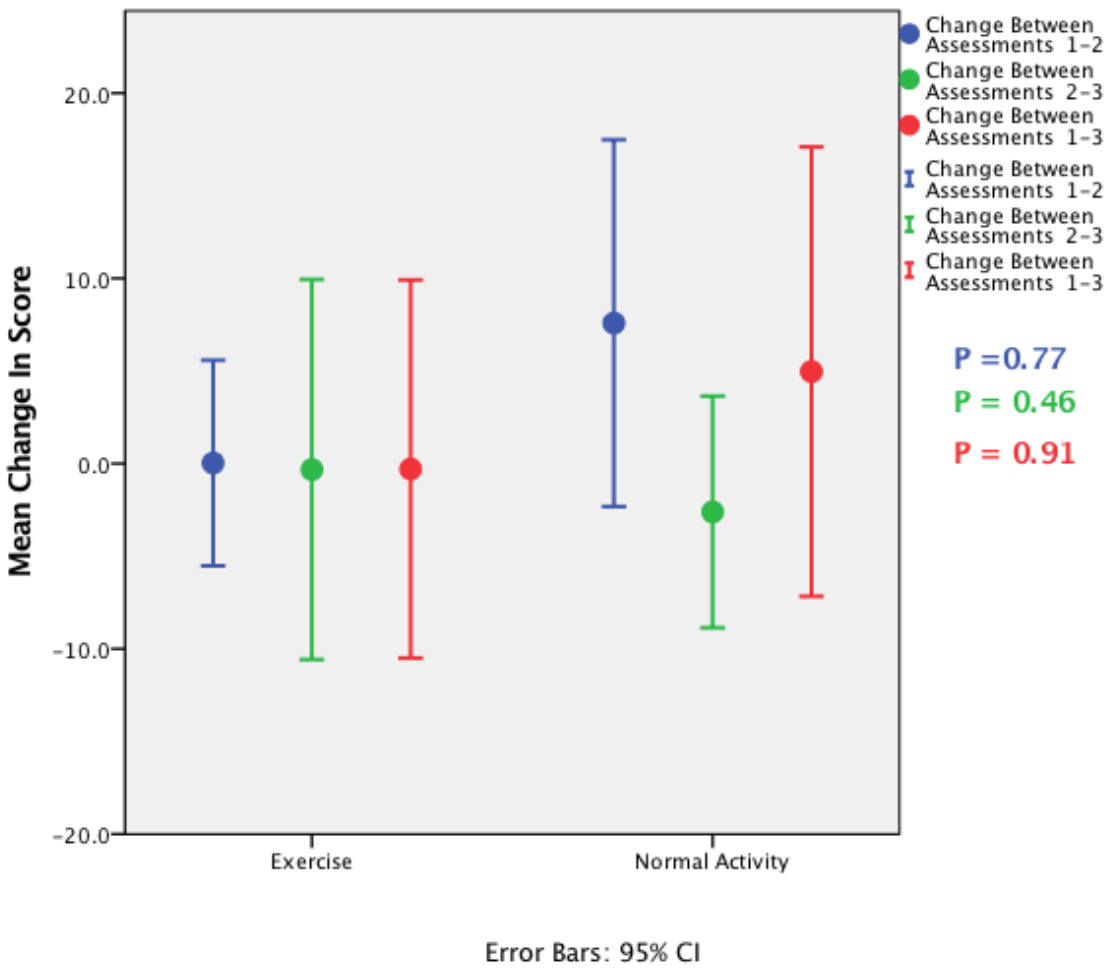


Figure 55: Mean change in EQ-5D-5L score, exercise vs. normal activity

Change in EORTC QLQ-C30 Combined Quality of life and Heath in last week

Perceived combined quality of life and health in the last week score health ranges from 0-100 with a higher score equating to a feeling of better quality of life and health over the last week. Test for normal distribution of data were a Shapiro-Wilk Test where distribution was considered normal at >0.05. Normally distributed data were analysed with independent samples T-Test and non-normally distributed data with Mann-Whitney U Test. Table 6 Displays results and Figure 56 displays the mean change between assessments with error bars showing the 95% CI.

	Exercise	Normal Activity	Test	P-Value
1st to 2nd Assessment (SD)	4.61 (14.8)	2.19 (16.9)	M-W U* Test	0.18
2nd to 3rd Assessment (SD)	-2.21 (15.7)	-0.93 (18.9)	M-W U* Test	0.99
1st to 3rd Assessment (SD)	2.94 (16.1)	1.39 (30.0)	M-W U* Test	0.33

* Mann-Whitney U Test

Table 6: Change in EORTC QLQ-C30 score

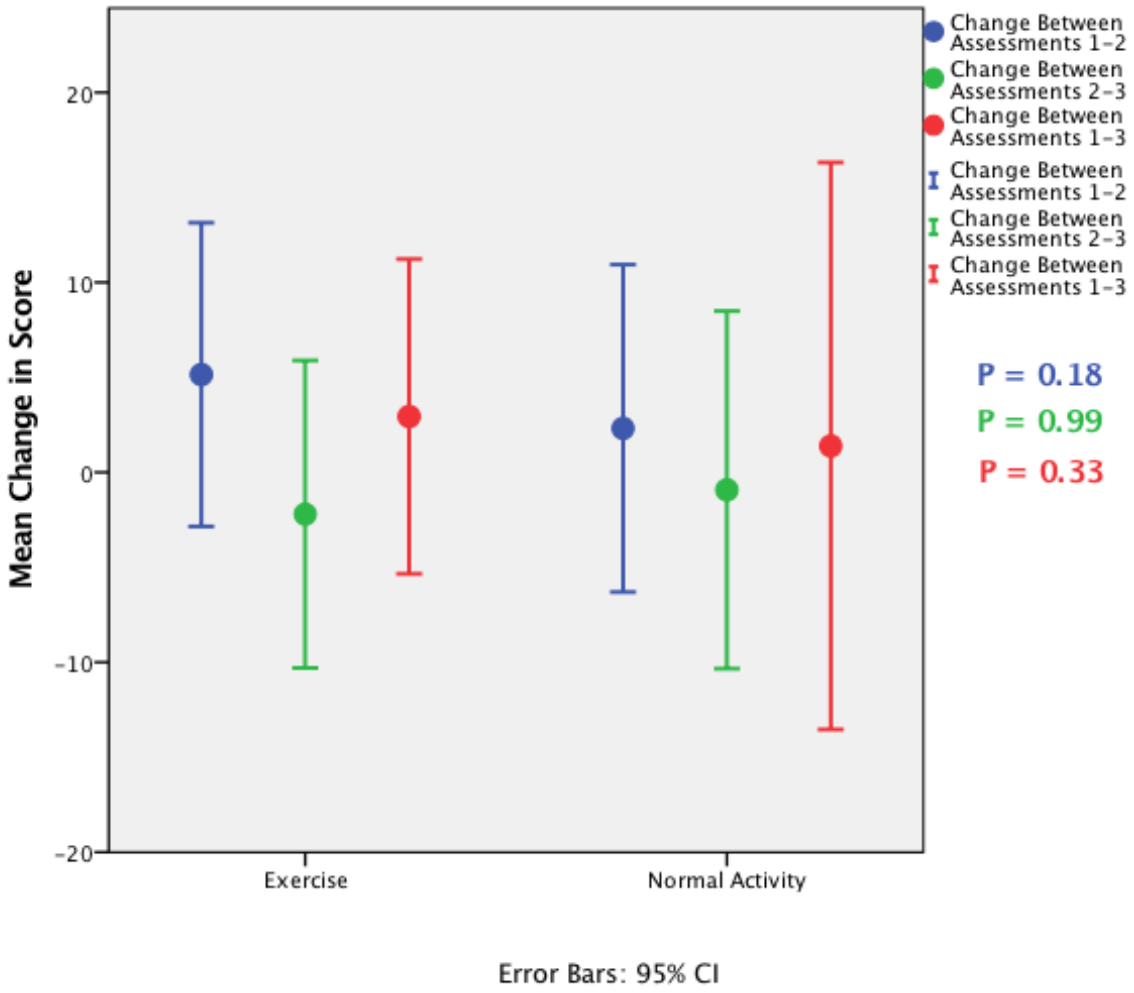


Figure 56: Mean change in EORTC QLQ-C30 score, exercise vs. normal activity

4. Discussion

4.1 Synopsis of Findings

Following analysis of the data, participation in this prehabilitation programme was found to be acceptable to those approached and achievable to those enrolled. The exercise arm of the study had a statistically significant increase in distance walked following the exercise programme compared with the normal activity group. There was a significant trend for lower BORG fatigue scores in the exercise group compared to the normal activity group, albeit very small. There was no difference detected in quality of life scores or heart rate recovery.

4.2 Discussion of Methods and Limitations

This study was a small pilot study and results have to be taken within the context of 40 participants. There is opinion that no matter what the potential of prehabilitation may or may not be, participation will not be adequate enough to realise any benefit. This question directed the primary outcomes and protocol design of this RCT. As such the exercise intervention was made as simple and inclusive as possible, the most ergonomic and simple activity monitor chosen and the least intrusive assessment schedule designed. By creating approachable prehabilitation we felt if the null hypothesis were proven then one could argue the viability of prehabilitation in general. We chose to approach those undergoing resectional bowel surgery for recruitment as many are subject to tight time targeted pathways that prehabilitation would have to exist within if used in the wider surgical care.

If the null hypothesis was disproven and this prehabilitation protocol appears to be achievable and acceptable is such a simple design of benefit? This question drove the secondary hypotheses. Markers of fitness and quality of life were chosen rather than major operative outcome measures such as length of stay as a trial powered to detect change in these is likely to be larger than was feasible in this setting.

Recruiting non-cancer participants for this study would have been simpler Cancer patients are subject to tight timelines that limit the space available for prehabilitation. However cancer patients undergo some of the largest of procedures carrying the most significant physiological insult. They are also less likely to be refused surgery due to lack of fitness than those with a benign diagnosis due to the life limiting nature of their disease. As such prehabilitation is likely to be most useful for this group. As the main aim of this study is look at acceptability and achievability, we felt it important to test this intervention in the group where it is most likely to be used in clinical practice.

The exercise intervention was tailored to each participant as an expansion of their performance in the 6MWT. However, the low BORG scores demonstrate that, for many, it may have been too low in intensity. This small study was not blinded meaning it is always possible to argue bias. However it is difficult to see a possible design that would allow blinding of participants to exercise.

4.3 Demographics

The majority of the participants had a cancer diagnosis and were therefore subject to a strict timeline for operative intervention. Participants were at differing points along this time line when listed

for surgery and as such had varying amounts of time available for exercise intervention. The clinical need for adjuvant chemotherapy also affected when it was possible for participants to undergo their three-month postoperative assessment. Time from first to second assessment and from second to third showed no statistical difference between arms. Not all participants were able to complete all assessments. This was for varied reasons such as inability to perform a 6MWT due to effects or complications of surgery, adjuvant chemotherapy, lack of motivation or due to travel to other centres out of region for further treatment. Again these potential confounding factors were spread equally between groups with an identical 80% of each arm completing all three assessments.

4.4 Primary Outcomes

Acceptability and Achievability

We set out to measure acceptability as previously published results have highlighted this as an issue. Here 45 potential participants were approached to recruit the 40 participants needed for the study. This gives an acceptability of this exercise programe of 89%. Acceptability of prehabilitation exercise programmes is of upmost importance if such interventions are to be of use in wider healthcare, outside of a trial environment. We decided to measure acceptability as recruitment rate alone rather than combine it with compliance. This was done as there was a feeling that many potential participants in a prehabilitation programme may not take part simply because the programme sounds too intensive. Separating like this allowed us to demonstrate how this form of prehabilitation is perceived (acceptability) and how it is tolerated (achievability).

It is worth remembering that in this study the MDT effectively pre-screened the patients removing those that were definitely not fit for surgery. It could be argued that these may of helped our recruitment as we did not have to approach the least fit. However as this study is investigating prehabilitation those who are deemed not fit for surgery are not of interest. Recruitment rates are not commonly published but those that are have shown mixed results. Those studies that have published similarly good recruitment rates have been of lower intensity and without the need for regular travel, similar to this RCT (Jones et al., 2004, Mock et al., 2005). In contrast those that have published poorer recruitment rates (Segal et al., 2001, Segal et al., 2003, Courneya et al., 2003a, Courneya et al., 2003b) have used more intensive exercise and supervision, often needing participants to travel to a gym. (Barakat et al., 2016) investigated the effect of a supervised exercise programe on patients awaiting abdominal aortic aneurysm repair. They approached 293 potential participants and documented a refusal rate of just over a third at 105 of the 293 approached. Their exercise programme consisted of 6 weeks gym based sessions and a significant reduction on post operative complications was detected in the 62 participants in the exercise group. This benefit however, was confined to the just over 50% (32/62) of the exercise arm that attended more than 75% of sessions. The 19 of the 62 in the exercise arm that attended some but less than 75% of sessions and the 11 that attended none showed no benefit in outcomes at all. This is one of the only studies to show an improvement in major post-operative outcomes such as length of stay and complications. This is interesting as it demonstrates the possiblity of obtaining real clinical benefit from prehabilitiation. However one third of those asked to participate declined and half of

those randomised to exercise did not attend enough sessions to demonstrate any outcome benefit. This raises questions about the feasibiltiy of such a 6 week programme that involves travel to a supervised programme, especially outside of a trial enviroment.

Given the age group, and known inactivety (Sport-England, 2017b, NBOCA, 2017) of the cohort that undergo bowel resection, an approachable home based programe of exercise may be key to generalisation into wider healthcare. Survey data (Ferreira et al., 2018) confirms that patients prefer low intesity, non-supervised programes. Carli et al., (Carli et al., 2010), also demostrated that larger improvents in fitness were noted with a walking based exercise programme versus higher intesity exercise. This adds weight to the argument that approachable prehabilitation programes, that do not involve travel or supervision, are better accepted. This falls in line with the acceptabilty rate of 89% that we found in our study.

The REx trial (Moug et al., 2019) confirmed that a walking based prehabilitation programme without direct supervision is feasable whilst undergoing neo-adjuvant chemotherapy before rectal cancer surgery. However despite this being a home-based programme recruitment rate was still only 62%, highlighting how difficult it is to motivate patients with a cancer diagnosis to exercise more. Those that did participate in this trial were however very satisfied with the experience and a non-significant increase in 6MWT distance was detected in the exercise group.

This study used activity monitoring technology in the form of a FBit Zip^{TM} to monitor the activity of the participants in the exercise group. Wareable divices such as the FitBIt Zip^{TM} are increasing in use and ability to give useful activity data to those who wear them. The use of FitBit ZIP^{TM} in a healthcare settting is only just being explored but their use as step counters has been validated (Cadmus-Bertram et al., 2015, Takacs et al., 2014, Diaz et al., 2015). Data from this study

confirms this with our cohort who wore the device and generated data for 93% of the days available. This study confirms that patients are able and willing to use activity-monitoring technology. Previous trials have used triaxial accelerometers to count steps (Moug et al., 2019) that needed to be glued to a participants leg and have no screen to give feedback to the user. There is some evidence that simply wearing an activity monitor that gives feedback (such as the device used in this study) may increase activity without any further intervention (Ummels et al.), although this effect may be short (Phan et al., 2018). The question for research is then; does wearing an interactive fitness device like used here count as an intervention in itself? This is still an unanswered question and is why the control group in this study were not given a FitBit Zip^{TM}. If we had given the control group a device as well we would have avoided the potential Hawthorne effect that may be associated with it. This would have needed the addition of a third arm to this trial as a true representation of current care (no device and no exercise intervention) is needed. Within the confines of a small pilot study this would not have been practical and likely need the addition of extra sites. However for the purposes of future intervention outside of a study this issue is moot, as any added benefit from the device alone above an exercise intervention can only serve to improve outcome.

The median compliance with the exercise programe was 68%. As expected there was a trend for fitter participants (classified as walking further in the initial 6MWT) to be more compliant with the exercise programme (Figure 14). This is what one would expect, as if someone is already motivated to keep fit then they are likely to be more motivated to undertake a further exercise programme. It could be argued, however, that this “already fit” group are least likely to benefit from an exercise intervention, as they may already have the physiological buffer necessary to cope with an upcoming surgical insult. Despite the observation of higher compliance with participants who were able to walk further at baseline, there was a trend

demonstrated for those who were less fit to benefit more from the exercise programme. A medium strength negative correlation between distance walked in the initial assessment and percentage increase in walking distance following the exercise programme can be seen in the results (Figure 15). This demonstrates the less fit you are the more you stand to gain. The p value for this spearman correlation was 0.131 but this is likely to be a function of the small sample size as is commonly the case with correlation calculations.

4.6 Secondary Outcomes

Distance Walked in 6MWT

Data were analysed to look for difference in distance walked between the exercise and normal activity arms for all assessments. The only statistically significant change in distance walked between the groups was from the 1st and 2nd assessments. Here the exercise group managed to increase the distance walked, where the normal activity group did not manage to match the distance walked in their first assessment. The mean difference between the groups was 29.4m. This is interesting as the intervention not only seems to have produced extra fitness but also seems to have prevented the deconditioning seen in the normal activity group. The mechanism for this deconditioning in the control group is likely to be multifactorial. Without specific guidance telling patients to exercise before a major surgical intervention it is likely that they do the opposite. This may be the result of stress induced by a cancer diagnosis or an upcoming major surgical procedure. Whatever the reason the prevention of deconditioning may be even more important than the attainment of extra fitness with prehabilitation. Those who undergo pre-operative assessment, and are just over the threshold for fitness to undergo surgery, may decondition to a point where they are no longer fit

between that assessment and surgery. The prevention of this is likely to do more to prevent adverse outcome than increasing the fitness of an individual who is already robust enough to withstand surgery. There was a similar trend between the 1st and 3rd assessments. Both groups lost fitness but the exercise group lost less. The exercise group lost less fitness than the normal activity group walking 10.8m less in the 6MWT, where the normal activity group walked 38.8m less on average. This 28m difference is a similar to the statistically significant difference seen between the 1st and 2nd assessments. It did not reach statistical significance here due to a much higher variance in distance within groups giving a larger standard deviation of the mean and wider confidence intervals. This could be explained by the addition of the confounding factor of adjuvant chemotherapy for some participants. At three months postop those participants who received chemotherapy were mid treatment. Chemotherapy has large and well documented deleterious effects on fitness (van Waart et al., 2015) and this was certainly observed in this study, where some participants cited this as a reason for not wishing to complete their final assessment.

Median time from 1st to second assessment was almost identical between the 2 groups and 11.5 and 12 days with no statistical significant difference (p=0.559). As displayed in figures 3 and 4 there was a larger range of time between the 1st and 2nd assessments in the exercise group. This group contained the 2 participants who spent longest in this study period at 77 and 100 days. Despite there being no significant difference between groups it could be argued that these two participants made the exercise intervention seem more effective. However it seems that there was deconditioning in the normal activity group so one could assume that if these participants had been randomised to the normal activity arm they would have made the deconditioning appear worse. In this case the fitness gap between the groups may have stayed the same.

Both the the REx trial and Barakats study had significantly longer exercise programmes (14 and 6 weeks) than the mean time in study of 21 days for the exercise arm of this trial. The short time period we encountered was a product of the tight target driven pathway that cancer patients are on. Given a longer period of study benefits from this study may have been more significant.

Heart Rate Recovery

A lower rise of heart-rate in response to an identical exercise and quicker recovery following it, demonstrates increased cardiovascular fitness. No such change was detected in this study between exercise and normal activity groups. The exercise intervention was purposely designed to be of lower intensity to maximise compliance. As such heart rate increases were not large following exercise perhaps negating ability to detect differences between groups. The extra fitness seen in distance walked by the exercise group may not have been enough to alter changes in heart rate. Of the studies discussed in the introduction only two had HR as a marker of outcome for prehabilitation, (Kim et al., 2009, Dunne et al., 2016). Kim et al., (2009) found a lower heart rate when participants were exercising to VO_2 Max in the exercise arm of the study after an exercise intervention of home-based exercise bike. Dunne et al., (2016) found no difference in HR between groups despite finding other improvements on CPET in the exercise group. Both the studies of Kim et al., (2009) and Dunne et al., (2016) used CPET to measure fitness rather than less intensive exercise such as the 6MWT used in this study. At higher intensity exercise it should be easier to detect differences in HR due to physical fitness. However despite improved CPET parameters Dunne et al., (2016) showed no difference in HR in the exercise group. The lack of HR as an outcome in all other studies on prehabilitation in abdominal surgery and the very mixed results in the two that used it, even on maximal testing, leads to a

conclusion that it is not a very reliable marker of fitness outcome in this setting. One could argue therefore the rationale for including it as an outcome at all in in this study. Heart rate is probably the most common metric measured in fitness training in any setting. Its ubiquity as a measure of exertion made it seem odd to complete an exercise based study without HR measurement. Many of the patients included in this trail had significant comorbidity and heart rate measurement in response to an exercise intervention is likely wise as a safety precaution.

BORG Scale

Despite a median change of BORG score of 0 for both groups between assessments 1 and 2 there was a statistically significant difference between groups. A trend for a decreased level of self determined exertion in the exercise group and increased in the normal activity group was found. As observed in the distance walked data, there was a similar trend between the 1st and 3rd assessments that did not reach statistical significance.

Quality Of Life Assessments

There was no significant difference between groups on quality of life assessment. Change in scores varied widely generating large 95% confidence intervals for the mean scores. Participants were largely dealing with a cancer diagnoses, the prospect of major surgery and not insignificant mortality rates. Following such surgery 2 participants were still inpatients and many undergoing adjuvant chemotherapy by the time the 3rd assessment was due. Given this it is not surprising then that any change in quality of life score that an exercise intervention may create was not detectible, if present. The results from the literature reviewed in the introduction show only 5 of the 21 comparable published studies on prehabilitation used quality of life assessments. The results from these 5 show broadly similar results

to this study. Of the 5 that included quality of life assessments as outcome measures 4 showed no difference between groups (Dunne et al., 2016, Barberan-Garcia et al., 2018, Dronkers et al., 2008, Dronkers et al., 2010, Gillis et al., 2014). The only study to show an improvement in quality of life scores was Dunne et al., (2016) where a significant difference in SF36^{TM} was detected. This study used a supervised high intensity interval cycling programme over a 4-week period before liver resection. Despite improved CPET and SF36^{TM} scores no improvement in operative outcomes including complications, length of critical care stay or total length of stay were found. The recruitment rate for this exercise programme was not good with 104 eligible patients needing to be approached to recruit 38 participants. Therefore despite results that seem promising with an increase in quality of life scores this exercise regime seems to be relatively unacceptable to those approached. Our results with regard to quality of life outcomes seem to sit well with that already observed in the published literature.

4.7 Overall Discussion of Findings

This study adds to the published work by demonstrating that patients are willing to participate in a walking based exercise programme and use activity monitoring technology to achieve set goals. It also demonstrates that a low intensity programme can increase fitness despite a short time period. The obvious question that has arisen from this study is whether the 29.2m statistically significant difference in distance walked reflects any clinical significance? This pilot study did not include clinical outcome measures such as length of stay and is not powered sufficiently to detect them. A similar question has arisen when single elements of multi-modal optimisation, or as it has commonly become known “enhanced recovery”, packages are tested alone. The benefits originally seen by Kehlet’s group and conformed

by others (Gatt et al., 2005, Anderson et al., 2003, McNaught and MacFie, 2002) are accepted to be the sum of many parts. It is not often possible to demonstrate that individual parts of these packages of care make differences to headline clinical outcomes. The failure of the literature so far to conclusively prove that prehabilitation can deliver improved outcome may mean that it should be seen as the next addition multi-modal optimisation of care. Such packages have been proven to deliver outcome benefit, but up to this point concentrated on postoperative care and treatment of the patient on the preoperative day only. Prehabilitation packages such as this could be a useful addition to usual care around major surgery. Those patients who are already fit are likely to benefit the least and interventions may be better targeted to those who are on the borderline of fitness for surgery.

In this study there was a trend for these less–fit patients to benefit more from the exercise programme. As previously demonstrated in the published literature (Carli et al., 2010, Ferreira et al., 2018), lower intensity exercise programmes without the need for direct supervision, are better accepted by participants. This study sits alongside these findings with a high acceptability and reasonable achievability rates.

This study demonstrated that the use of wearable activity monitoring technology is useable by both patients and as a method of monitoring by healthcare workers. Timing of prehabilitation and how it fits alongside existing time-targeted surgery is likely to be key to its success. The exercise group of this study had on average 21 days available to them between recruitment and surgery with many having considerably less than that. Barakat et al. who showed an improvement in operative outcome following prehabilitation had a 6 week period for the participants to undergo training. Their participants were awaiting elective aortic aneurysm repair that is not as tightly time-targeted as cancer resection and as such gives more

space for prehabilitation. If the evidence were stronger for the benefits of prehabilitation one could argue that in those of borderline fitness, surgery should be delayed to undergo prehabilitation. However the evidence from this study and that so far published is not strong enough to argue a case for they delay of cancer treatment.

There is no study that has found or reported any harm coming from a prehabilitation programme. Therefore moving the timing of any exercise intervention to before a decision has been made for surgery should be acceptable. By doing this many would be included that don't actually undergo surgery and could therefore gain no benefit if terms of operative outcome. However if there is no appreciable risk of harm to these patients and real benefits can be demonstrated, the approach of prehabilitation in those who have symptoms that may lead to surgery may produce greater benefit than waiting until the decision is made, and time is limited.

Exactly this premise was investigated by Barlow in a study (Barlow et al., 2018) that recruited 189 patients that were being referred on to secondary care for a possible diagnosis of cancer. Of these 189 recruited 163 attended secondary care for investigation and of those 15 had a cancer diagnosis following investigation. This study did not however include exercise in its prehabilitation bundle but rather optimization of behaviour such as smoking cessation and moderation of alcohol intake, tighter control of hypertension, blood glucose and co-morbid conditions as well as nutritional supplementation where necessary. The "care bundle" proposed for prehabilitation was well received with only 6 refusing to take part, although this cannot be compared with recruitment to an exercise intervention that requires far more from a participant. Adding an exercise intervention to a pre-diagnosis prehabilitation programme such as that used by Barlow et al., (2018) would need a low cost and non-supervised programme given the much larger numbers that would be included. As such a protocol like the one used in this study, monitored by wearable

technology may be ideal and may help avoid the time constrains encountered with prehabilitation commenced after a decision to operate has been made.

4.8 Further Work

This pilot study validates the method and demonstrates that it is possible to improve patient care with wider use but does not prove that. Larger studies will always provide more powerful evidence; however simply enlarging this study would not be a logical next step. As previously mentioned it may not be possible to demonstrate an effect on key outcomes such as mortality. Advanced exercise testing such as CPET gives detailed analysis of a patient's fitness for surgery. A study using a similar programme to this in a cohort who could spend more time in an exercise programme, with fitness assessed by CPET, may yield interesting results. This may also get round the variance in follow up time that was seen in this study. As mentioned previously there was a wide range in the timing of the 3rd assessment. This was in the majority due to the need for further treatment in the form of adjuvant chemotherapy or further surgery at distant sites.

Recruiting participants from primary care referral for surgery or from those with a benign diagnosis may give more time to investigate the limits of benefit that can be gained from prehabilitation. The time restraints governed by cancer waiting list targets could be altered if prehabilitation was considered part of the treatment pathway. The PREPARE ABC (Hernon, 2018) is currently recruiting to a 4-week exercise prehabilitation trial prior to colorectal cancer resection. This 4-week period has been secured by permission by regulators for the exercise intervention to be considered the first stage in treatment for

colorectal cancer and as such the delay in surgery will not breach targets. If results from this are favourable it may open a larger window for prehabilitation, particularly for those who are on the borderline of fitness for surgery and risk of delay in surgery is outweighed by gains in fitness.

In addition to looking at elective surgery, there may be large gains to be made in the application of some prehabilitation to the emergency surgery group. This has obvious difficulties as the emergency group are, particularly the general surgical cohort, often unstable and inappropriate for all but resuscitation prior to surgery. The first step in assessing this group is likely to be highlighting those who would benefit the most. Fitness testing in those who have acute intrabdominal pathology can be almost impossible. In place of, and in addition to, fitness testing there has been much interest in frailty assessment. Frailty has been associated with poor outcomes in both medical and surgically managed pathology. Specific frailty assessment and interventions (discussed in the next chapter) have been shown to improve outcome markedly in medical and surgical cohorts. The identification of frail patients within the emergency surgical cohort at least may enable post operative intervention but, dependant on the urgency of the operative treatment, may allow some prehabilitation in this group.

4.9 Conclusion

In conclusion we have demonstrated some interesting answers to the questions this study set out to answer. It appears both acceptable and achievable for patients listed for resectional bowel surgery to undertake a walking based exercise programme. Even within the confines of a strict cancer target pathway fitness measures can be improved. The use of simple wearable activity-monitoring technology is acceptable to patients.

Below, I have listed the initial hypotheses and whether or not the null hypothesis has been rejected from using the results of this study:

Primary Hypothesis
It is acceptable and achievable for patients to undertake a simple walking-based prehabilitation programe, monitored by wearable technology, in the waiting time before colorectal surgery.
- Hypothesis is accepted, null hypothesis rejected

Secondary Hypothesis
It is possible to increase fitness levels and quality of life scores of a simple walking-based prehabilitation programe, monitored by wearable technology, in the waiting time before colorectal surgery.
- **Results obtained show evidence to reject the null hypothesis, but the results are not conclusive. Fitness levels did show improvement on distance walked but quality of life scores were unaffected by the exercise programme.**

Effect of Frailty measured by Sarcopenia In Emergency Surgery

1 Introduction

1.1 Frailty

Frailty is a widely used term that does not have an agreed clinical definition (Rodriguez-Manas et al., 2013) and, as such, does not appear in the WHO international statistical classification of diseases. Despite this “frailty”, however measured, has been used to predict adverse events and outcome (Fried et al., 2001, Rockwood et al., 2004, Bandeen-Roche et al., 2006), particularly in elderly patients.

Examples of definitions of frailty include “a biologic syndrome of decreased reserve and resistance to stressors, resulting from cumulative declines across multiple physiologic systems, and causing vulnerability to adverse outcomes”(Fried et al., 2001), or “ a state of increased vulnerability to poor resolution of homoeostasis after a stressor event, which increases the risk of adverse outcomes, including falls, delirium, and disability”(Clegg et al., 2013). Neither is precise or reproducible. Both are wordy and non-specific emphasizing the difficulties with using this term in clinical research.

Frailty has not been extensively investigated in surgical patients despite the fact that intuitively one would expect the “frail” to adapt less well to a surgical insult than the physically fit.

The gold standard in frailty measurement is comprehensive geriatric assessment (CGA). This involves a multidisciplinary team (MDT) approach including a geriatrician, physiotherapist, occupational therapist, dietician, pharmacist, and speech and language therapists.

CGA has been demonstrated in some studies to have major outcome benefits when applied to the frail hospitalised population (Ellis et al., 2017). However, CGA is cumbersome and unsuitable for routine preoperative assessment. Clearly, as described, it is inappropriate in the emergency surgical setting. In fact any kind of assessment that includes a functional element is difficult in the acutely unwell surgical patient. Those presenting with acute intrabdominal pathology commonly have pain as a predominant feature. This combined with the physiological instability that is associated with patients needing an emergency laparotomy means the results from any functional assessment are unlikely to give much insight into the patients baseline functioning. As such frailty assessments that including these are of little use. Time is also a factor, as decisions frequently need to be made out of normal working hours or without the time required for a multidisciplinary assessment such as CGA.

Other frailty scores have been described. These include, for example, those describe by (Fried et al., 2001) or (Mitnitski, 2001) that have been validated in retrospective reviews of thousands of patients. However even these modified scores can be cumbersome and are often more suited to geriatric medical specialties than surgical ones. Shortened frailty scales such as the clinical frailty scale by (Rockwood et al., 2005) have been used in frailty scoring in the surgical population and most recently in those undergoing emergency laparotomy in the ELF study (Parmar et al., 2019). The clinical frailty scale is now included in the National Emergency Laparotomy Audit (NELA). Shortened frailty assessment scores such as these are well validated but include an element of subjectivity as the user must decide which frailty category the person falls into.

1.2 Sarcopenia as a Marker of Frailty

Sarcopenia as a term was proposed by Irwin Rosenberg in 1989 (Greek ‘sarx’ meaning flesh plus ‘penia’ meaning loss) to describe this age-related decrease of muscle mass (Rosenberg, 1989). It is now defined as a combination of low muscle mass and weakness in older adults that causes functional problems. From the 1st of October 2016 it has been recognised as an independent reportable condition in the Clinical Modification (ICD-10-CM) provided by the Centers for Medicare and Medicaid Services and the National Center for Health Statistics for medical coding and reporting in the United States. The ICD-10-CM is a morbidity classification for classifying diagnoses and reason for visits in all American health care settings.

Global frailty assessment takes account of more than just the physical aspects of sarcopenia including social and physiological factors. Despite this sarcopenia is often found in combination with frailty and measures such as the ‘time-to-up-and-go test’ and gait speed have been used to measure both (Cruz-Jentoft et al., 2010, Sternberg et al., 2011).

Sarcopenia is the major contributor to physical frailty in the geriatric population. It is associated with morbidity and mortality from physical disability, falls, fractures, poor quality of life, depression and hospitalization (Beaudart et al., 2016, Janssen et al., 2002). However the loss of skeletal muscle mass affects more than just physical strength and functioning. Skeletal muscle is the major store of protein and in times of stress provides a source of amino acids for protein synthesis in other systems, making it integral in mechanisms of systemic inflammatory response (Malietzis et al., 2016). Skeletal muscle is the largest disposal site of glucose and energy and its loss is associated with age related fall in basal metabolic rate (Cooper et al., 2012). High levels of body fat coupled with sarcopenia have been shown to cause significant mobility difficulties in the elderly and a

term of sarcopenic obesity has been coined (Zamboni et al., 2008). Sarcopenia, although distinct from frailty, is felt to be a mediator and integral to the development of frailty, with both commonly occurring in combination (Cooper et al., 2012, Rolland et al., 2008). Development of sarcopenia especially in combination with increased adipose tissue undoubtedly impairs physical function and an individual ability to live independently. These are all hallmarks of vulnerability and frailty both as a cumulate deficit mode (Mitnitski, 2001), or a phenotype model (Fried et al., 2001). It is clear that both sarcopenia and frailty are interlinked and frail patients can be identified by a loss of skeletal muscle with or without obesity (Cesari et al., 2006).

The measurement of total body skeletal mass is possible in different ways. Whole body Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) are both accepted validated methods to measure this (Heymsfield et al., 1997). Whole body CT or MRI is not a tool that is used clinically and as such work has been carried out to find a “single slice” that would accurately represent total body composition. Work was published (Heymsfield et al., 1997, Shen et al., 2004) where different slices of whole body MRI in 328 healthy volunteers (including both sexes and varying BMI and ethnicity) were assessed to find the best area to represent total body composition. They looked at total body fat and muscle. It was found that the slice that best represents total body lean muscle volume was 5cm above L4-5 with a correlation to whole body skeletal muscle make up of 92.4%. For the group of patients whose disease brings them to a gastrointestinal surgeon they are far more likely to have a CT than an MRI. It is for this reason studies discussed below now use a single CT slice, at L3 level, to represent muscle mass and determine sarcopenia.

1.3 Review of Published Work on Effect of Sarcopenia in Surgery

Frailty is becoming an area of interest to the surgical community as it is felt it can be used to predict poor outcome. There are many different preoperative risk stratification scores that in the majority concentrate on laboratory biochemical markers, co-morbid conditions or formal cardiopulmonary exercise testing. Such assessments are of great use but fail to capture the frailty of a patient. Frailty describes the vulnerability of a patient and can be used to predict mortality, morbidity and dependence. Frailty scores are not suited to the surgical pre-operative pathway especially in the emergency setting. Sarcopenia has been used as a surrogate marker of frailty due to its common presence in the frail and its integral role in the development of frailty.

Below is a systematic review of the current published literature on CT assessment of sarcopenia as a marker of frailty in the surgical population. Given the heterogeneous nature of these studies Meta analysis was felt to be not possible.

All novel prospective or retrospective studies looking into the effect of CT assessed sarcopenia in a cohort of patients that had undergone surgical or radiological intervention were deemed eligible for this review. A search of the PubMed database on the 24th of April 2018 was performed with the search terms Sarcopenia and Surgery. Filters of English language and publication in the last 10 years were applied yielding 853 results. These were then reviewed as shown in the Prisma flow chart (Figure 57).

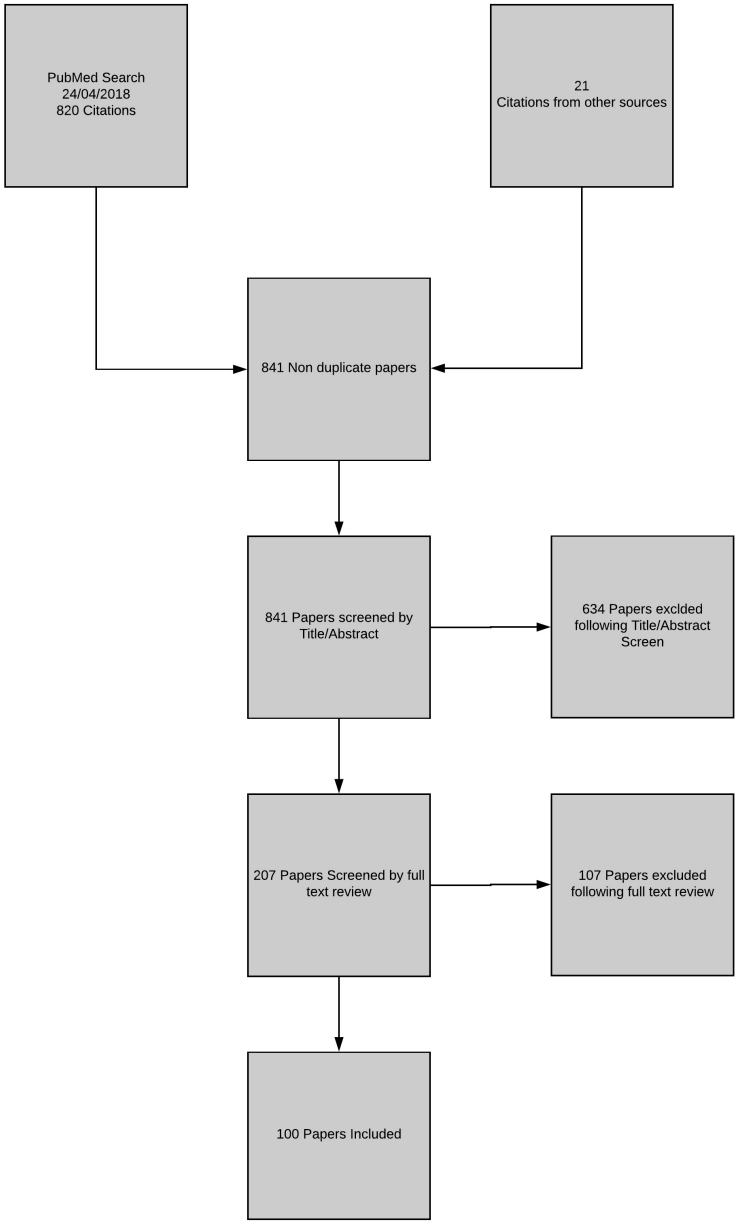


Figure 57: PRISMA Flow Chart

Following assessment (Figure51) 100 studies were included in this review that in total involved 29,404 patients. All but four contained exclusively elective patients. The majority (75/100) were of patients being treated for malignancy. 97 of the 100 were retrospective.

The majority (95/100) of the studies used “single slice” calculation of sarcopenia at L3 or L4 level. Of the five studies that did not use this

level, two used analysis of a “single slice” at T12 level, two used psoas volume calculation and one used a thoracic level to give pectoralis and erector spinae muscle cross-sectional area. The method described above by (Shen et al., 2004) for single slice estimation of total body skeletal muscle and determination of sarcopenia takes into account all skeletal muscle at this level. Although based on this, many published studies further simplify the method to include the psoas muscle only. The frequency of the different methods of CT sarcopenia assessment across the 100 studies is shown in table 7. Table 8 shows the effect that sarcopenia was found to have in these studies.

Cut-off values for sarcopenia varied widely between the different studies. This has an effect of varying the amount of a study population that is found to be sarcopenic from 25%, where the lowest sex specific quartile was used, to 73% (Ninomiya et al., 2017, Du et al., 2014) where values were used from other studies (Suzuki et al., 2016).

Many surgical specialties and different procedures were included in the published literature demonstrating the pan-speciality interest in this subject. Effect of sarcopenia on different procedures included: colorectal; pancreatic; liver resection and transplant; gastric; oesophageal; pulmonary; open and interventional radiological cardiac; open and interventional radiological vascular; renal, bladder; prostate; femoral and pelvic trauma; emergency abdominal; laryngeal.

Method of Sarcopenia Assessment on CT	Number of Studies Using this Method
Psoas Area	38
Total Muscle Area	34
Total Muscle Area and Adipose Tissue	8
Psoas Density	7
Psoas Intramuscular Adipose Tissue	3
Psoas Area and Density	3
Psoas Volume	2
Dorsal Muscle Area	2
Psoas Area and Volume	1
Ratio of Psoas Long to Short Axis	1
Pectoralis and Erector Spinus Area	1

Table 7: Methods of CT sarcopenia assessment

Effect of Sarcopenia	Numbers of Studies Reporting Effect
No Effect	7
Decreased Survival	50
Increased Complication	37
Increased Length of Stay	13
Increased Cost	5
Higher Recurrence Rate	5
Increased Chemotherapy toxicity	1

Table 8: Effects of CT defined sarcopenia

Measurement of sarcopenia has been done by differing methods at L3/4 level (Total skeletal muscle vs. psoas alone and area vs. density vs. volume). Total skeletal muscle is often used as this has been described and validated. Psoas area and density are easier to calculate as they do not require additional software over and above what is used to report CT images. Much of the published literature gives reference to sarcopenia and obesity and the reduction in quality of muscle as it is replaced with fat. A number of studies attempt to address this (Pecorelli et al., 2016a, Clegg et al., 2013, Okumura et al., 2016, Kobayashi et al., 2016, Nishigori et al., 2016, Hamaguchi et al., 2016, Tamandl et al., 2016, Lieffers et al., 2012) by

including visceral fat or intra-muscular fat in sarcopenic assessment. Psoas density, and the studies that use it (Wagner et al., 2016, Joglekar et al., 2015, Buettner et al., 2016), can be included in this as they are assessing muscle quality not just size or volume.

1.4 Sarcopenia measurement in Emergency Abdominal Surgery

Such assessment of sarcopenia has been shown to be a useful independent marker of outcome in the elective setting. Over the last decade there has been increasing interest into the care of elderly patients need in emergency gastrointestinal surgery (Wilkinson, 2010). The service provided to patients undergoing emergency surgery (laparotomy) is now being audited and published freely at a national level by the National Emergency Laparotomy Audit (NELA). As NELA moves past its first few years, its data is becoming more complete and as such, variances in outcome are being looked at with greater interest. No real work has been published looking at frailty in this group. In the elective setting modern surgical assessment and MDT will often risk assess the elderly and frail and deem them inappropriate for surgical management. In the emergency setting assessments such as cardiopulmonary assessment are not possible. Decisions on operative fitness in life in the emergency situation are made on the judgment of consultant surgeons and anesthetists, taking into account of patient and family wishes. This is not bad practice but in an era where results are coming under increased and appropriate scrutiny, risk adverse behavior can develop. Some objective assessment of survival is made, such as P Possum (Copeland et al., 1991) scoring, but these fail to assess the frailty of patients. P Possum was not designed for individual risk assessment and although it can add some objective assessment, it relies heavily on individual's vital parameters such as heart rate and blood pressure that can vary widely in the pre-operative period.

The European Working Group on Sarcopenia in Older People published a consensus statement on the measurement of sarcopenia in 2010 (Cruz-Jentoft et al., 2010). Here they stated that to investigate sarcopenia an assessment of muscle mass and muscle function should be made. It is stated that this is optimal as the relation between muscle mass and function is not linear. This may be optimal in an elective or research setting but assessment of muscle function in an emergency setting is not possible. Assessments such as gait speed or handgrip strength are not at all reliable in any acutely unwell patient especially not in those who need an emergent laparotomy.

CT assessment of sarcopenia has clear advantages as it is reproducible and takes advantage of an investigation that is almost universally performed in patients undergoing major gastrointestinal surgery. It appears to be a useful predictor of poor outcome for both morbidity and mortality. This assessment of frailty may be of greatest use in the emergency surgical patient where such luxuries as careful pre-operative physiological assessment such as cardio pulmonary exercise testing are not available. Any assessment of frailty by measurement of sarcopenia would have to be easy to use and give rapid results in the emergency situation. Detailed specialist CT analysis using additional software above the standard reporting set used are not likely to be appropriate in the emergency situation.

1.5 Aims of the Research Study and Hypotheses

Aim: To assess the effect of sarcopenia, measured by psoas density and area on perioperative CT, on outcomes following emergency laparotomy.

Primary Hypothesis

Frailty defined by sarcopenia on CT scanning, independently worsens mortality following Emergency Laparotomy

Secondary Hypothesis

Frailty defined by sarcopenia on CT scanning, independently worsens social dependence following Emergency Laparotomy

Rationale

It has been demonstrated that increasing Frailty has an adverse effect on outcomes following surgery. There is no time for complex frailty scoring in the emergency setting. CT measured sarcopenia has been shown to be a surrogate marker of frailty. No functional assessment of muscle function is possible in this cohort due to their acute pathology. The simplest of all sarcopenia measurements are single slice psoas area and density measurements making them most appropriate for usage in this group in whom the majority have perioperative CT scanning. Should an association be found between sarcopenia and mortality or social dependence post emergency laparotomy it could be used to aid peri-operative decision-making.

Approach

Patients who have undergone an emergency laparotomy will be identified from the NELA database from multiple sites. Retrospective analysis of those who had a perioperative CT scan will be undertaken and those in lower quartile of psoas area and density will be deemed sarcopenic. The outcomes of 30-day and one-year

mortality rates as well as postoperative social dependence will be compared between sarcopenic and non-sarcopenic groups

2. Methods

2.1 Design

This was a retrospective multisite study including 4 NHS acute hospitals:

- Scarborough General Infirmary
- York District Hospital
- Bradford General Hospital
- Royal Derby Hospital

Patients who had undergone an abdominal CT within 30 days of an emergency non-resuscitative laparotomy under the general surgical team in a 12-month period from October 2014 were included. Included then had their CT scans analysed and clinical data and outcomes recorded.

2.2 Patient Identification

Patients included in the study were identified from the National emergency Laparotomy Audit (NELA) database. NELAs aim is "to enable the improvement of the quality of care for patients undergoing emergency laparotomy, through the provision of high quality comparative data from all providers of emergency laparotomy. Each provider was asked to input all emergency laparotomies into the database prospectively. NELA is part of the National Clinical Audit and Patient Outcomes Programme (NCAPOP), overseen by the Healthcare Quality Improvement Partnership (HQIP). NELA was officially launched in 2012 but most acute trusts did not start data entry until 2013. Data entry was sporadic in year 1 which is why 2013 was not used in data collection, as it may not provide an accurate representation of emergency laparotomy workload.

2.3 CT sarcopenia assessment

Once patients had been identified their perioperative CT sarcopenia was performed. Assessed was carried out at L3 level on the first slice in which both transverse processes were visible. Sarcopenia was assessed in two ways, using average psoas density (PD) as well as psoas area (PA). PD was calculated by the creation of a 'region of interest' around each psoas muscle on the chosen slice. The imbedded software then calculated an average density for each psoas muscle in Hounsfield units. The average of these values was taken to give the PD value as previously described (Buettner et al., 2016, Mok et al., 2016)[1]. PA was calculated by again creating a region of interest around both psoas muscles. The software displays an area for both and an average value is taken, Figure 58. For PA standardisation for body surface area is needed to give a psoas area in cm^2 per m^2 of body surface area. This normalisation of psoas area requires the documentation of anthropometric data in the form of height, weight and BMI. Patients in whom this anthropometric data were not available were excluded from PA assessment. The lowest quartile was used to define the sarcopenic group. This could be considered arbitrary. However, as there are no published cut-off values for patients from the United Kingdom undergoing emergency laparotomy this was felt to be wiser.

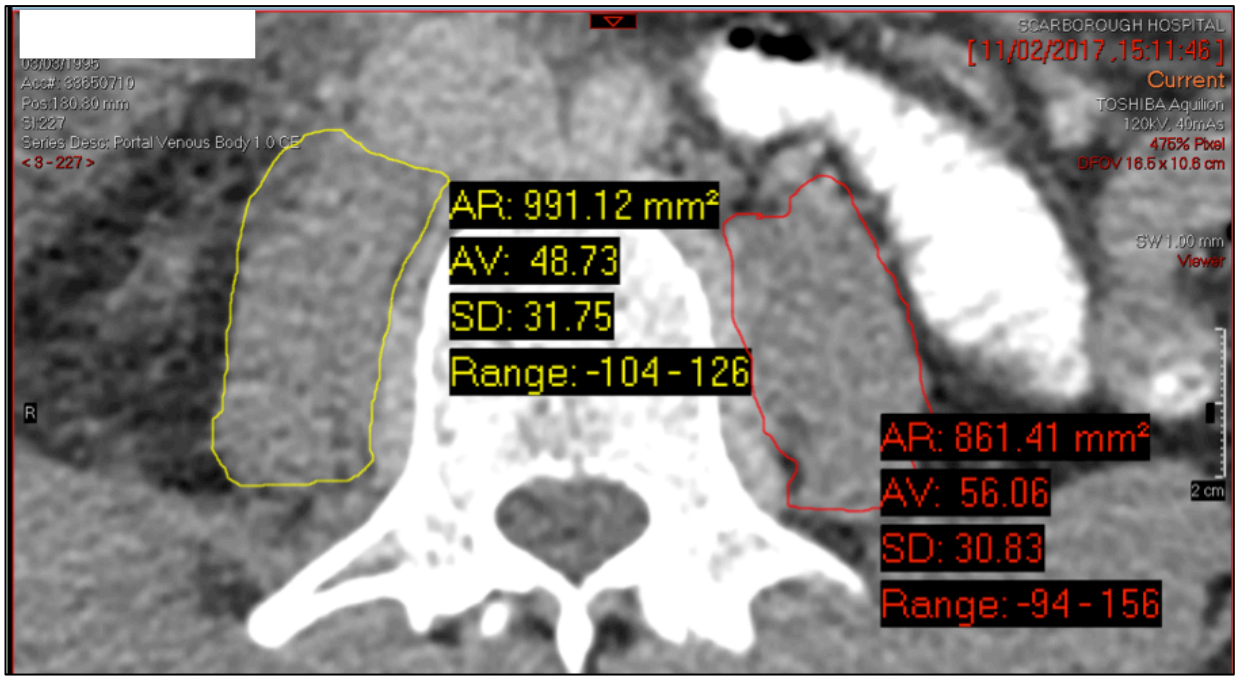


Figure 58: Measurement of PA and PD

Scarborough and York sarcopenia assessment was undertaken by myself following training by a consultant radiologist. CT sarcopenia analysis from Bradford was undertaken by another researcher who was similarly trained. Analysis from Derby was undertaken by a number of different researchers. These were trained by instruction by myself, some not in person. The Sarcopenia analysis from this site was therefore validated before inclusion in the final dataset.

2.4 Inter and Intra-operator Variability

Inter and intra-operator variability was analysed by asking 8 different operators to calculate the PA and PD of 10 different CTs. For this analysis 2 consultants, 2 registrars (SpR), 2 senior house officers (SHO) and 2 house officers (HO) were used. Each operator was asked to complete the analysis for each patient 3 separate times. The operators that were picked to do this were not part of the team who performed the data collection. This was done to give a worst-case scenario view of what the variability between operators may be. It is reasonable to assume that there will be a learning curve to

analysis of PA and PD on CT. Not using the more experienced team who had calculated these values on the hundreds of patients for the study gives an insight into what the error or variance rate may be for a clinical using this method for the first time in actual practice following a brief training episode.

2.5 Data Management and Statistical Analysis

A Microsoft Excel spreadsheet (Excel for Windows, Microsoft Corporation, Redmond, Washington, USA) was created that contained fields of the data needed and sent of secure NHS trust email servers to the 4 sites. Patients were identified from the NELA database and the relevant data gathered from hospital data systems and entered into the spreadsheet along with calculated PA and PD readings. The complete datasets for each trust were then sent back and combined into a single spreadsheet. Before these were transmitted all patient identifiable data were removed and no such information was sent out of each parent trust.

Once collated data were analysed using the IBM Statistical Package for the Social Sciences (SPSS) version 24. To assess for the distribution of the data, a histogram was drawn. This enabled easy identification of whether the data being analysed was normally distributed or skewed. Where it is unclear if a histogram is displaying normal distribution a Shapiro-Wilk test will be used with a value of >0.05 denoting a normal distribution.

Demographics of the two groups were compared for any significant difference with a p value set at 0.05. Data were analysed for difference in demographics between sarcopenic and non-sarcopenic groups as and for differences between sites. Normally distributed data were analysed with either a Chi-Squared test with Yates'

continuity correction or an independent samples T test. Non-normally distributed data were analysed by a Mann-Whitney U test.

Demographics analysed were:

- Age
- Sex
- Malignancy
- Preoperative predicted P-POSSUM mortality
- BMI

Analysis of differences in outcome between sarcopenic and non-sarcopenic groups were analysed with binary logistic regression. An odds ratio (OR) was calculated with 95% confidence intervals where significance was identified.

Outcomes measured were:

- 30 day Mortality
- 1 year Mortality
- Loss of independent living.

Loss of independent living was defined as; a patient who had been admitted from their own home that was discharged from their acute bed to either permanent or temporary supported living. This did not include patients who had been discharged home with support. Patients who had been admitted from supported living or who did not survive to discharge were not included in this analysis.

Outcomes such as length of stay, admission to critical care or return to critical care were not included. Data were available for this but as this study is multi-site and institution they were not used. Each institution has different admission criteria for intensive care. Some place all acute laparotomies in at least level 2 care post operatively and some only when needed. Any analysis for these softer endpoints may then represent differences between studied sites rather than any frailty of the patients studied.

Kaplan-Meier analysis was performed to generate a survival curve for the sarcopenic and non-sarcopenic groups as defined by PA and PD analysis. Follow up for some patients was larger than one year so survival analysis was continued until all patients had either died or been censored.

The cut off for sarcopenia was taken as the bottom quartile of the PA and PD calculations. To ascertain if a different cut off would be more clinically useful as a screening tool, a receiver-operating characteristic (ROC) curve was calculated for PA and PD. This gave an optimum trade-off between specificity and sensitivity that may or may not be of use.

Inter-operator variability was found for each CT scan by calculating the coefficient of variance that is reported as a percentage. Statistical analysis for significant variation was calculated by either ANOVA testing or by Kruskal-Wallis testing depending on the distribution of the data. Inter-operator variability was calculated for all grades of the team and for the consultant and registrar grades alone as these would be the likely ones to complete such analysis in clinical practice. Intra-operator variability was calculated in the same fashion but analysis for statistical significance was not possible as each operator calculated only 3 readings per CT.

3. Results

3.1 Intra-Operator and Inter-Operator Variability of CT Sarcopenia Calculation

Inter-Operator Variability

Each of the 8 operators calculated the PA and PD for 10 CT scans. Inter-operator variability is displayed in Table 9 as mean coefficient of variance across all grades and for consultant and registrar grades only. Figure 59 is a scatter plot of the mean percentage variation for each scan (PD and PA combined) for all grades and for consultant and registrar grades only. Horizontal line represents the mean percentage error across all scans. Mean inter-operator variability was 6.38% for consultant and registrar graded only and 9.83% for all grades.

Scan Number	All Grades	Consultants and Registrars Only
1	4.2%	3.0%
2	11.2%	5.5%
3	19.8%	7.5%
4	3.9%	3.5%
5	10.4%	9.1%
6	11.9%	14.7%
7	9.5%	4.7%
8	8.3%	4.5%
9	12.6%	3.6%
10	6.5%	5.7%

Table 9: Mean variance for each CT

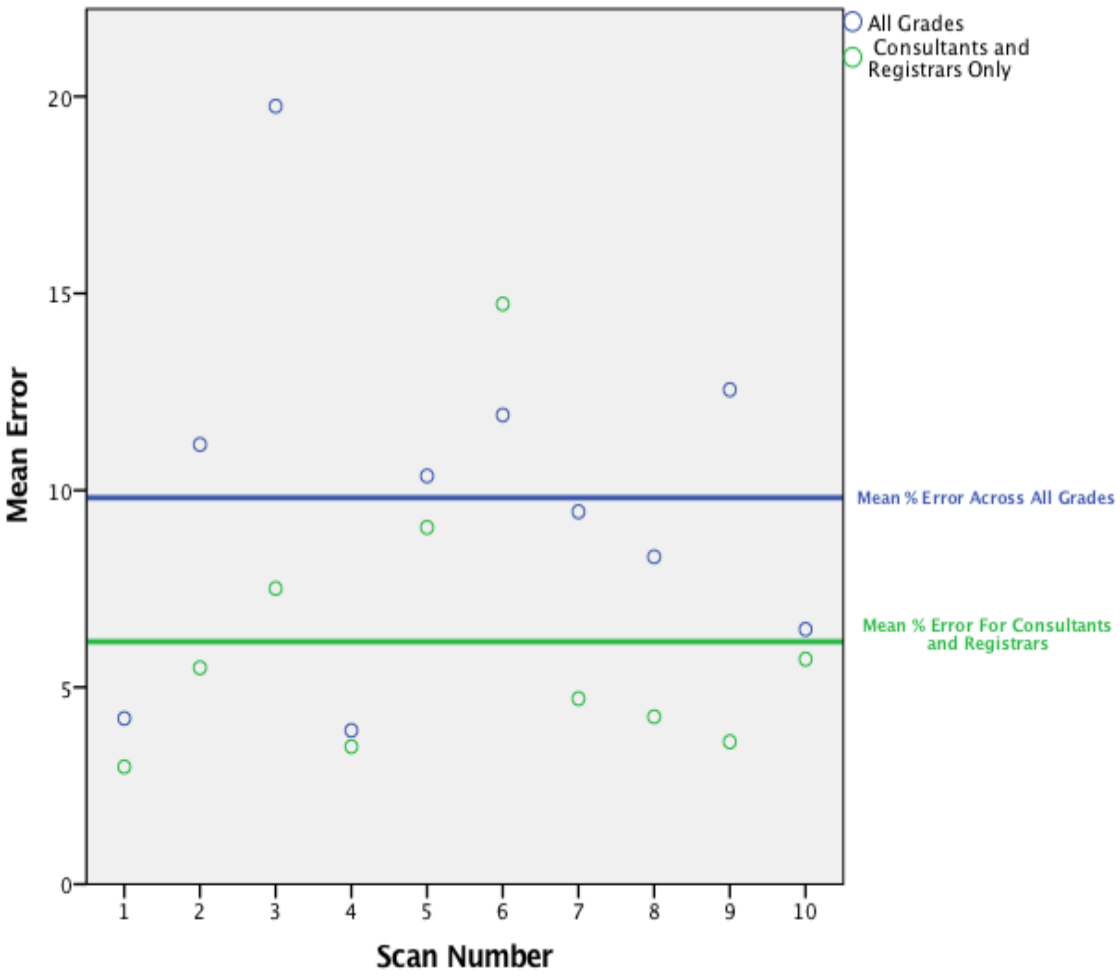


Figure 59: Mean percentage variation for all grades vs. consultant and registrars only

Table 10 and 11 display statistical significance of the variation in readings across all grades for PA and PD measurements respectively.

Scan Number	Shapiro-Wilk Result	Test Used	P-Value
1	Normal Distribution	One-Way ANOVA	<0.001
2	Non-Normal Distribution	Kruskal-Wallis	0.007
3	Non-Normal Distribution	Kruskal-Wallis	0.004
4	Normal Distribution	One-Way ANOVA	0.541
5	Normal Distribution	One-Way ANOVA	<0.001
6	Non-Normal Distribution	Kruskal-Wallis	0.005
7	Normal Distribution	One-Way ANOVA	0.009
8	Normal Distribution	One-Way ANOVA	0.109
9	Non-Normal Distribution	Kruskal-Wallis	0.012
10	Non-Normal Distribution	Kruskal-Wallis	0.065

Table 10: Significance of variation PA

Scan Number	Shapiro-Wilk Result	Test Used	P-Value
1	Non-Normal Distribution	Kruskal-Wallis	0.161
2	Normal Distribution	One-Way ANOVA	<0.001
3	Normal Distribution	One-Way ANOVA	0.009
4	Normal Distribution	One-Way ANOVA	<0.001
5	Normal Distribution	One-Way ANOVA	<0.001
6	Non-Normal Distribution	Kruskal-Wallis	0.005
7	Non-Normal Distribution	Kruskal-Wallis	0.011
8	Non-Normal Distribution	Kruskal-Wallis	0.049
9	Non-Normal Distribution	Kruskal-Wallis	0.007
10	Normal Distribution	One-Way ANOVA	<0.001

Table 11: Significance of variation PD

Table 12 and 13 display statistical significance of the variation across consultant and registrar grades only for PA and PD measurements respectively.

Scan Number	Shapiro-Wilk Result	Test Used	P-Value
1	Normal Distribution	One-Way ANOVA	0.034
2	Normal Distribution	One-Way ANOVA	0.527
3	Normal Distribution	One-Way ANOVA	0.480
4	Normal Distribution	One-Way ANOVA	0.204
5	Normal Distribution	One-Way ANOVA	0.065
6	Non-Normal Distribution	Kruskal-Wallis	0.025
7	Normal Distribution	One-Way ANOVA	0.209
8	Normal Distribution	One-Way ANOVA	0.346
9	Normal Distribution	One-Way ANOVA	0.481
10	Normal Distribution	One-Way ANOVA	0.437

Table 12: Significance of variation PA consultants and registrars only

Scan Number	Shapiro-Wilk Result	Test Used	P-Value
1	Normal Distribution	One-Way ANOVA	0.219
2	Normal Distribution	One-Way ANOVA	0.001
3	Normal Distribution	One-Way ANOVA	0.437
4	Normal Distribution	One-Way ANOVA	0.070
5	Normal Distribution	One-Way ANOVA	0.010
6	Non-Normal Distribution	Kruskal-Wallis	0.032
7	Normal Distribution	One-Way ANOVA	0.707
8	Non-Normal Distribution	Kruskal-Wallis	0.871
9	Normal Distribution	One-Way ANOVA	0.037
10	Normal Distribution	One-Way ANOVA	0.114

Table 13: Significance of variation PA consultants and registrars only

Intra-Operator Variability

Each operator calculated the PD and PA for each of the 10 scans 3 times. The coefficient of variance was calculated to give a percentage variance for each operator. Table 14 displays Mean percentage variance for all 10 scans for PA and PD calculation. Mean intra-operator variability was 3% for PA and 3.9% for PD.

Operator	Psoas Area	Psoas Density
Cons1	2.0%	2.8%
Cons2	3.6%	3.9%
Spr1	4.6%	7.1%
Spr2	3.7%	3.9%
SHO1	3.6%	3.0%
SHO2	1.9%	2.8%
FY1 1	1.9%	3.5%
FY1 2	2.95	4.1%

Table 14: Mean intra-operator variability for each operator

3.2 Total Data Collected for Analysis

Total number of patients collected from each site is shown in table 15. Numbers analysed from Scarborough, York and Bradford for PD were larger than for PA due to inconsistent recording of height and weight on the varying hospital systems. As these demographics are required to normalise PA for body surface area, patients without this recorded could not be analysed. Data on where patients had been admitted from and discharged to was only available from York NHS FT systems (Scarborough and York Sites). The data collected from Derby represented 20 rather than 12 months of activity at that site.

Site	Number analysed for PD	Number analysed for PA
Scarborough	87	80
York	172	168
Bradford	146	0
Derby	440	440
Total	845	688

Table 15: Numbers of patients from each site

3.3 Variance of Demographics Between sites

Due to population variance between the areas each site serves there is likely to be demographic differences between sites. The variance between sites of the key demographics is displayed in table 16. There is no BMI data for Bradford, as their hospital systems do not routinely store this information.

	Scarborough	York	Bradford	Derby	P-Value
Mean Age (SD)	69.5 (13.7)	66.1 (16.9)	62.9 (16.5)	61.7 (18.2)	<0.001
BMI (SD)	27 (7.2)	25.8 (6)	-	26.1 (5.8)	0.2
Malignancy	18.4%	21.5%	19.9%	18.4%	0.84
Sex (M:F)	46:41	79:93	76:70	209:231	0.57
Mean P-possum Mortality (SD)	20.8 (25)	17.2 (22.2)	17.5 (23.3)	20.4 (23.7)	0.079

Table 16: Variance of demographics between sites

Age

The age of the patients from each site is not normally distributed data and has therefore been analysed with the Kruskal-Wallis Test. Between all 4 sites there is a significant difference in age, $p<0.001$. Post Hoc analysis between individual sites (Bonferroni adjustment applied to give a p value for statistical significance of <0.008) reveals that there is significant difference of age between, Scarborough and Bradford; Scarborough and Derby; York and Derby.

BMI

The BMI of the patients from each site is normally distributed data and therefore has been analysed using ANOVA testing. This did not reveal any significant difference between sites.

Malignancy

The rate of malignancy of the patients from each site has been analysed using binary logistic regression. This did not reveal any significant difference between sites.

Sex

The sex of the patients from each site has been analysed using binary logistic regression. This did not reveal any significant difference between sites.

P-Possum

The pre-operative predicted P-possum mortality of the patients from each site has been analysed with the Kruskal-Wallis Test. This did not reveal any significant difference between sites.

3.4 Psoas Density Analysis

Data on 845 patients was available for psoas density analysis for mortality at 30 days and 1 year. Data on where these patients had been admitted from and discharged to was available on 259 patients (Scarborough and York Sites). As previously detailed the bottom quartile of psoas muscle density measurement were deemed Sarcopenic. Within the 845 patients analysed for psoas density this left a sarcopenic population of 211 and non-sarcopenic of 634.

3.41 Demographics

The key demographics of the sarcopenic and non-sarcopenic population, as defined by PD, are displayed in table 17.

Demographic	Sarcopenic	Non-Sarcopenic	P-value
Age (SD)	71.8 (12.9)	61 (17.9)	<0.001
BMI (SD)	26.8 (5.3)	25.9 (6.3)	0.102
Malignant diagnosis	23.2%	18%	0.116
Sex M:F	111:100	299:335	0.197
P-Possum mortality (SD)	27.8% (27)	16.5% (21.7)	<0.001

Table 17: Demographics PD

Age

Figure 60 is a stacked histogram displaying the non-normal distribution of age in the analysed population. Figure 61 displays the mean age for the sarcopenic and non-sarcopenic groups with bars displaying the 95% CI. Data were analysed using the Mann-Whitney U test.

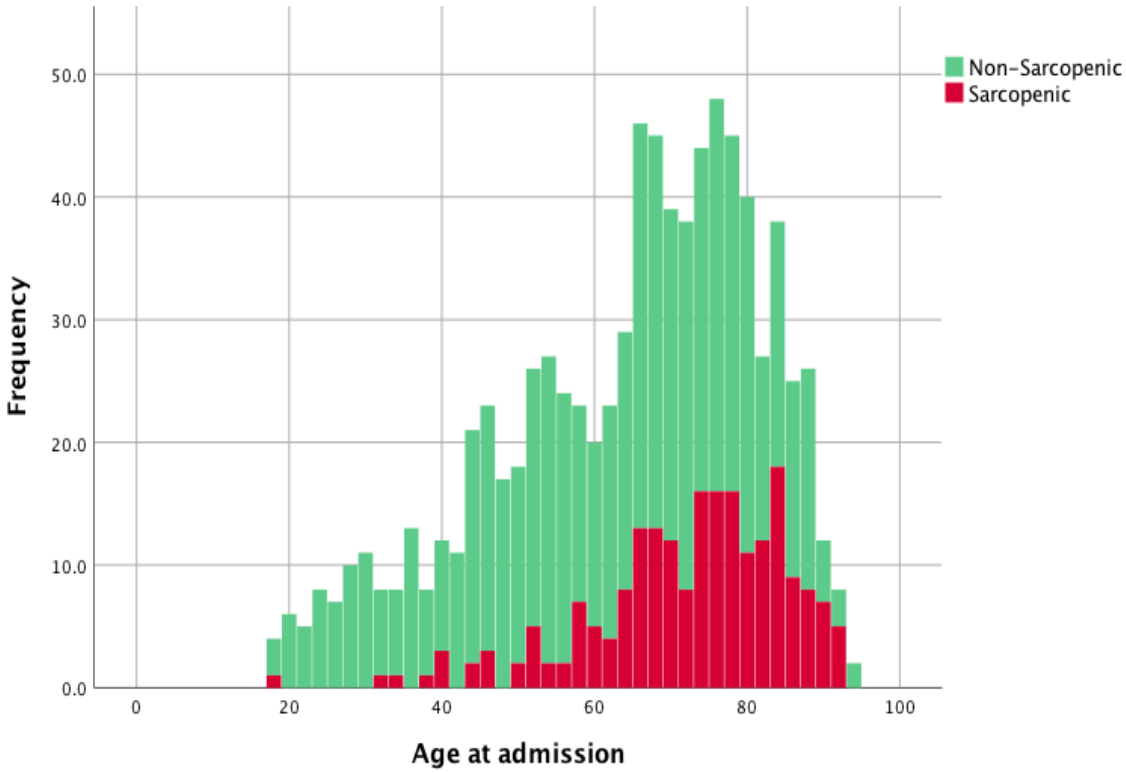


Figure 60: Stacked histogram of ages for PD

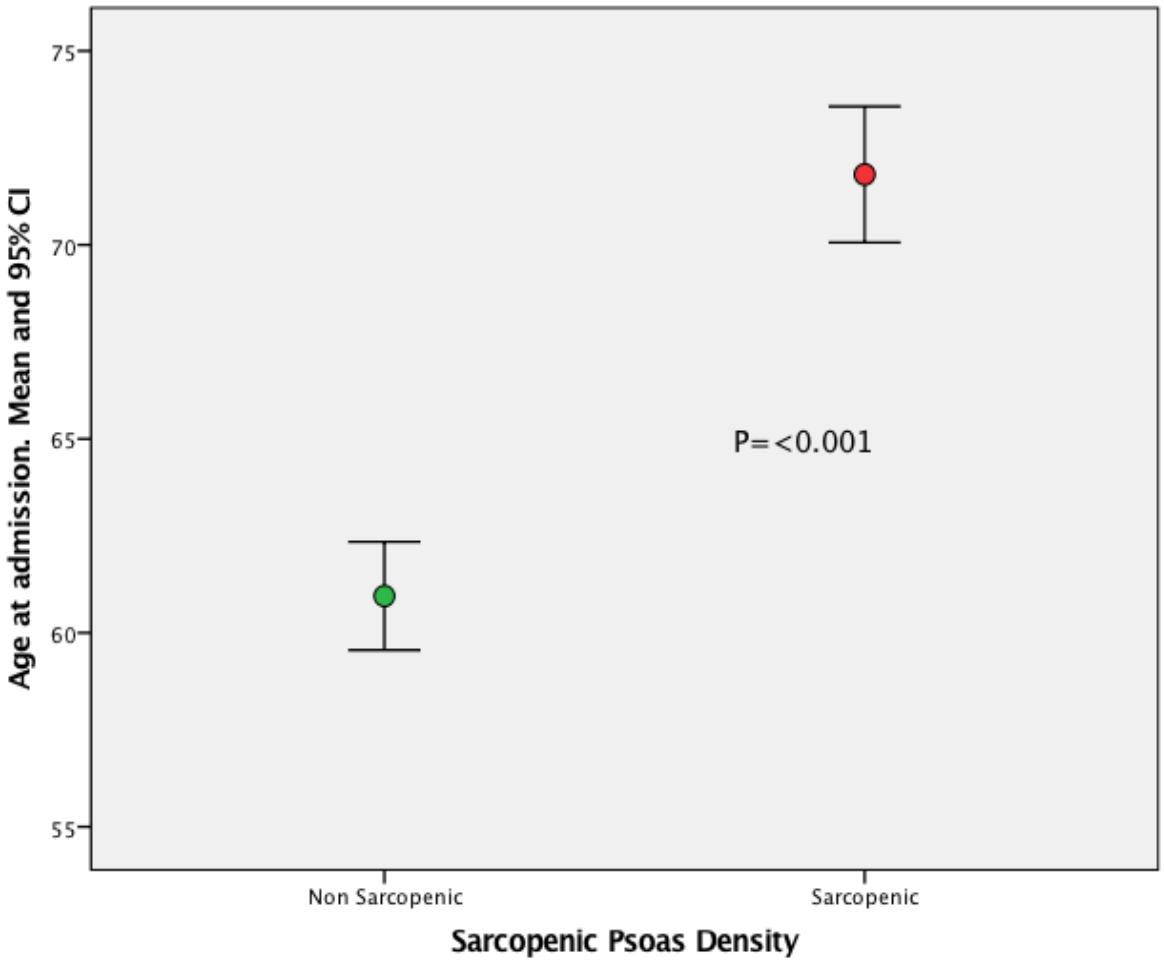
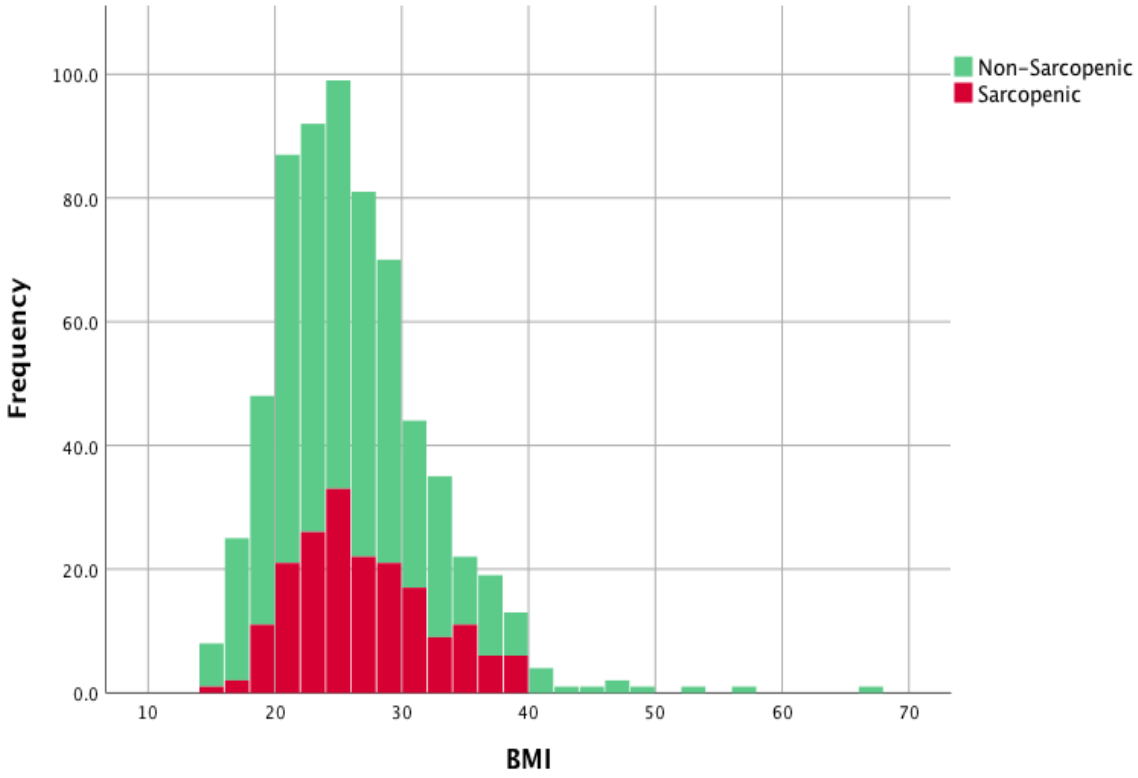


Figure 61: Mean age of sarcopenic and non-sarcopenic, PD

BMI

Figure 62 is a stacked histogram displaying the normal distribution of BMI in the analysed population. Figure 63 displays the mean BMI for the sarcopenic and non-sarcopenic groups with bars displaying the 95% CI. The BMI of the patients from each has been analysed with a T-test. This did not reveal any significant difference between sites.



62: Stacked Histogram of BMI for PD

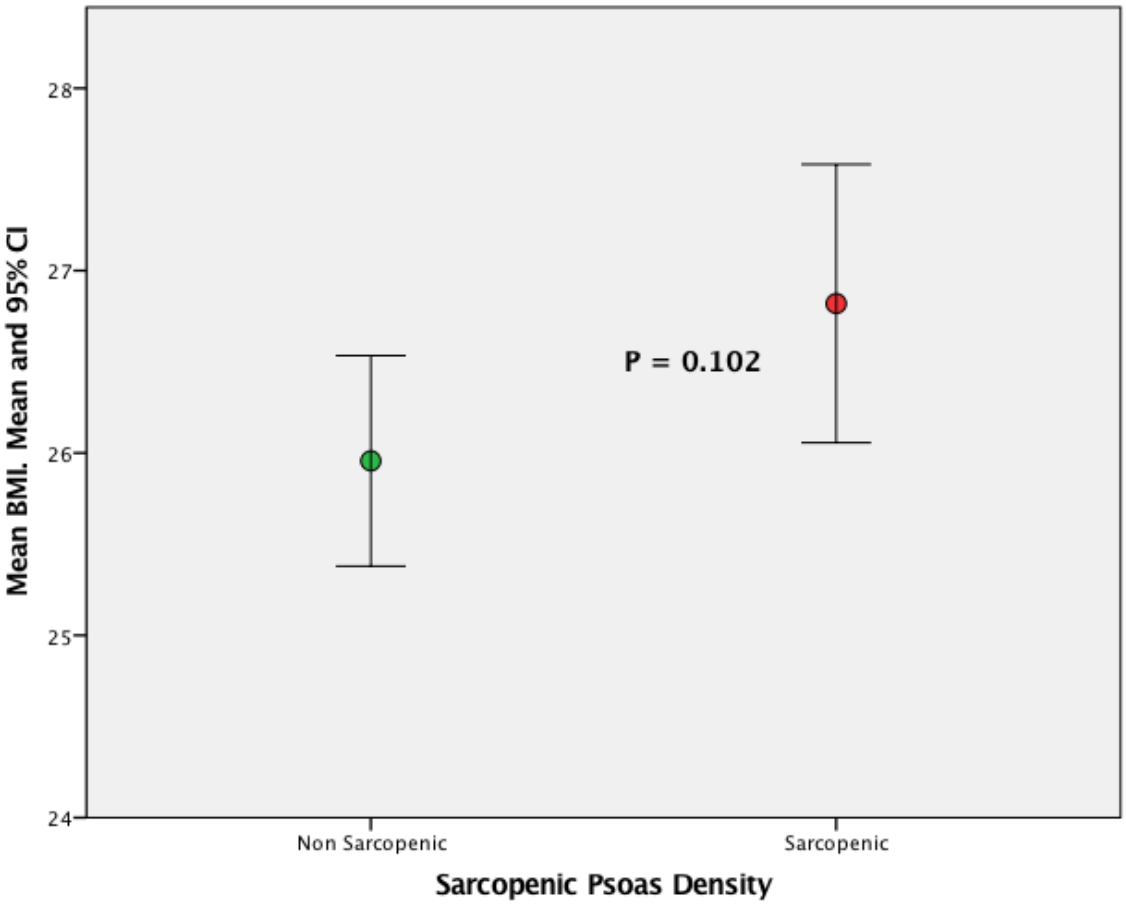


Figure 63: Mean BMI of sarcopenic and non-sarcopenic, PD

Malignant Histology

Figure 64 displays the percentage of the Sarcopenic and Non-Sarcopenic groups that had a malignant histology. A Chi-Square Test for independence (with Yates Continuity Correction) indicated no significant association between Sarcopenia and Mortality at 30 days, $\chi^2 = 2.47$, $p = 0.116$

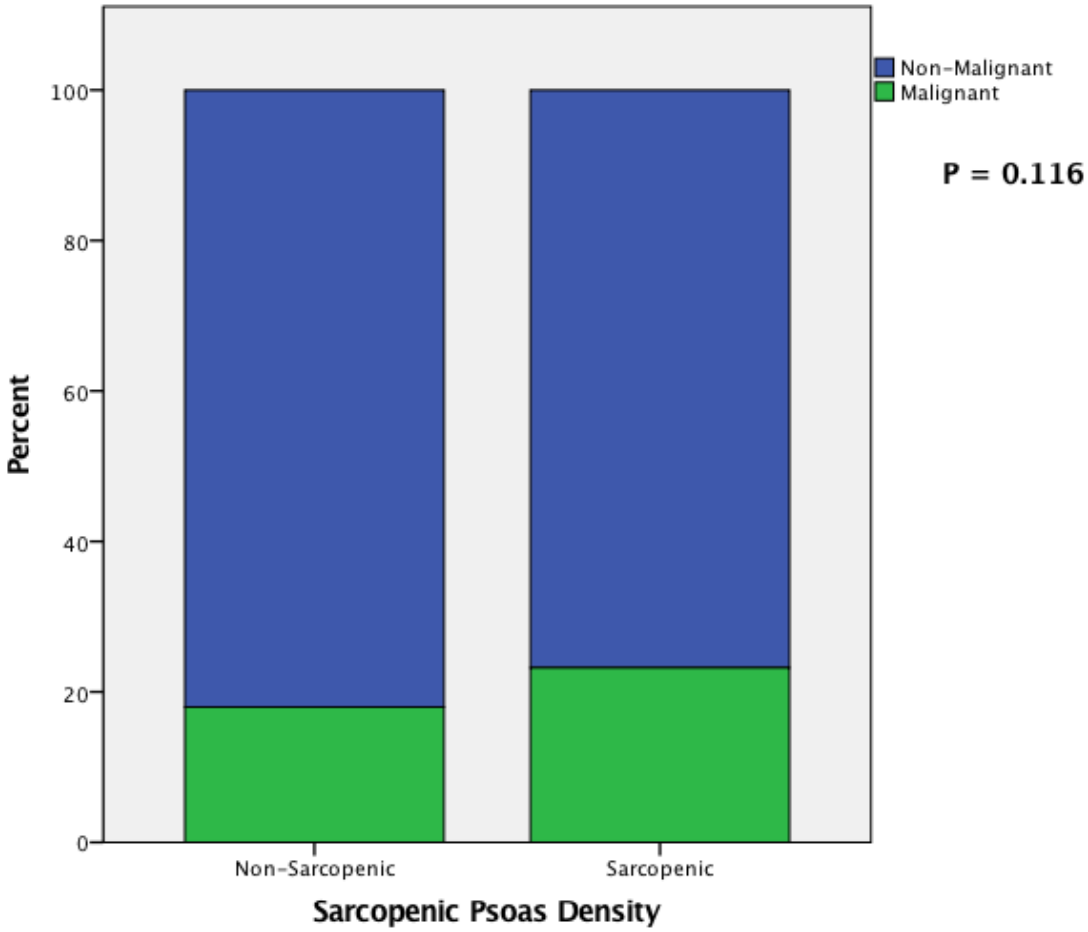


Figure 64: Percentage malignancy of non-sarcopenic and sarcopenic groups, PD

Sex

Figure 65 displays the percentage of the Sarcopenic and Non-Sarcopenic groups that were male or female. A Chi-Square Test for independence (with Yates Continuity Correction) indicated no significant association between Sarcopenia and Mortality at 30 days, $\chi^2 = 1.67$, $p = 0.197$

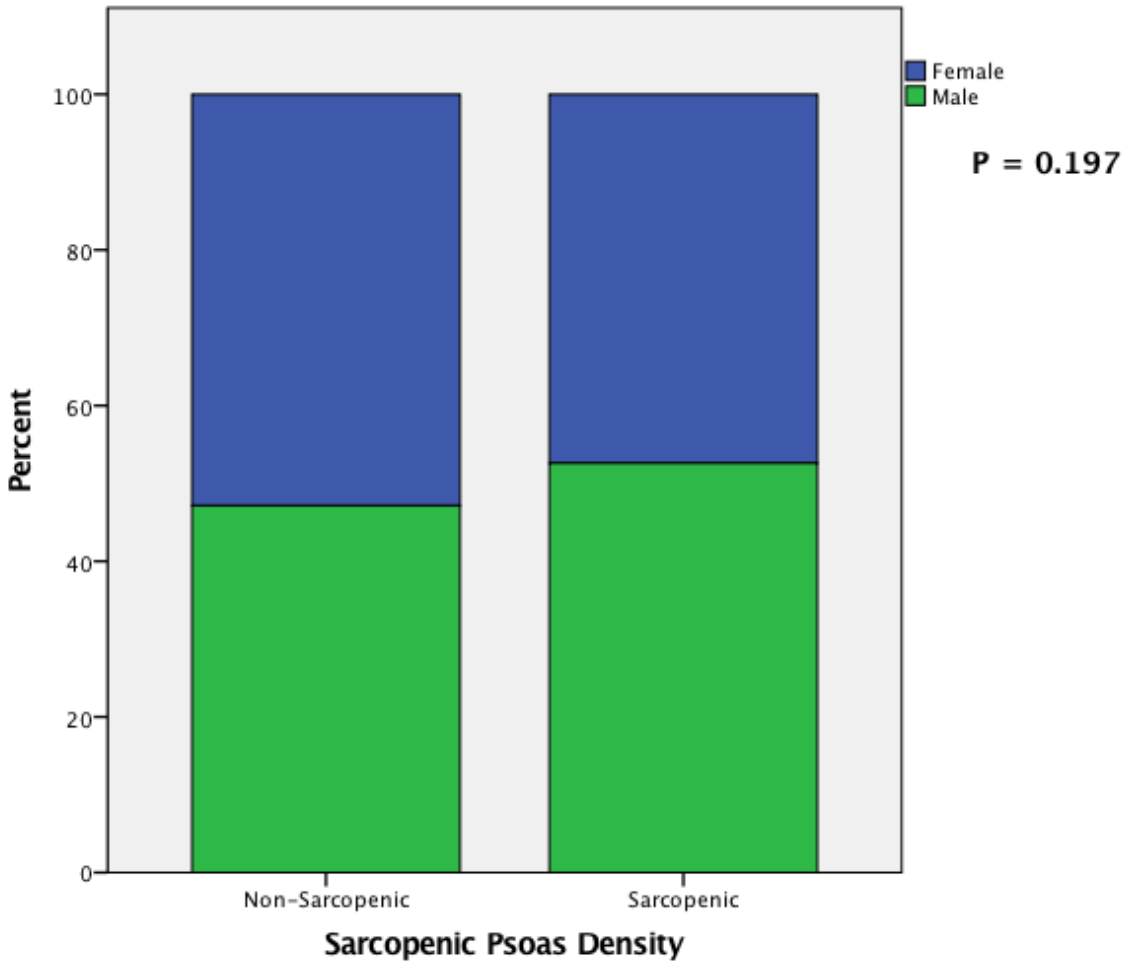


Figure 65: Percentage of male vs. female for non-sarcopenic ad sarcopenic groups, PD

Predicted P-Possum Mortality

Figure 66 is a stacked histogram displaying the non-normal distribution of p-possum mortality in the analysed population. Figure 67 displays the mean P-possum mortality for the sarcopenic and non-sarcopenic groups with bars displaying the 95% CI. Data were analysed using the Mann-Whitney U test.

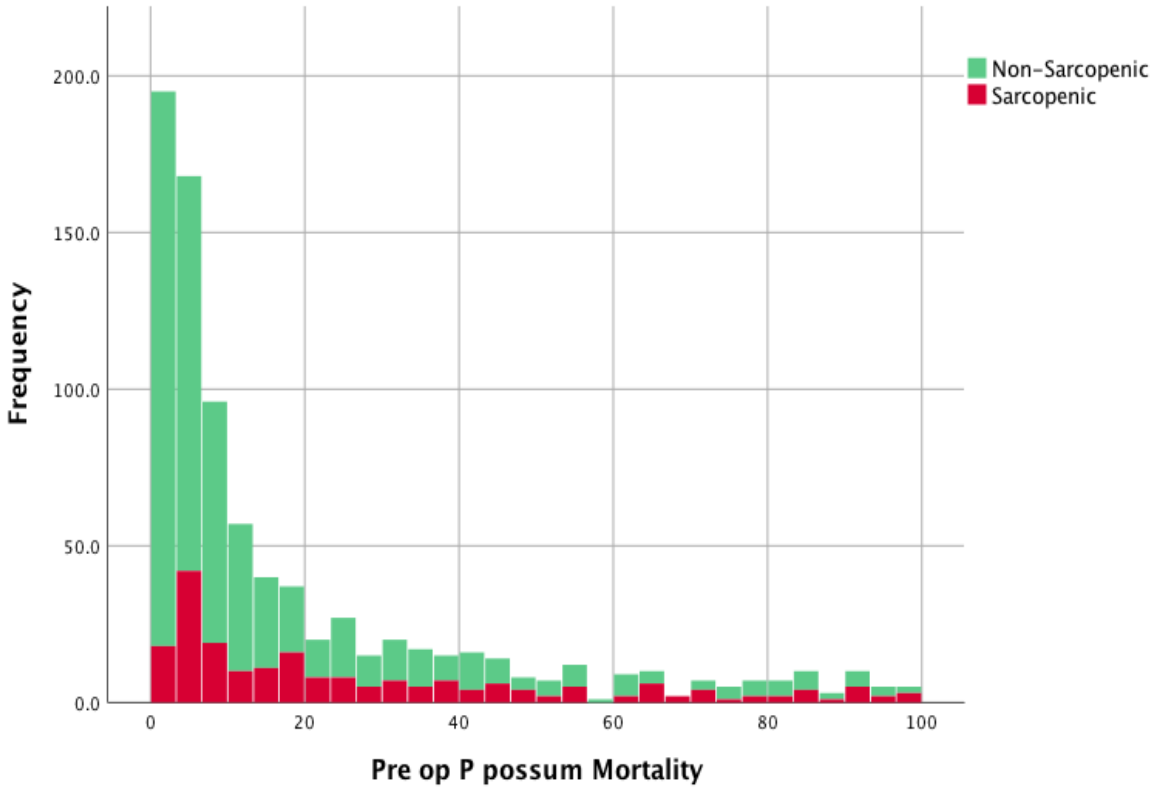


Figure 66: Stacked histogram of P-Possum scores, PD

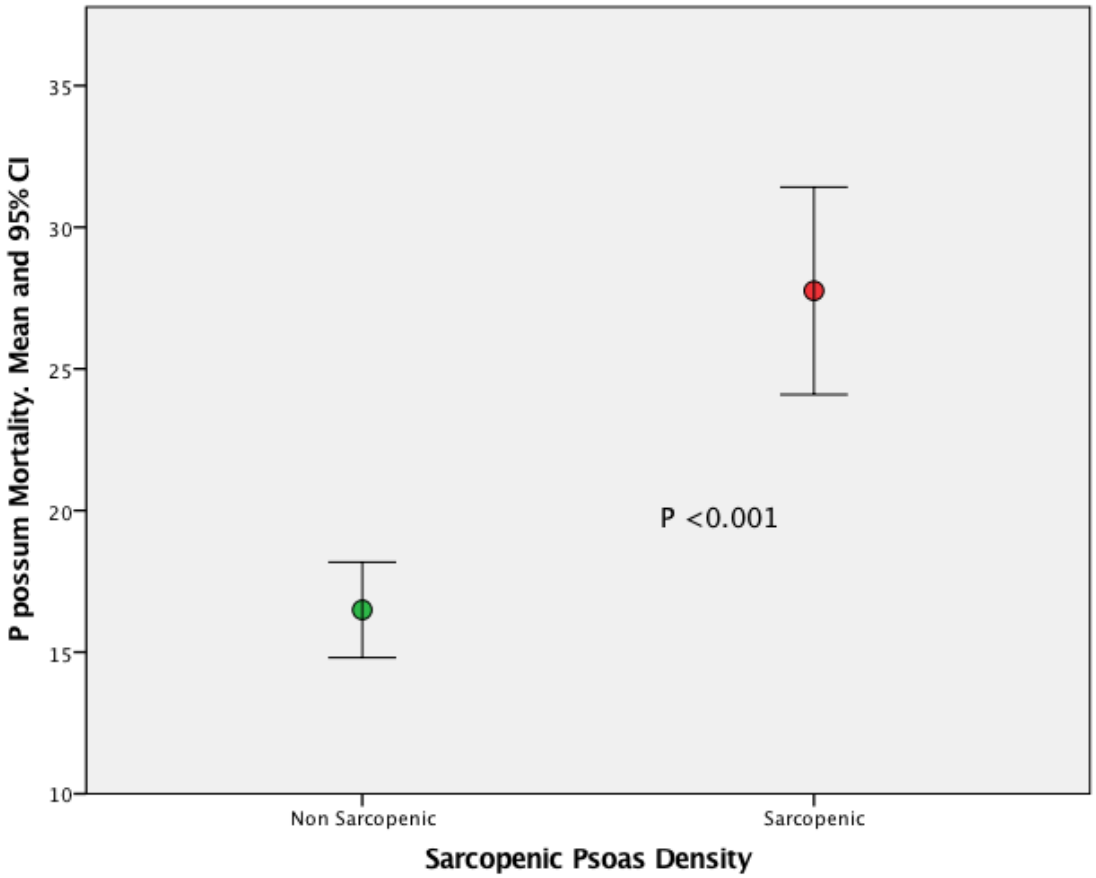


Figure 67: Mean of P-Possum score for non-sarcopenic group vs. sarcopenic group, PD

3.42 30-Day Mortality

Table 18 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by psoas density, that were alive or dead at 30 days. Figure 68 displays this as a bar chart. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 30-day mortality with OR of 2.84 with 95% CI of 1.88-4.29, P<0.001

	Sarcopenic	Non-Sarcopenic
Dead at 30 days	23.8% (49/211)	9.6% (61/634)
Alive at 30 days	76.2% (162/211)	90.4% (573/634)

Table 18: 30-day mortality PD

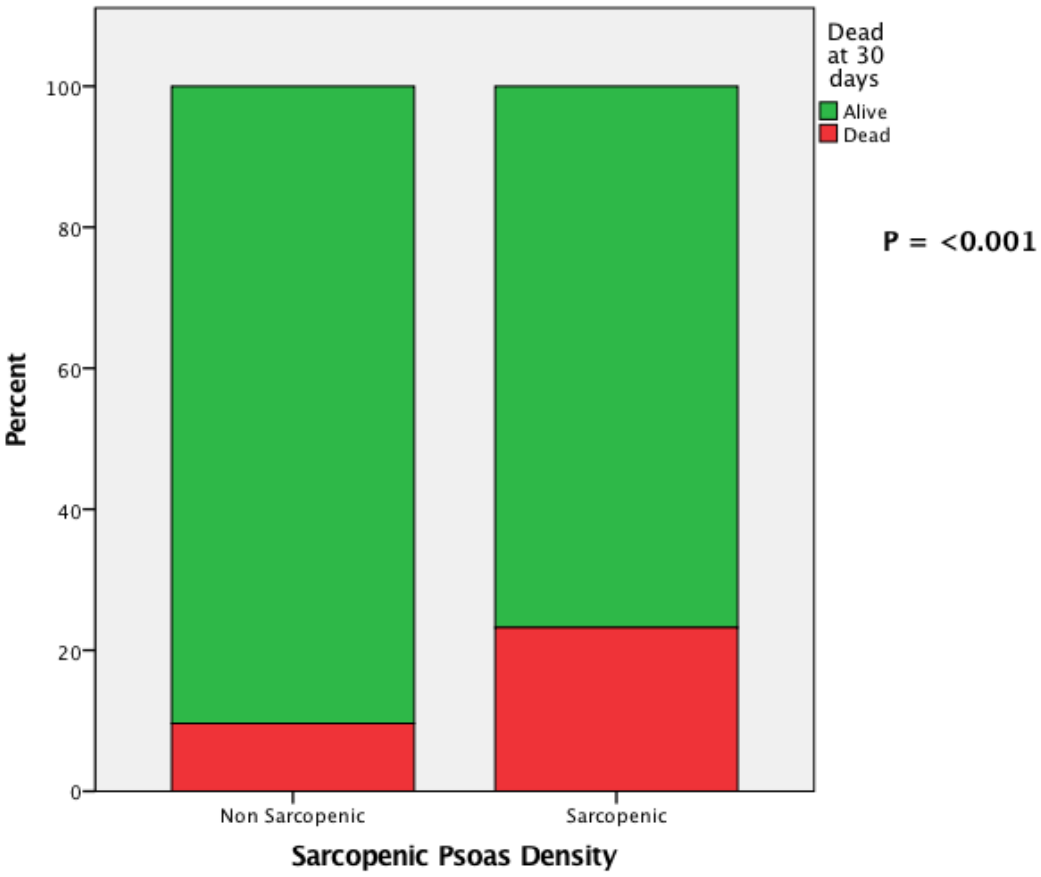


Figure 68: 30-day mortality sarcopenic vs. non-sarcopenic, PD

3.43 1-Year Mortality

Table 19 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by psoas density, that were alive or dead at 1-year. Figure 69 displays this as a bar chart. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 1-year mortality with OR of 2.46 with 95% CI of 1.75-3.46, P<0.001

	Sarcopenic	Non-Sarcopenic
Dead at 1-Year	37% (78/211)	19.2% (122/634)
Alive at 1-Year	63% (133/211)	80.8% (512/634)

Table 19: 1-year mortality PD

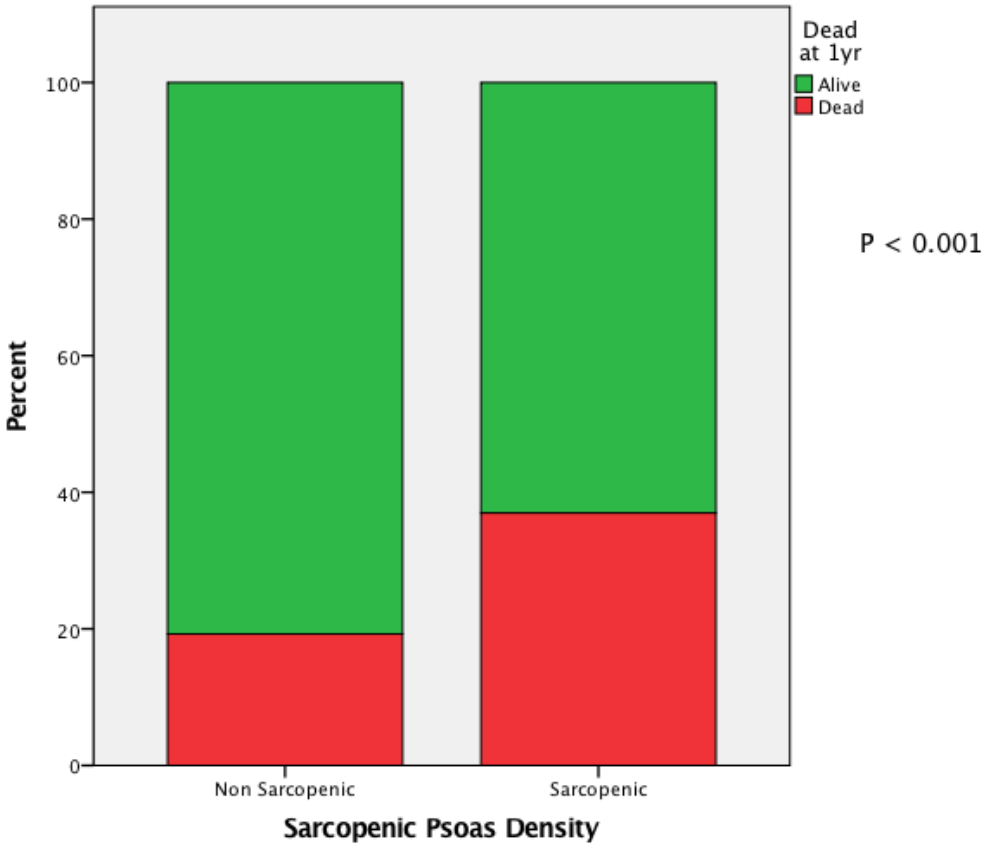


Figure 69: 1-year mortality sarcopenic vs. non-sarcopenic, PD

3.44 Multi Variant Analysis

As displayed in table 17, there is a significant difference in the rates of age and p-possum mortality between the sarcopenic and non-sarcopenic groups defined by psoas density.

30-day mortality

Binary logistic regression analysis was performed to assess the impact of sarcopenia on 30-day mortality adjusting for predicted p-possum mortality and age. The adjusted odds ratio for sarcopenia defined by psoas density is 1.83 with a 95% CI of 1.16 – 2.87, P = 0.008.

1-Year Mortality

Binary logistic regression analysis was performed to assess the impact of sarcopenia on 1-year mortality adjusting for predicted p-possum mortality and age. The adjusted odds ratio for sarcopenia defined by psoas density is 1.58 with a 95% CI of 1.09 – 2.29, P = 0.016.

3.45 Survival Analysis

Kaplan-Meier survival analysis for the sarcopenic and non-sarcopenic groups as defined by psoas density displays a significant difference between the survival curve for the sarcopenic and non-sarcopenic groups, P < 0.001. Figure 70 displays the first 100 days where the survival curve for the sarcopenic group falls away rapidly from the non-sarcopenic group. Figure 71 displays the curve out to 1000 days of follow up where, although not as rapidly divergent, the survival curve for the sarcopenic group is still falling away from the non-sarcopenic group.

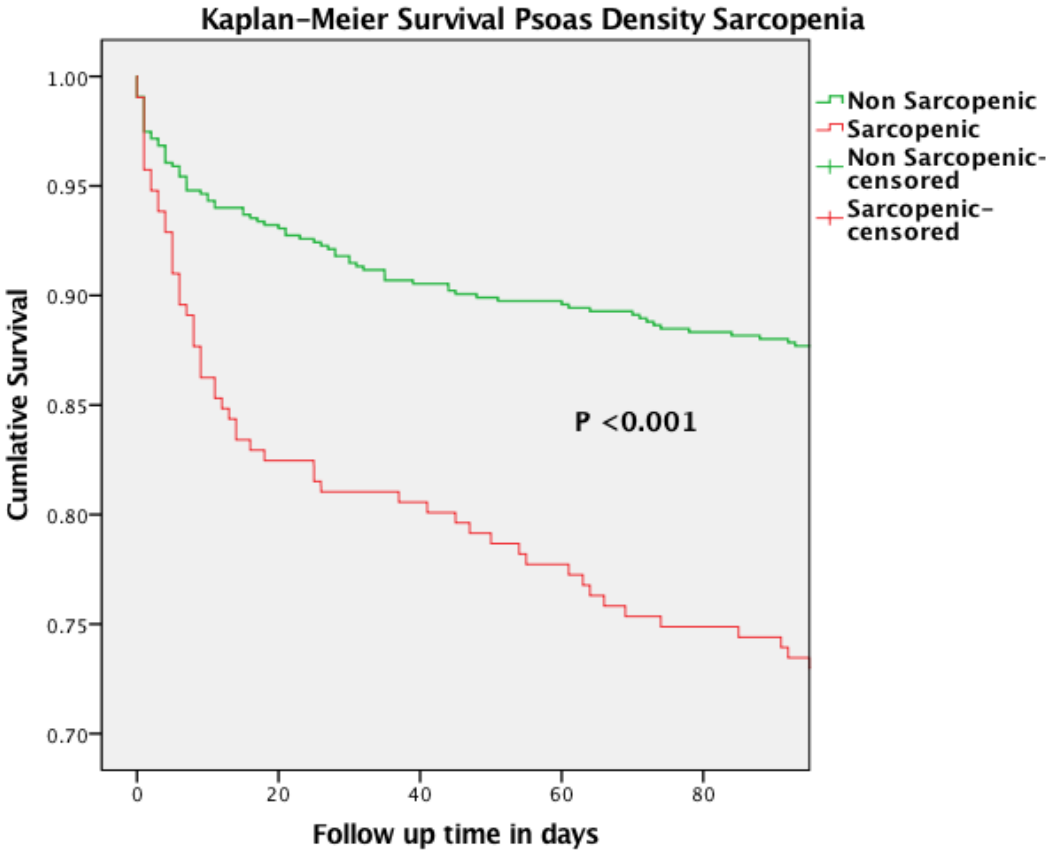


Figure 70: Kaplan-Meier survival 1st 100 days, PD

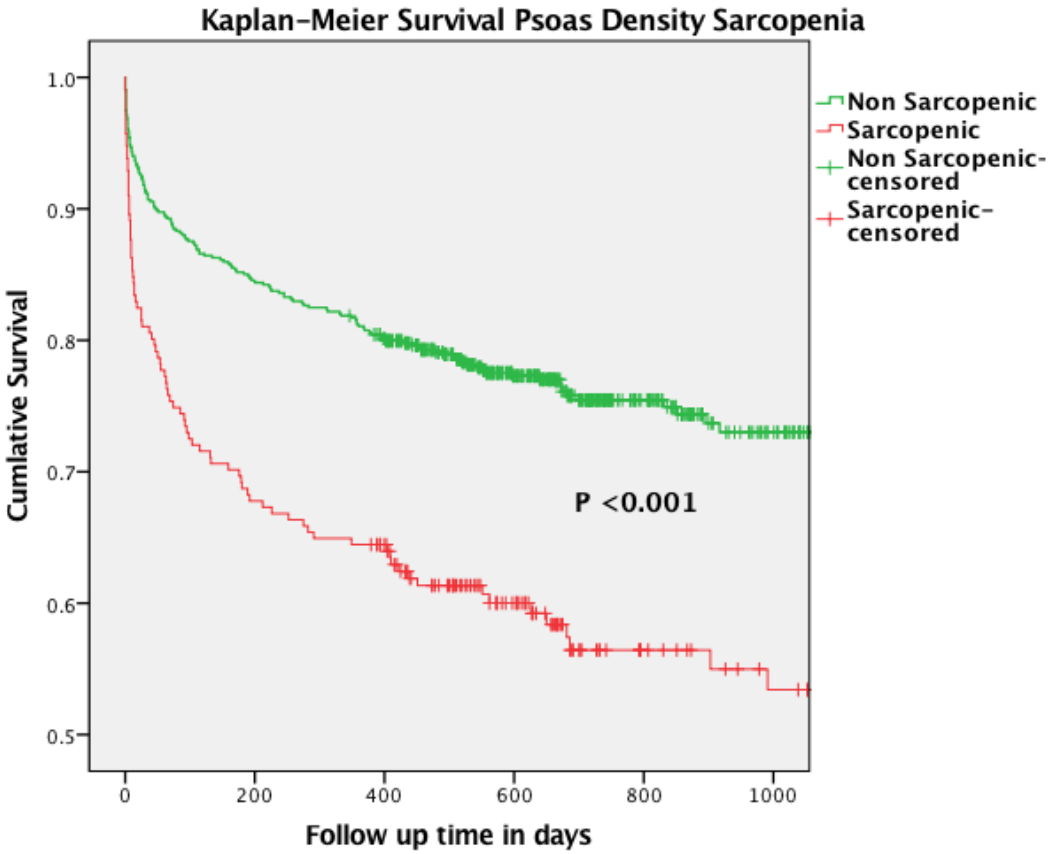


Figure 71: Kaplan-Meier survival 1000 days, PD

3.46 Receiver Operator Characteristic Curve PD

To ascertain the effectiveness of PD as a screening tool to detect mortality within the follow-up period, a ROC curve was calculated as displayed in figure 72. This shows an area under the graph of 0.640 with 95% CI 0f 0.551-0.682, p = <0.001. Ideal trade off between sensitivity and specificity was a PD of 35.99HU. This would put the sarcopenic population in the 845 patients analysed at 422 or 50%.

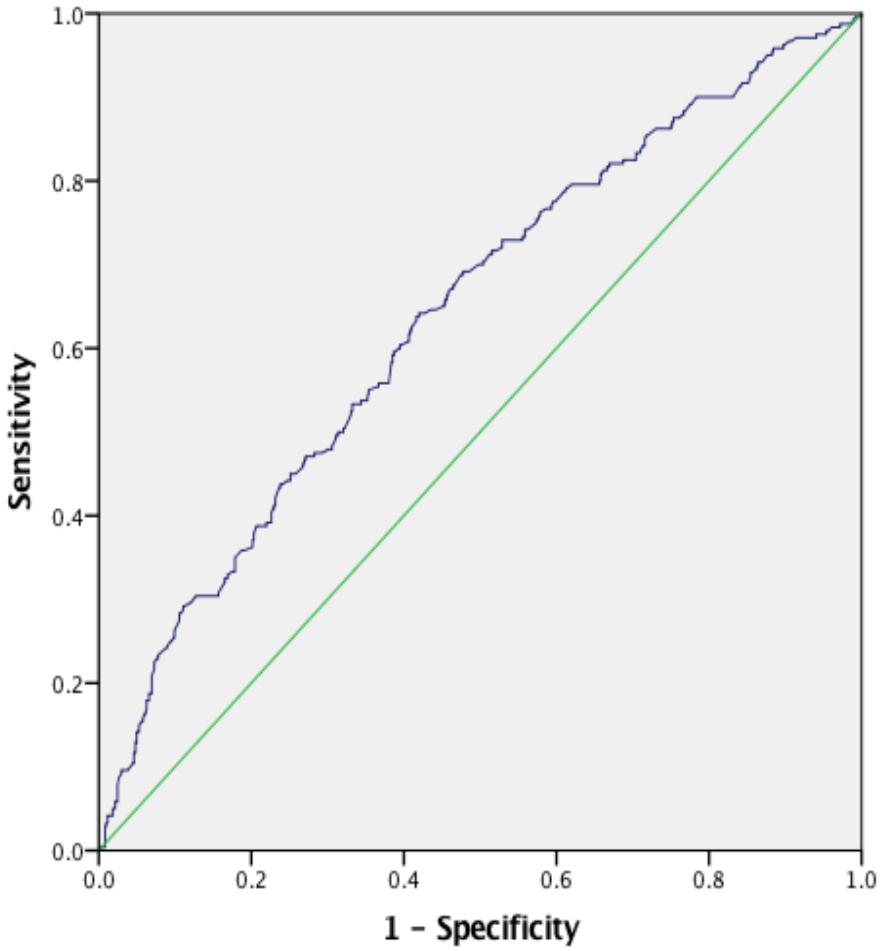


Figure 72: ROC curve PD

Table 20 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by ROC curve derived cut-off for psoas density, that were alive or dead at 30 days. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 30-day mortality with OR of 2.22 with 95% CI of 1.43-3.32, P<0.001

Table 21 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by ROC curve derived cut-off for psoas density, that were alive or dead at 1-year. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 1-year mortality with OR of 2.24 with 95% CI of 1.62-3.13, P<0.001

	Sarcopenic	Non-Sarcopenic
Dead at 30 days	17.3% (73/422)	8.7% (37/423)
Alive at 30 days	82.7% (349/422)	91.3% (386/423)

Table 20: 30-day mortality ROC curve PD

	Sarcopenic	Non-Sarcopenic
Dead at 1-year	30.8% (130/422)	16.5% (70/423)
Alive at 30 days	69.2% (292/422)	83.5% (353/423)

Table 21: 1-year mortality ROC curve PD

Demographic	Sarcopenic	Non-Sarcopenic	P-value
Age (SD)	69.9 (13.6)	57.5 (18.6)	<0.001
BMI (SD)	26.8 (6.2)	25.4 (5.8)	0.003
Malignant diagnosis	22.7% (96/422)	15.8% (67/423)	0.014
Sex M:F	205:217	205:218	1.0
P-Possum mortality (SD)	24.4 (25.7)	14.3 (24.5)	<0.001

Table 22: Demographics ROC curve PD

3.47 Increased dependency

Table 20 displays the number and percentage of the sarcopenic and non-sarcopenic groups that were admitted from independent living and discharged from the acute ward to non-independent living. Figure 73 displays this as a bar chart. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and increased dependency with OR of 2.50 with 95% CI of 1.15-5.46, P = 0.021

	Sarcopenic	Non-Sarcopenic
Remains in Independent Living	78% (48/61)	90.4% (169/187)
Home to Non-Independent Living	21.3% (13/61)	9.6%% (18/187)

Table 23: Loss of independence PD

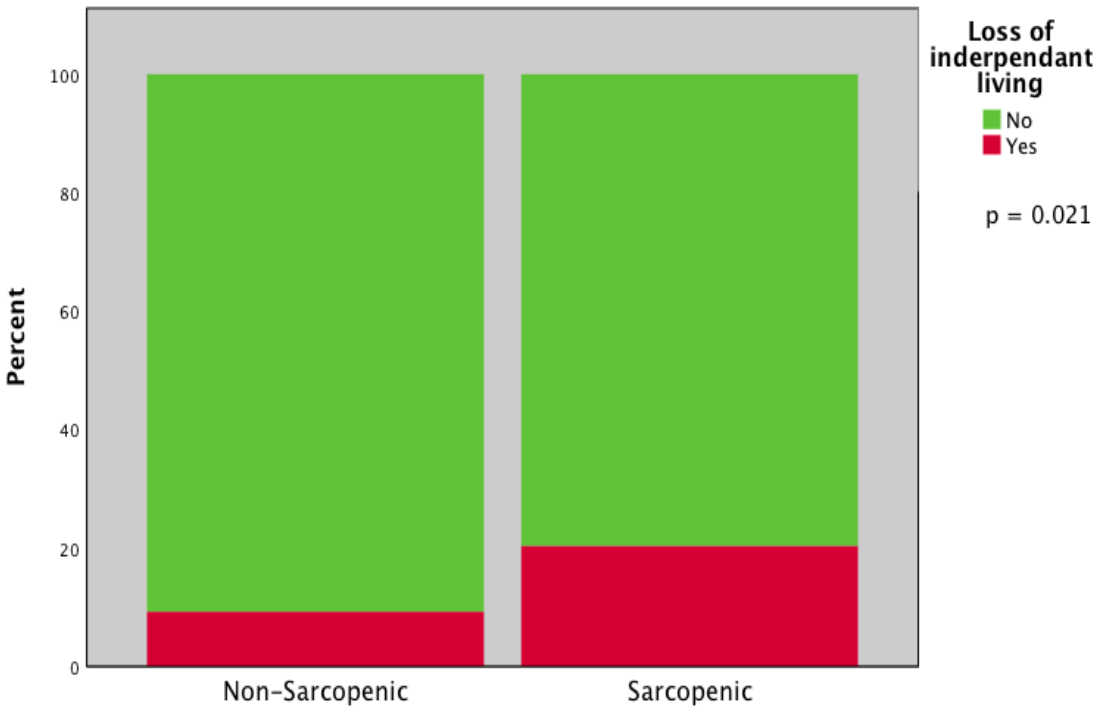


Figure 73: Loss of independent living PD

3.5 Psoas Area Analysis

Data on 687 patients was available for psoas density analysis for mortality at 30 days and 1 year. Data on where these patients had been admitted from and discharged to was available on 259 patients (Scarborough and York Sites). As previously detailed the bottom quartile of psoas muscle density measurement were deemed Sarcopenic. Within the 687 patients analysed for psoas area this left a sarcopenic population of 172 and non-sarcopenic of 515. The total patient number analysed for PA is lower than that for PD due to incomplete recording of height and weight data that included the entirety of the Bradford cohort.

3.51 Demographics

The key demographics of the sarcopenic and non-sarcopenic population, as defined by PA, are displayed in table 21.

Demographic	Sarcopenic	Non-Sarcopenic	P-value
Age (SD)	65.3 (16.3)	63.4 (17.3)	0.343
BMI (SD)	26.5 (7.1)	25.9 (5.7)	0.247
Malignant diagnosis	24.4% (42/172)	16.7% (86/515)	0.033
Sex M:F	82/90	245/270	1.0
P-Possum mortality (SD)	18.3 (20.6)	20.0 (24.5)	0.307

Table 24: Demographics PA

Age

Figure 74 is a stacked histogram displaying the non-normal distribution of age in the analysed population. Figure 75 displays the mean age for the sarcopenic and non-sarcopenic groups with bars displaying the 95% CI. Data were analysed using the Mann-Whitney U test.

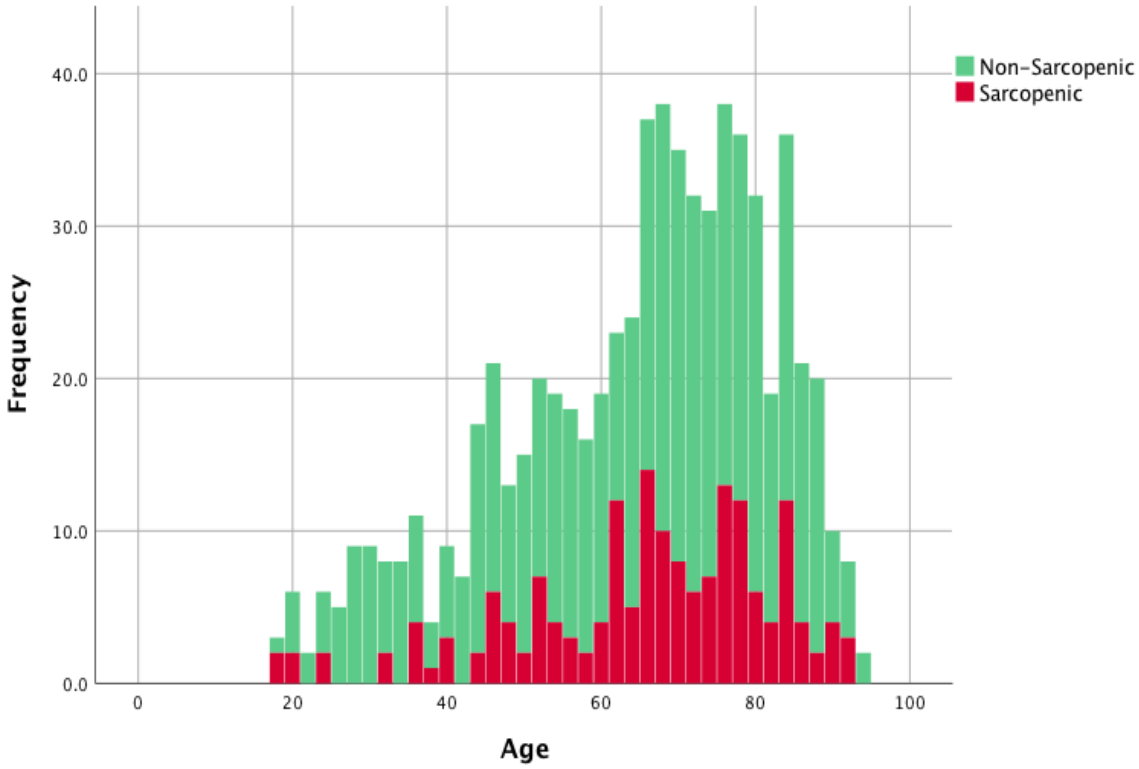


Figure 74: Stacked histogram of age PA

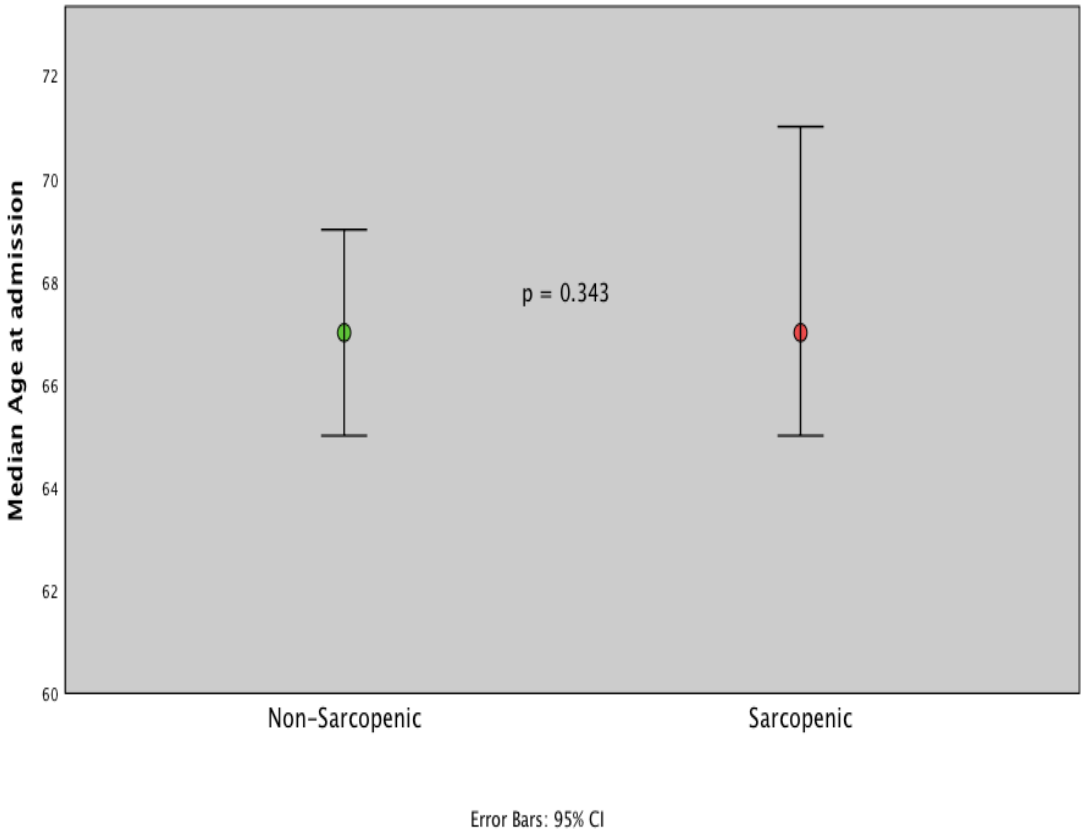


Figure 75: Median age sarcopenia vs. non-sarcopenia PA

BMI

Figure 76 is a stacked histogram displaying the normal distribution of BMI in the analysed population. Figure 77 displays the mean BMI for the sarcopenic and non-sarcopenic groups with bars displaying the 95% CI. The BMI of the patients from each has been analysed with a T-test. This did not reveal any significant difference between sites.

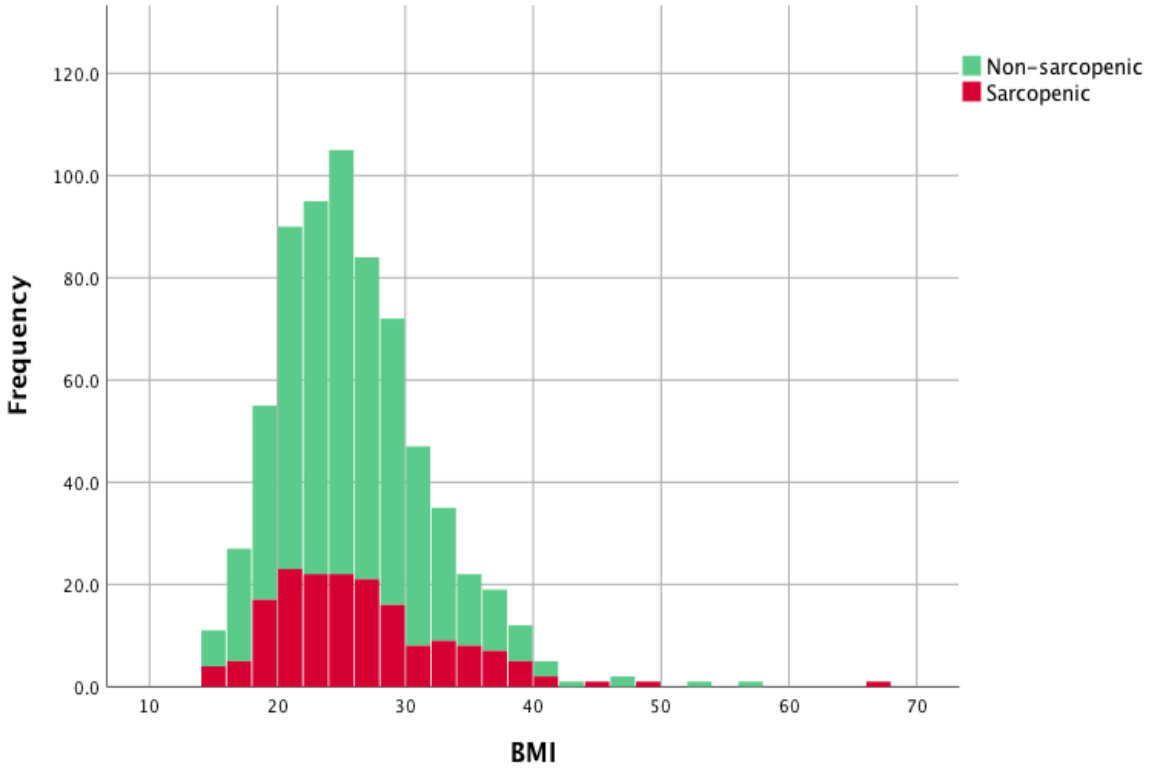


Figure 76: Stacked histogram of BMI PA

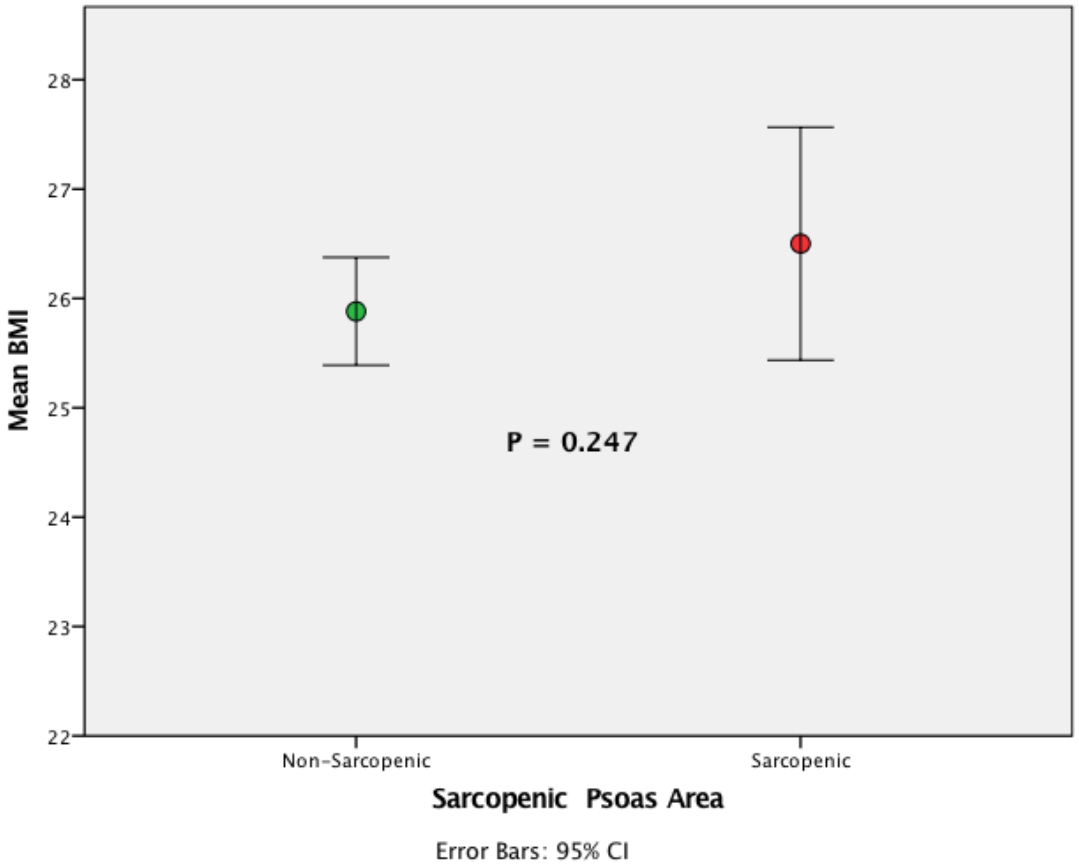


Figure 77: Mean BMI sarcopenia vs. non-sarcopenia PA

Malignant Histology

Figure 78 displays the percentage of the Sarcopenic and Non-Sarcopenic groups that had a malignant histology. A Chi-Square Test for independence (with Yates Continuity Correction) indicated no significant association between Sarcopenia and Mortality at 30 days, $\chi^2 = 4.57$, p = 0.033, Phi 0.86.

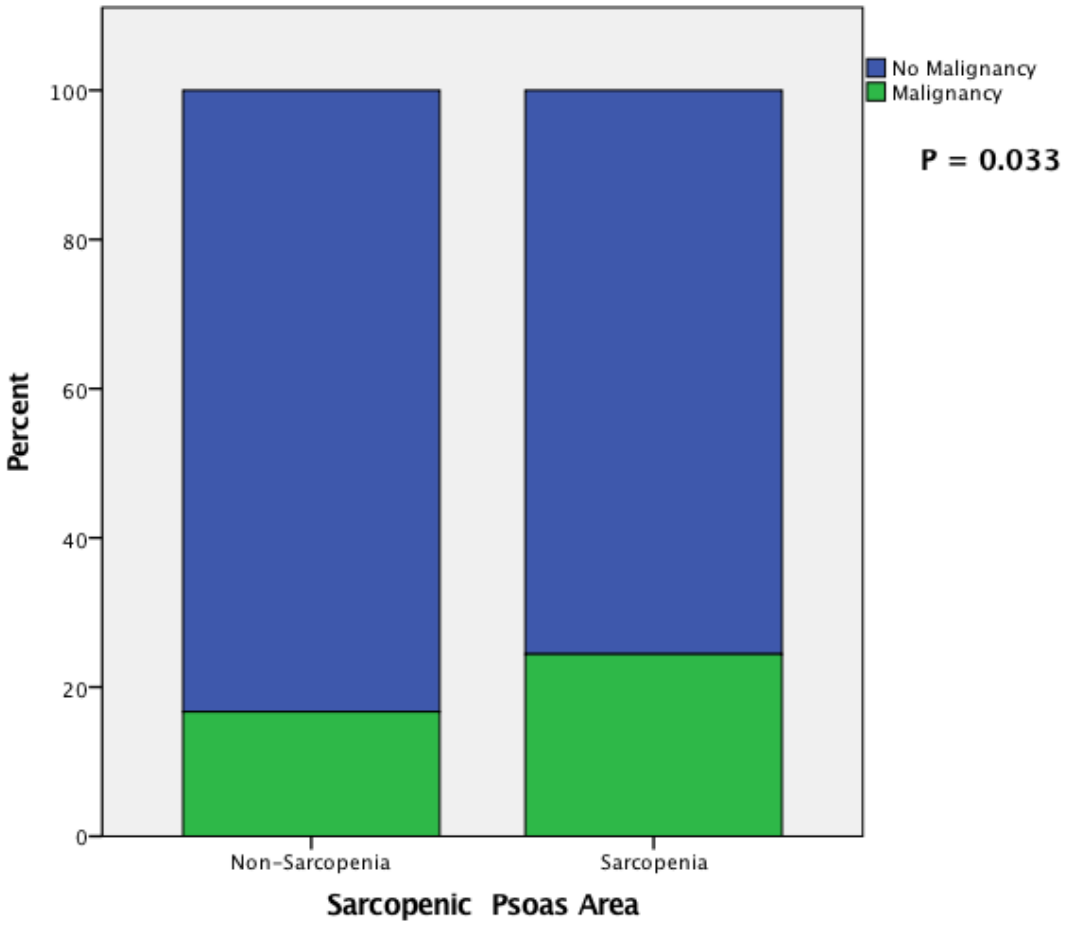


Figure 78: Malignancy sarcopenia vs. non-sarcopenia PA

Sex

Figure 79 displays the percentage of the Sarcopenic and Non-Sarcopenic groups that were male or female. A Chi-Square Test for independence (with Yates Continuity Correction) indicated no significant association between Sarcopenia and Mortality at 30 days, $\chi^2 = 0.0$, $p = 1$.

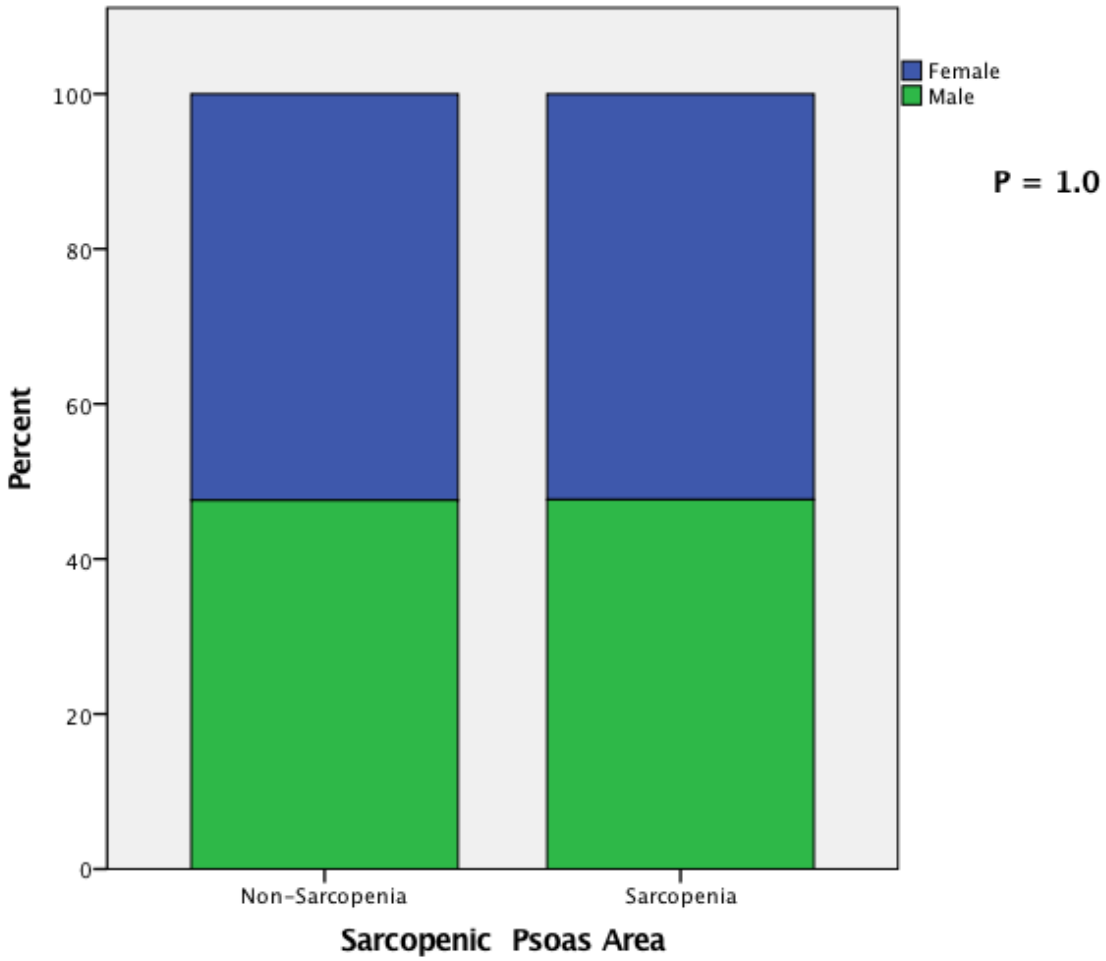


Figure 79: Sex sarcopenia vs. non-sarcopenia PA

Predicted P-Possum Mortality

Figure 80 is a stacked histogram displaying the non-normal distribution of p-possum mortality in the analysed population. Figure 81 displays the mean P-possum mortality for the sarcopenic and non-sarcopenic groups with bars displaying the 95% CI. Data were analysed using the Mann-Whitney U test.

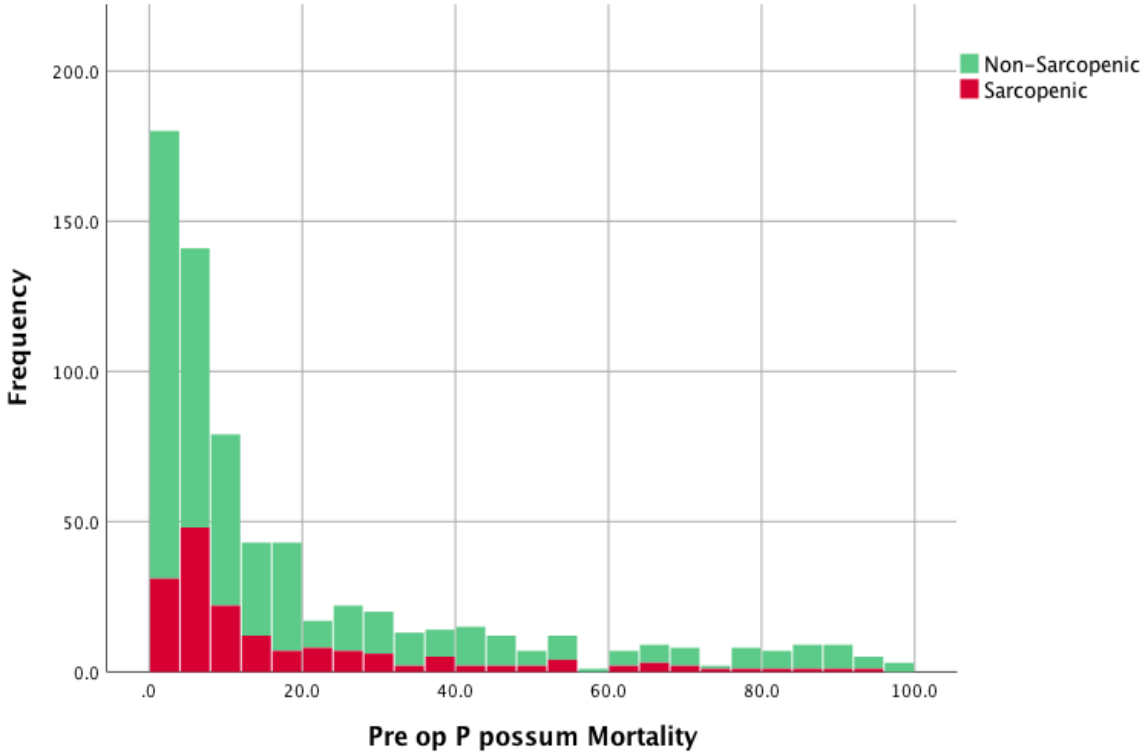


Figure 80: Stacked histogram of P-Possum PA

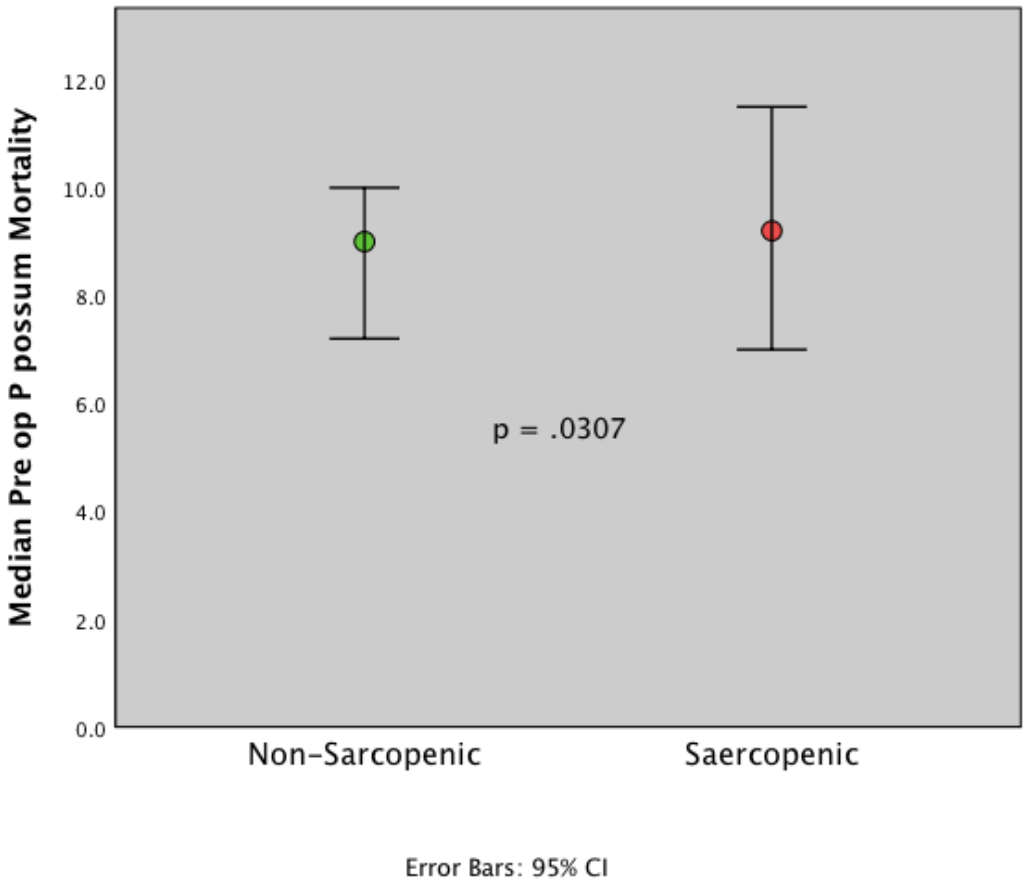


Figure 81: Median P-Possum Score sarcopenia vs. non-sarcopenia PA

3.52 30-Day Mortality

Table 22 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by psoas area, that were alive or dead at 30 days. Figure 82 displays this as a bar chart. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 1-year mortality with OR of 2.31 with 95% CI of 1.38-3.88, P = 0.002

	Sarcopenic	Non-Sarcopenic
Dead at 30 days	16.3% (28/172)	7.8% (40/515)
Alive at 30 days	83.7% (144/172)	92.2% (475/515)

Table 25: 30-day mortality PA

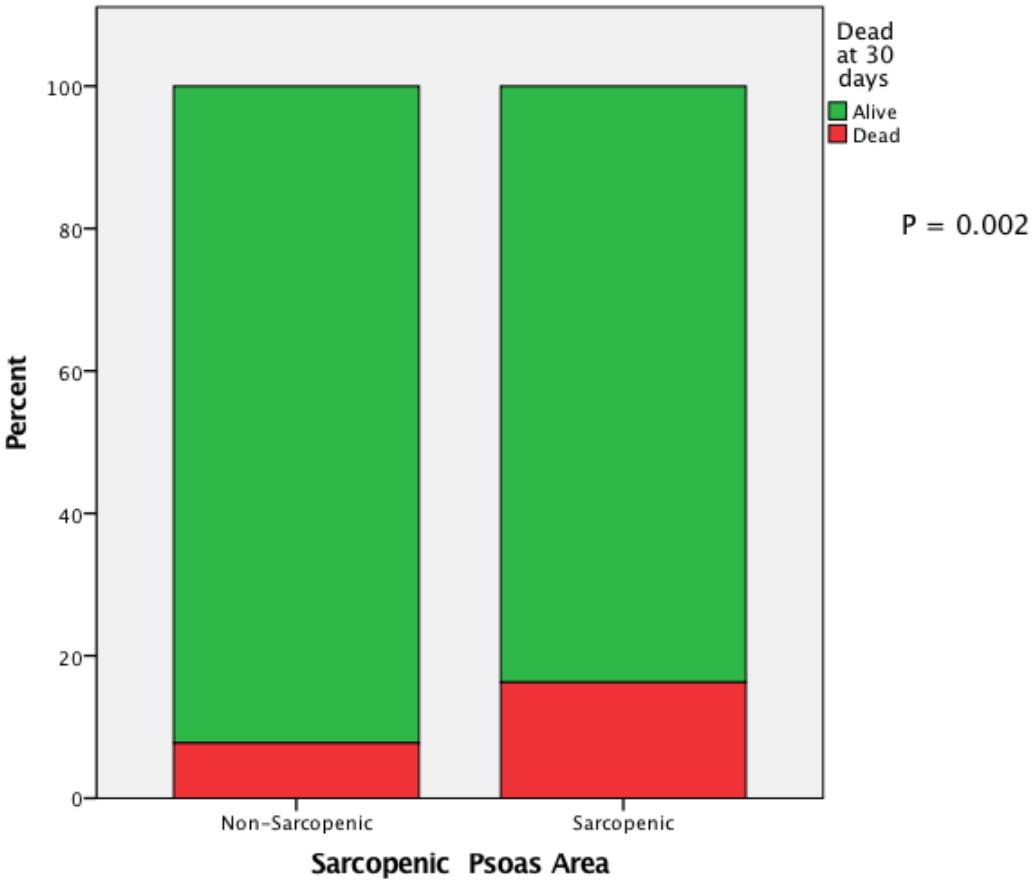


Figure 82: 30-Day mortality sarcopenia vs. non-sarcopenia PA

3.53 1-Year Mortality

Table 23 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by psoas area, that were alive or dead at 1 year. Figure 83 displays this as a bar chart. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 1-year mortality with OR of 2.25 with 95% CI of 1.52-3.33, P<0.001

	Sarcopenic	Non-Sarcopenic
Dead at 1 year	32% (55/172)	17.3% (89/515)
Alive at 1 year	68% (117/172)	82.7% (425.515)

Table 26: 1-year mortality PA

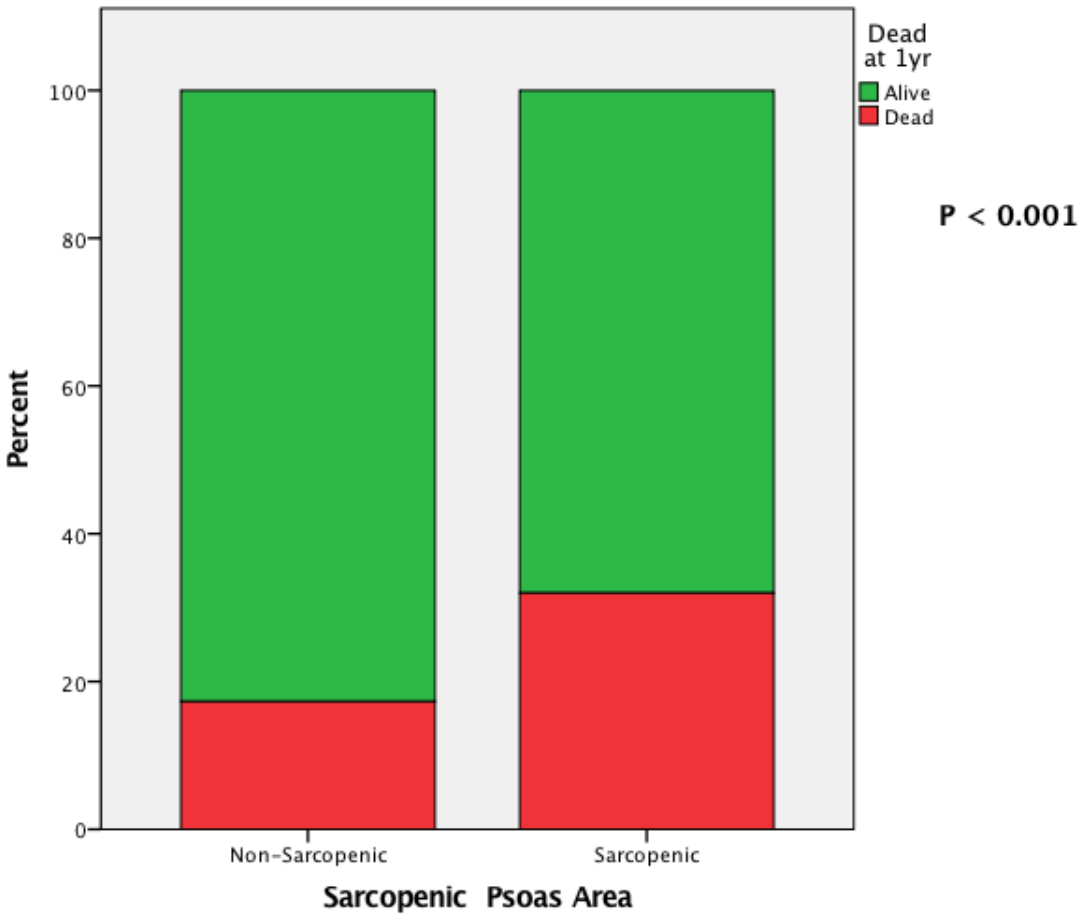


Figure 83: 1-Year mortality sarcopenia vs. non-sarcopenia PA

3.54 Multi Variant Analysis

As displayed in table 21, there is a significant difference in the rates malignancy between the sarcopenic and non-sarcopenic groups defined by psoas area.

30-day mortality

Binary logistic regression analysis was performed to assess the impact of sarcopenia on 30-day mortality adjusting for malignancy. The adjusted odds ratio for sarcopenia defined by psoas area is 2.25 with a 95% CI of 1.34 – 3.79, p = 0.002.

1-Year Mortality

Binary logistic regression analysis was performed to assess the impact of sarcopenia on 1-year mortality adjusting for predicted p-possum mortality and age. The adjusted odds ratio for sarcopenia defined by psoas area is 2.10 with a 95% CI of 1.39 – 3.17, p <0.001.

3.55 Survival Analysis

Kaplan-Meier survival analysis for the sarcopenic and non-sarcopenic groups as defined by psoas area displays a significant difference between the survival curve for the sarcopenic and non-sarcopenic groups, P < 0.001. Figure 84 displays the first 100 days where the survival curve for the sarcopenic group falls away rapidly from the non-sarcopenic group. Figure 85 displays the curve out to 1000 days of follow up where, although not as rapidly divergent, the survival curve for the sarcopenic group is still falling away from the non-sarcopenic group.

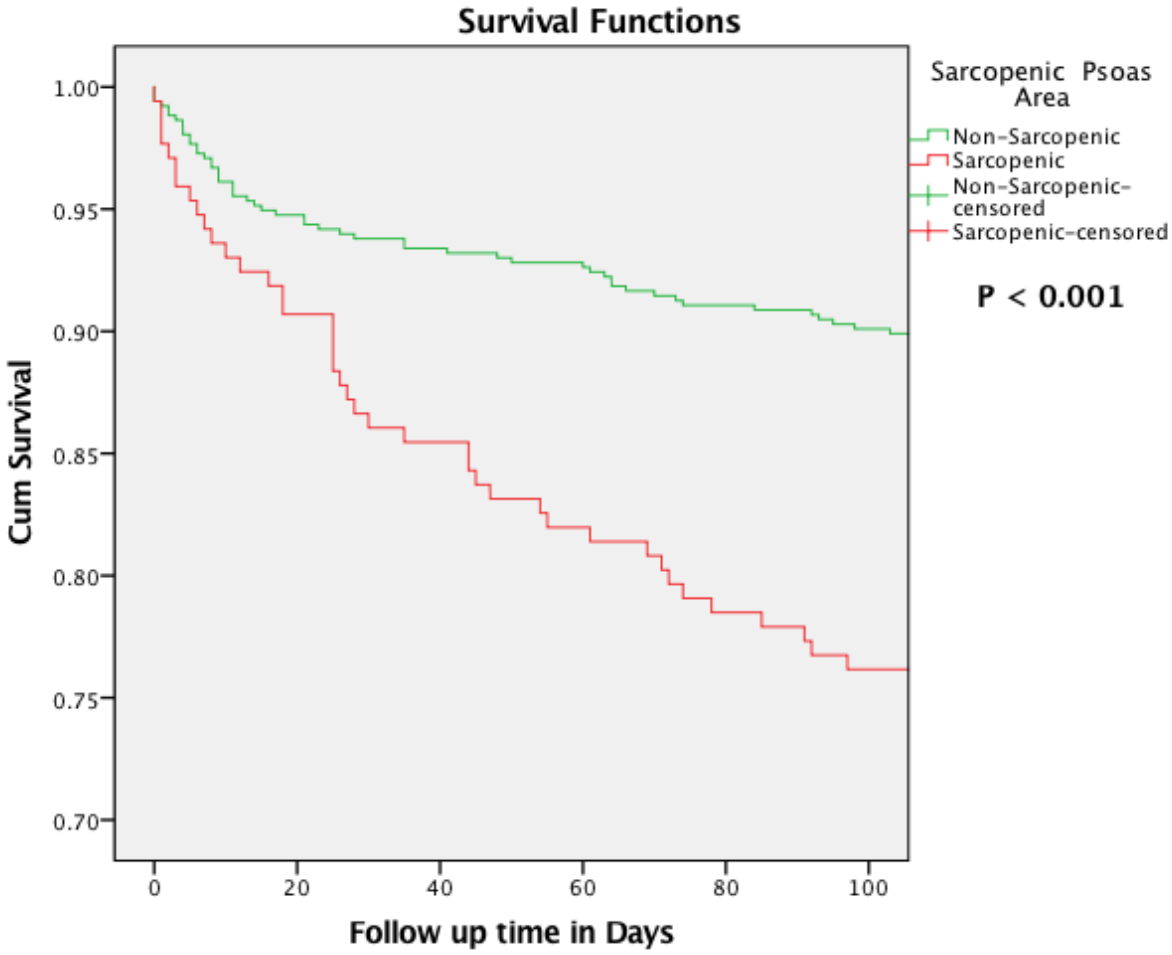


Figure 84: Kaplan-Meier survival 1st 100 days, PA

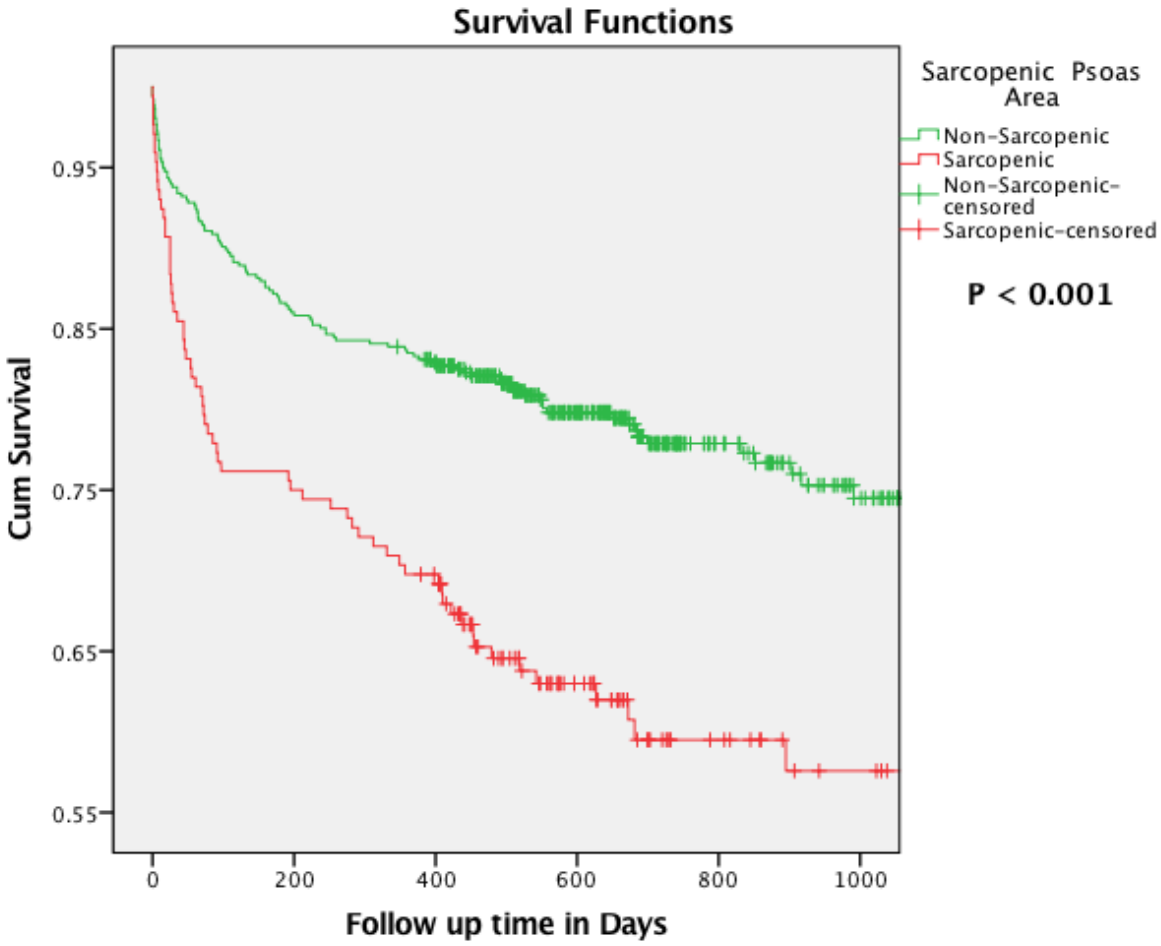


Figure 85: Kaplan-Meier survival 1000 days, PA

3.56 Receiver Operator Characteristic Curve PA

To ascertain the effectiveness PA as a screening tool to detect mortality within the follow-up period a ROC curve was calculated as displayed in figure 86. This shows an area under the graph of 0.639 with 95%CI 0f 0.593-0.686, p = <0.001. Ideal trade off between sensitivity and specificity was a PA of 4.17cm^2/m^2 for men and 3.17cm^2/m^2 for women. This would put the sarcopenic population in the 687 patients analysed at 289 or 42% (37.8% for women and 46% for men)

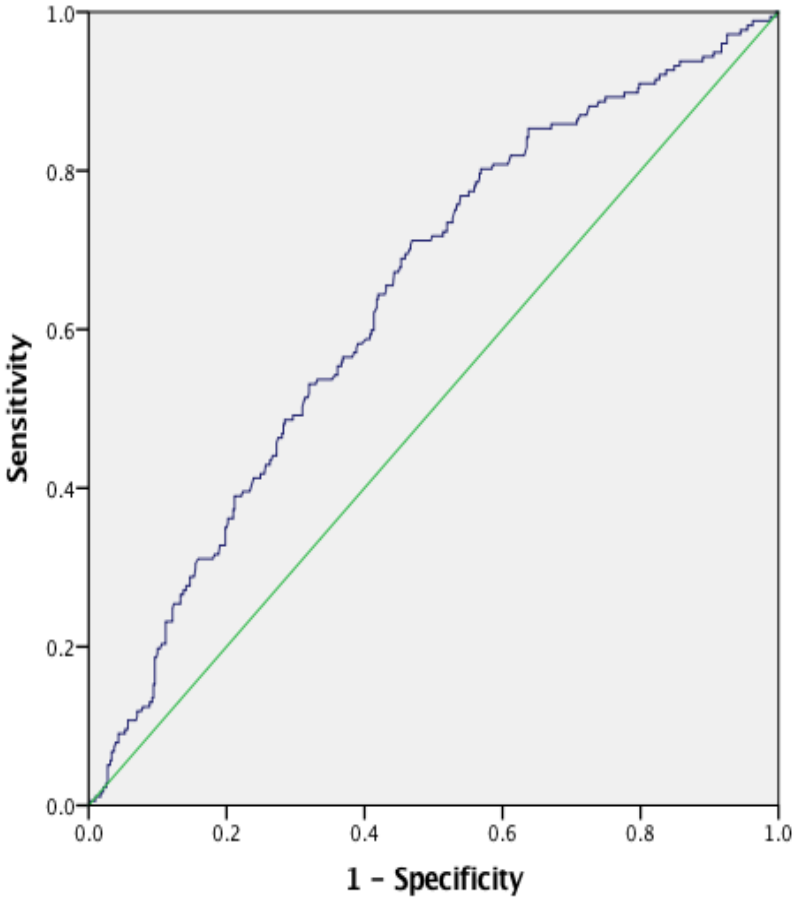


Figure 86: ROC curve sarcopenia PA

Table 20 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by ROC curve derived cut-off for psoas density, that were alive or dead at 30 days. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 30-day mortality with OR of 2.43 with 95% CI of 1.45-4.07, P<0.001

Table 21 Displays the percentage of sarcopenic and non-sarcopenic patients, defined by ROC curve derived cut-off for psoas density, that were alive or dead at 1-year. Data were analysed with binary logistic regression and a significant association was found between sarcopenia and 1-year mortality with OR of 2.31 with 95% CI of 1.58-3.36, P<0.001

	Sarcopenic	Non-Sarcopenic
Dead at 30 days	14.5% (42/289)	6.5% (26/398)
Alive at 30 days	85.5% (247/289)	93.5% (372/398)

Table 27: 30-day mortality ROC curve PA

	Sarcopenic	Non-Sarcopenic
Dead at 1-year	29.1% (84/289)	15.1% (60/398)
Alive at 30 days	70.7% (205/289)	84.9% (338/398)

Table 28: 1-year mortality ROC curve PA

Demographic	Sarcopenic	Non-Sarcopenic	P-value
Age (SD)	66.02 (16.02)	62.34 (18.02)	0.006
BMI (SD)	26.1 (6.3)	26.0 (5.9)	0.952
Malignant diagnosis	22.8% (66/289)	18.4% (62/336)	0.021
Sex M:F	152:137	175:223	0.031
P-Possum mortality (SD)	21.1 (20.6)	18.4 (24.5)	0.149

Table 29: Demographics ROC curve PA

3.57 Increased dependency

Table 24 displays the number and percentage of the sarcopenic and non-sarcopenic groups, defined by psoas area, that were admitted from independent living and discharged from the acute ward to non-independent living. Figure 87 displays this as a bar chart. Data were analysed with binary logistic regression and no significant association was found p = 0.74.

	Sarcopenic	Non-Sarcopenic
Remains in Independent Living	88.7% (55/62)	87.1% (162/186)
Home to Non-Independent Living	11.3% (7/62)	12.9% (24/186)

Table 30: Loss of independence PA

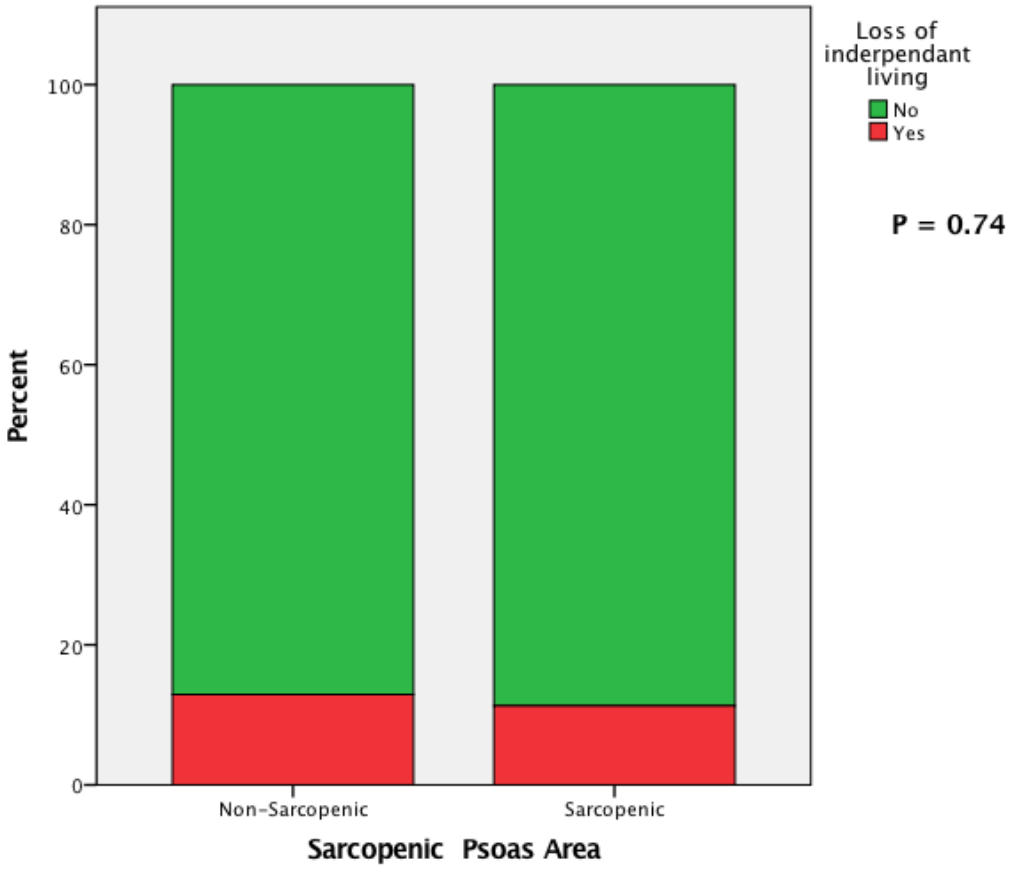


Figure 87: Loss of independent living PA

4. Data Validation from Multiple Sites

PD and PA calculation from the Derby and Bradford groups was collected by a number of different individuals. Once data collection was complete data validation was needed to ensure psoas area and density calculation were accurate from these sites.

No issue was found with Bradford data. Once received the data set for Derby underwent initial analysis that revealed no significant difference in mortality at 30-days or 1-year between sarcopenic and non-sarcopenic groups using either PD or PA (table 25). This either represented an issue with the Derby data or that the difference seen in Scarborough, York and Bradford is not seen elsewhere.

A visit to Derby hospital was made to validate the data and calculation of PD and PA. A sample of 50 patents was selected from the Derby dataset and their PD and PA values calculated and compared to those calculated by the Derby team. There was a large difference between values calculated by the Derby team and those from the trained validator. Error (difference between Derby team and validator) was calculated for each of the sample 50 patients for PD and PA. This error in either Hounsfield units (PD) or cm^2 (PA) was then calculated as a percentage of the validator's measurement to give a percentage error for each measurement by the Derby team. The Median error rate from the Derby team was 36.8% (IQR 53.9) for PD and 28.2% (IQR 33.3%) for PA. This large error may have accounted for the lack of a significant difference for increased mortality associated with sarcopenia in the Derby data that was seen elsewhere.

The Derby data had been combined from three separate sources: a NELA database download; an output from the hospital data system; an excel spreadsheet where PA and PD values were entered. These

had then been combined into a final data set. On review of these individual source data sheets a copy error was revealed. The first patients PD and PA measurements had not been merged to the final data sheet and as such all PA and PD values were shifted to the patient above. Once this was corrected the percentage error of the Derby PD and PA values in reference to the validators was recalculated. Median error was now 6.6% (IQR 15.3%) and 5% (IQR 9.3%) respectively. A median error of 6.6% and 5% would be clinically acceptable. The entire data-set from derby was re-analysed with a Chi-Squared test looking for the effect of sarcopenia calculated by PD and PA on 30-day and 1-year mortality. The statistically significant trend seen in all other sites was then found in the recalculated data (Table 26). The Derby data were then combined with that of Scarborough York and Bradford for further analysis. Figure 88 displays the Median % error between the Derby and Validators PA and PD values before and after the copy error was corrected.

	30-day Mortality	1-Year Mortality
PD	0.27	0.59
PA	0.84	0.73

Table 31: P-values from Chi-Squared test for non-validated raw data from Derby

	30-Day mortality	1-Year Mortality
PD	0.007	0.04
PA	0.011	0.003

Table 32: P-values from Chi-Squared test for validated raw data from Derby

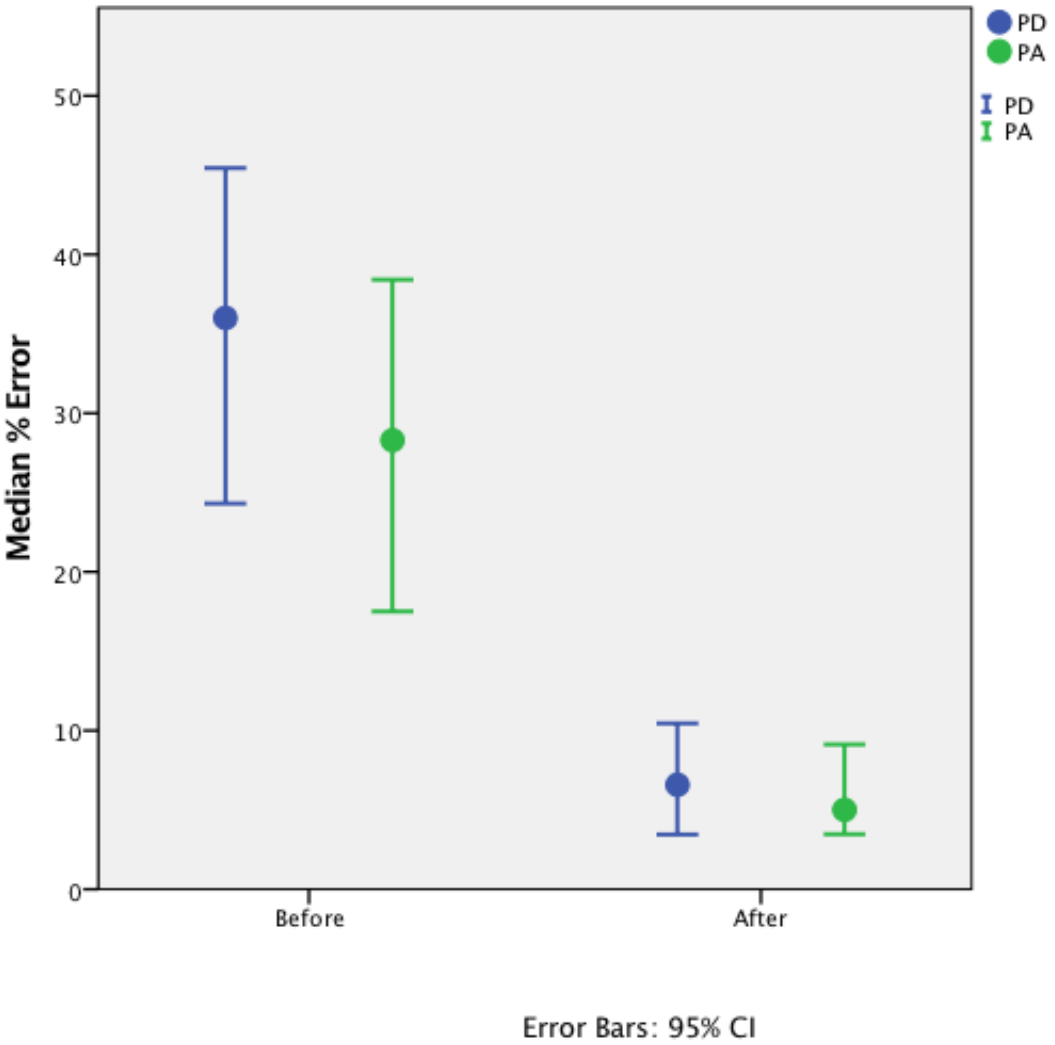


Figure 88: Median error before and after data validation

5. Discussion

5.1 Synopsis of Findings

Both PD and PA sarcopenic groups have significantly worse mortality at 30 days and at 1 year. Following multi variant analysis odds ratios for 30-day mortality for were 1.83 and 2.25 for PD and PA respectively. Odds ratios for 1-year mortality were 1.58 and 2.10. Analysis of the sarcopenic group defined by PA revealed a significant increase in post laparotomy dependency OR of 2.50.

5.2 Discussion of Methods and Limitations

The main limitation of this study is that uses retrospective data. We have attempted to combat this by only picking “hard” endpoints for analysis that are less prone to error such as mortality. This however has meant that we have not been able to assess endpoints such as complication rate that are widely reported in other studies.

Cases were identified through the NELA database where collection is good but cases are missed. In 2015-16 83% of the approximately 30,000 laparotomy cases performed in England and Wales were collected (NELA, 2017).

The cut-off we used for sarcopenia was the lowest quartile of PD or PA. This could be considered arbitrary. However, as there are no published cut-off values for patients from the United Kingdom undergoing emergency laparotomy this was felt to be wiser. Whilst sex-specific cut-off values have now been identified by both the European Working Group on Sarcopenia in Older People (EWGSOP)

and the Asian working group (Chen et al., 2014) for performance base measurement of sarcopenia, the same is not true for muscle mass or quality measurement. The latest publication from the European Working Group on Sarcopenia in Older People (EWGSOP) (Cruz-Jentoft, 2019) commented that ‘cut-off points for low muscle mass are not yet well defined’. It has been demonstrated (Zhong et al., 2012) that there are significant differences of muscle mass in healthy individuals between ethnic groups. It would be reasonable therefore to express caution when applying cut-off values from one study to another and this may explain the huge variance in sarcopenia prevalence in some studies.

ROC curve analysis for both PA and PD produced a cut-off value that included a significantly greater proportion of the observed cohort in the sarcopenic group. With the new cut-off values 42% of patients were sarcopenic in the PA analysis and 50% in the PD analysis. Despite this large increase in the proportion of the group deemed sarcopenic, differences in survival were preserved with very similar odds ratios. However other changes occurred when the cut-offs were moved. For PA derived sarcopenia a significant difference in the age between groups was now seen. For PD this is also the case with the sarcopenic group now having more malignancy and a higher BMI where there not before. The differences previously seen in predicted P-Possum mortality and age were still present but larger with the ROC curve cut-off. It would seem that despite maintaining the predictive value of PA and PD for mortality moving the cut-offs make it less independent of patient demographics. This makes it less useful as an addition to tests and information that are already used in clinical practice. Using the bottom quartile as the cut-off for PA and PD analysis produced a more clinically useful test.

5.3 Demographics

Differences between sites

There were 4 separate acute sites involved in data collection. Although these are all UK acute trusts they serve different populations. As has been highlighted in the published literature (Wigodski et al., 2018, Yoowannakul et al., 2018b, Yoowannakul et al., 2018a) generic cut-offs for sarcopenia measurement (Cruz-Jentoft et al., 2010, Fielding et al., 2011) may not be relevant to all populations. Although this may not apply down to the level of individual towns within a single country, differences within our data should be considered. Unfortunately no data on ethnicity was available from hospital IT systems. On analysis of other relevant demographics only age was found to be different between sites. The maximum difference in mean age was 7.8 years and occurred between Derby and Scarborough sites. A mean age difference of 7.8 years reflects the population and is likely to make considerable clinical difference in outcome following emergency laparotomy. The bottom quartile of PD and PA measurements was used as a cut-off for sarcopenia in this study. This was calculated with data from all sites together. The difference in mean age between sites makes one consider if cut-offs should be calculated for each individual site. However given that a significant proportion of the published literature is content with using values from single nationality studies on an international basis it seems reasonable to use one cut off.

PA versus PD

Key demographics of the analysed population for PD and PA analysis were investigated for difference between the sarcopenic and

non-sarcopenic groups. There were significant differences found but there were not the same for PD and PA. The sarcopenic group defined by PD analysis were significantly older (10.8 years) and sicker (11.3% higher P-POSSUM,) than the non-sarcopenic group. The sarcopenic group as defined by PA analysis did not display this with no difference found in age or P-POSSUM score. The Sarcopenic group as defined by PA had significantly higher (7.7%) malignancy rates. This difference between groups and analysis methods is interesting and changes how we should interpret results. It may be that PA is a more useful marker of frailty than PD. PD appears to be a function of age and sickness as defined by P-POSSUM. However, this needs to be interpreted with caution as there is a degree of co-linearity as age is used in P-POSSUM predicted mortality, alongside markers of physiological compromise. Although PA defined sarcopenic patients are more likely to carry a malignant diagnosis it being independent of age and the physiological instability that makes up the majority of P-POSSUM scoring may make it more of a useful adjunct to current practice.

The vast majority of the CT scans used for this study were done with intravenous contrast and analysed in the portal-venous phase. There is evidence (Boutin at al. 2016) of a significant difference in psoas density measurement based on the phase of an intravenous contrast enhanced CT. This was most marked between non-contrast and delayed phase (post portal-venous phase), where a difference of 28% was observed. Furthermore the age of the patient had a significant effect on the amount of density change seen between phases of contrast enhanced scans. Our finding that PD is related to physiological compromise (P-Possum predicted mortality) and age may be explained by this. In a physiologically compromised state circulation of intravenous contrast may be hindered thus altering the observed psoas density on a post contrast scan. This further confirms that in an acute situation PA may be of more use than PD in sarcopenia assessment.

5.4 Mortality

Both PD and PA sarcopenic groups have significantly worse mortality at 30 days and at 1 year. Following multi variant analysis odds ratios for 30-day mortality for were 1.83 (95% CI of 1.16 – 2.87, P = 0.008) and 2.25 (95% CI of 1.34 – 3.79, p = 0.002) for PD and PA respectively. Odds ratios for 1-year mortality were 1.58 (95% CI of 1.09 – 2.29, P = 0.016) and 2.10 (95% CI of 1.39 – 3.17, p <0.001) for PD and PA respectively.

Survival curves for both PA and PD analysis show the cumulative survival of sarcopenic groups falling away rapidly for the first 100 days from that of the non-sarcopenic groups. This rapid fall slows but even at 1000 days, when the last patients have been censored or have died, the curves are still divergent. This analysis confirms that the same trend for poor outcome found in the vast majority of published literature on elective surgery is also found in the emergency cohort. PA defined sarcopenia appears to be the stronger predictor of mortality.

5.5 Increased Dependency

Analysis of the sarcopenic group defined by PA revealed an increase in post laparotomy dependency. The sarcopenic group had an OR of 2.50 (95% CI of 1.15-5.46, P = 0.021) of being discharged to dependant living if admitted from independent living. There was no significant trend found on PD analysis. This is despite having 157 (19%) less numbers in the PA group. This is inline with mortality analysis where PA sarcopenia was a better predictor of mortality than PD.

5.6 Inter and Intra-Operator Variability

Inter-operator variability was 9.83% for all grades, and 6.38% when analysis was confined to operators of consultant and registrar grade only. Mean intra-operator variability was 3% for PA and 3.9% for PD. It is likely that if sarcopenia analysis were to be used in routine practice for emergency laparotomies a more senior member of the surgical team, if not the reporting radiologist, would calculate it. These values are analogous to error rates for radiologists reporting CTs with a large meta-analysis (Wu et al., 2013) finding a pooled error rate of 7.7% (95% CI 5.6-10.3%) across 58 studies and 388,000 CT examinations. Given the wide use and trust placed in CT reports in wider clinical practice a finding of a variability of 6.38% for these grades with an intra operator variably of 3% is likely to be acceptable variance for more widespread use. In addition to this, those chosen to do this validation work had received training but not undertaken any PD or PA calculation previously. This means that if a learning curve exists for calculation this group of assessors would be at the start of that curve. As such the variability quoted here represents a worst-case scenario.

5.7 Overall Discussion of findings

Findings from this multicentre study broadly fall in-line with those we have previously published from a single centre (Trotter et al., 2018). The results suggest that CT defined sarcopenia is associated with an increased 30-day and one year mortality rate following emergency laparotomy. Further, sarcopenic patients frequently fail to be discharged back to independent living following their acute stay.

Pre-operative assessment of patients scheduled for elective surgery is now standard practice. This is driven to some extent by the publication of surgeons' outcome data (NBOCA, 2017). Patients are now subjected to intensive investigation and assessment in the form of Cardio-Pulmonary Exercise testing (CPET), echography and Consultant anaesthetic evaluation. In some units (Partridge et al., 2017), elderly patients undergo a formal CGA as there is increasing recognition that frailty syndrome is associated with a poor outcome, but may be manipulated positively through the CGA process. These assessment techniques, however, are often time consuming and may be inappropriate in the emergency situation. Currently clinicians rely on experience and risk scoring formulae such as P-POSSUM to aid decision-making. P-POSSUM, and similar prognostic indices based on physiological parameters have been validated (Copeland et al., 1991). However they suffer the common drawback that results are very variable depending on the state of physiological compromise of the patient and the efficacy of resuscitation received.

There are many techniques described in the literature to measure sarcopenia as outlined in the introduction. PA and PD analysis has the advantage that it does not require specialist radiological software and is rapid to perform. With inter and intra operator variability rates from minimally trained staff falling in clinically acceptable levels, PA

and PD measurement may be useful in real-time clinical practice. Although there is no accepted definition of sarcopenia the consensus statements (Fielding et al., 2011, Morley et al., 2011, Chen et al., 2014) suggest that it is best defined with a both a measure of muscle mass and function such as gait speed. Again in an elective or research setting this may be ideal. It is, however, improbable that a measure of muscle function such as gait speed or hand-grip strength can be relied upon in a clinically unstable patient that requires an emergency laparotomy.

PA seems the most useful marker of frailty from this study. Its apparent independence from age and the clinical markers of instability contained in P-POSSUM scoring could make it a useful adjunct to clinical care. The next question is therefore how this measure of frailty would then be used to alter care. A 1-year mortality rate of 32% is unlikely to change a decision for operative intervention. However, in the new post Montgomery world of informed consent (Smith et al., 2002), this information may help patients, families and surgeons weigh up the risk/benefit ratio of emergency laparotomy.

One of the key NELA recommendations is that patients over 70 years of age would have routine input from care of the elderly physicians. The reality is that only 10% of patients receive this care (NELA, 2017). If it is not possible to provide this service to all patients who undergo an emergency laparotomy, frailty assessment by CT sarcopenia may help target this assessment to those who might potentially benefit the most. These patients may then be able to access some of the benefit that target multidisciplinary interventions such as CGA have been proven to deliver (Ellis et al., 2017, Partridge et al., 2017). Since this study has finished data collection NELA has begun to capture some data on frailty in the form of a clinical frailty scale.

5.8 Conclusion

Sarcopenia defined by single slice measurement of psoas area and density is associated with a significant rise in 30- mortality, 1-year mortality and loss of independence. Of the two measurement techniques psoas area is the more predictive of these poor outcomes, and more independent of other confounding variables. The technique of psoas area measurement appears to be simple and reliable when performed by members of the surgical team without the need for extra resource or extensive training. The difference in outcome between sarcopenic and non-sarcopenic groups is not likely enough to turn a patient down for an emergency laparotomy but it may be a useful adjunct to clinical decision-making. In particular it can highlight patients who may benefit most from specialist input such as intensive care or formal frailty assessment with CGA in the postoperative phase.

Below, I have listed the initial hypotheses and whether or not the null hypothesis has been rejected from using the results of this study:

Primary Hypothesis

Frailty defined by sarcopenia on CT scanning, interpedently worsens mortality following Emergency Laparotomy

- Hypothesis is accepted, null hypothesis rejected

Secondary Hypothesis

Frailty defined by sarcopenia on CT scanning, interpedently worsens social dependence following Emergency Laparotomy

- Hypothesis is accepted, null hypothesis rejected

5.9 Further Work

The results from this retrospective study of 845 patients in 4 centres have shown some interesting results. The logical next step would be to test these results in a prospective study. This need not be confined to laparotomy patents but could be expanded to all surgical admissions across many specialties. Such a study would allow collection of other outcome makers such as complication rate. If results were similar to this study, they would serve of further evidence for routine frailty screening and targeted multidisciplinary specialist frailty input for this group.

Thesis Conclusion

This thesis intended to investigate a method of identification of frailty in the surgical population, CT defined sarcopenia, and a possible method to attenuate its effects in the preoperative period, prehabilitation.

It appears both acceptable and achievable for patients listed for resectional bowel surgery to undertake a walking based exercise programme. The use of simple wearable activity-monitoring technology is acceptable to patients and even within the confinement of a strict cancer target pathway fitness measures can be improved.

Objective CT sarcopenia measurement is associated with a significant rise in 30- mortality, 1-year mortality and loss of independence. The technique used here of psoas area is reliable when performed by senior members of the surgical team without the need for extra resource or extensive training.

In a result driven modern healthcare system where more complex surgical procedures are being performed, on an aging population, the identification of those who are at risk of poor outcomes is key. The decision on fitness for surgery is multifaceted and not simply reliant on age. When time is short in emergency care simple tools like CT sarcopenia assessment may be of use to aid decision making for patients, family and clinicians. In the elective setting where time is short the application of preoperative strategies such as exercise could lead to the amelioration of some of the effects of frailty by increasing the physiological robustness of patients. One of the key elements for success of such programmes is engagement of those who require it and here we have demonstrated a programme that

meets that requirement. Further work with a programme such as this may allow its modification to become part of preoperative care, most likely for those who are frail but still possibly fit for surgery.

References

AARONSON, N. K., AHMEDZAI, S., BERGMAN, B., BULLINGER, M., CULL, A., DUEZ, N. J., FILIBERTI, A., FLECHTNER, H., FLEISHMAN, S. B., DEHAES, J., KAASA, S., KLEE, M., OSOBA, D., RAZAVI, D., ROFE, P. B., SCHRAUB, S., SNEEUW, K., SULLIVAN, M. & TAKEDA, F. 1993. THE EUROPEAN-ORGANIZATION-FOR-RESEARCH-AND-TREATMENT-OF-CANCER QLQ-C30 - A QUALITY-OF-LIFE INSTRUMENT FOR USE IN INTERNATIONAL CLINICAL-TRIALS IN ONCOLOGY. *Journal of the National Cancer Institute,* 85**,** 365-376.

ANDERSON, A. D. G., MCNAUGHT, C. E., MACFIE, J., TRING, I., BARKER, P. & MITCHELL, C. J. 2003. Randomized clinical trial of multimodal optimization and standard perioperative surgical care. *British Journal of Surgery,* 90**,** 1497-1504.

ANTONESCU, I., SCOTT, S., TRAN, T. T., MAYO, N. E. & FELDMAN, L. S. 2014. Measuring postoperative recovery: What are clinically meaningful differences? *Surgery,* 156**,** 319-327.

BALADY, G. J., ARENA, R., SIETSEMA, K., MYERS, J., COKE, L., FLETCHER, G. F., FORMAN, D., FRANKLIN, B., GUAZZI, M., GULATI, M., KETEYIAN, S. J., LAVIE, C. J., MACKO, R., MANCINI, D., MILANI, R. V., AMER HEART ASSOC EXERCISE, C., COUNCIL EPIDEMIOLOGY, P., COUNCIL PERIPHERAL VASC, D. & INTERDISCIPLINARY COUNCIL, Q. 2010. Clinician's Guide to Cardiopulmonary Exercise Testing in Adults A Scientific Statement From the American Heart Association. *Circulation,* 122**,** 191-225.

BANDEEN-ROCHE, K., XUE, Q. L., FERRUCCI, L., WALSTON, J., GURALNIK, J. M., CHAVES, P., ZEGER, S. L. & FRIED, L. P. 2006. Phenotype of frailty: Characterization in the women's health and aging studies. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences,* 61**,** 262-266.

BARAKAT, H. M., SHAHIN, Y., KHAN, J. A., MCCOLLUM, P. T. & CHETTER, I. C. 2016. Preoperative Supervised Exercise Improves Outcomes After Elective Abdominal Aortic Aneurysm Repair A Randomized Controlled Trial. *Annals of Surgery,* 264**,** 47-53.

BARBALHO-MOULIM, M. C., SOARES MIGUEL, G. P., PAZZIANOTTO FORTI, E. M., CAMPOS, F. D. A. & COSTA, D. 2011. Effects of preoperative inspiratory muscle training in obese women undergoing open bariatric surgery: respiratory muscle strength, lung volumes, and diaphragmatic excursion. *Clinics,* 66**,** 1721-1727.

BARBERAN-GARCIA, A., UBRE, M., ROCA, J., LACY, A. M., BURGOS, F., RISCO, R., MOMBLAN, D., BALUST, J.,

BLANCO, I. & MARTINEZ-PALLI, G. 2018. Personalised Prehabilitation in High-risk Patients Undergoing Elective Major Abdominal Surgery: A Randomized Blinded Controlled Trial. *Annals of Surgery,* 267**,** 50-56.
BARDRAM, L., FUNCHJENSEN, P., JENSEN, P., CRAWFORD, M. E. & KEHLET, H. 1995. RECOVERY AFTER LAPAROSCOPIC COLONIC SURGERY WITH EPIDURAL ANALGESIA, AND EARLY ORAL NUTRITION AND MOBILIZATION. *Lancet,* 345**,** 763-764.
BARLOW, R. C., CHAN, D. S. Y., MAYOR, S., PERKINS, C., LAWTON, H. L., POWELL, A. G. & LEWIS, W. G. 2018. Fit for Cancer Treatment: a prospective feasibility study of primary care initiated prehabilitation for patients with suspected cancer. *BJGP open,* 2**,** bjgpopen18X101608-bjgpopen18X101608.
BEAUDART, C., MCCLOSKEY, E., BRUYERE, O., CESARI, M., ROLLAND, Y., RIZZOLI, R., DE CARVALHO, I. A., THIYAGARAJAN, J. A., BAUTMANS, I., BERTIERE, M. C., BRANDI, M. L., AL-DAGHRI, N. M., BURLET, N., CAVALIER, E., CERRETA, F., CHERUBINI, A., FIELDING, R., GIELEN, E., LANDI, F., PETERMANS, J., REGINSTER, J. Y., VISSER, M., KANIS, J. & COOPER, C. 2016. Sarcopenia in daily practice: assessment and management. *Bmc Geriatrics,* 16.
BITTNER, V., WEINER, D. H., YUSUF, S., ROGERS, W. J., MCINTYRE, K. M., BANGDIWALA, S. I., KRONENBERG, M. W., KOSTIS, J. B., KOHN, R. M., GUILLOTTE, M., GREENBERG, B., WOODS, P. A. & BOURASSA, M. G. 1993. PREDICTION OF MORTALITY AND MORBIDITY WITH A 6-MINUTE WALK TEST IN PATIENTS WITH LEFT-VENTRICULAR DYSFUNCTION. *Jama-Journal of the American Medical Association,* 270**,** 1702-1707.
BORG, G. A. V. 1982. PSYCHOPHYSICAL BASES OF PERCEIVED EXERTION. *Medicine and Science in Sports and Exercise,* 14**,** 377-381.
BRAMLEY, P. & BROWN, J. 2018. Cause for scepticism regarding preoperative cardiopulmonary exercise testing. *British Journal of Anaesthesia,* 121**,** 496-497.
BRUBAKER, P. H. & KITZMAN, D. W. 2011. Chronotropic Incompetence Causes, Consequences, and Management. *Circulation,* 123**,** 1010-1020.
BUETTNER, S., WAGNER, D., KIM, Y., MARGONIS, G. A., MAKARY, M. A., WILSON, A., SASAKI, K., AMINI, N., GANI, F. & PAWLIK, T. M. 2016. 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,* 222**,** 397-407.
BUTLAND, R. J. A., PANG, J., GROSS, E. R., WOODCOCK, A. A. & GEDDES, D. M. 1982. 2-MINUTE, 6-MINUTE, AND 12-

MINUTE WALKING TESTS IN RESPIRATORY-DISEASE. *British Medical Journal,* 284**,** 1607-1608.
BOUTIN, R., KAPTUCH, J. M. 2016. Infulecne of IV contrast on CT Measures of Muscle and Bone Attenuation: Implications for Sarcopenia and Osteoporosis Evalusation. *American Journal of Roentgenology*, 207, 1046-1054
CADMUS-BERTRAM, L. A., MARCUS, B. H., PATTERSON, R. E., PARKER, B. A. & MOREY, B. L. 2015. Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women. *American Journal of Preventive Medicine,* 49**,** 414-418.
CAHALIN, L., PAPPAGIANOPOULOS, P., PREVOST, S., WAIN, J. & GINNS, L. 1995. THE RELATIONSHIP OF THE 6-MIN WALK TEST TO MAXIMAL OXYGEN-CONSUMPTION IN TRANSPLANT CANDIDATES WITH END-STAGE LUNG-DISEASE. *Chest,* 108**,** 452-459.
CARLI, F., CHARLEBOIS, P., STEIN, B., FELDMAN, L., ZAVORSKY, G., KIM, D. J., SCOTT, S. & MAYO, N. E. 2010. Randomized clinical trial of prehabilitation in colorectal surgery. *British Journal of Surgery,* 97**,** 1187-1197.
CARVALHO, V. O., BOCCHI, E. A. & GUIMARES, G. V. 2009. The Borg Scale as an Important Tool of Self-Monitoring and Self-Regulation of Exercise Prescription in Heart Failure Patients During Hydrotherapy - A Randomized Blinded Controlled Trial. *Circulation Journal,* 73**,** 1871-1876.
CESARI, M., LEEUWENBURGH, C., LAURETANI, F., ONDER, G., BANDINELLI, S., MARALDI, C., GURALNIK, J. M., PAHOR, M. & FERRUCCI, L. 2006. Frailty syndrome and skeletal muscle: results from the Invecchiare in Chianti study. *American Journal of Clinical Nutrition,* 83**,** 1142-1148.
CHAMARI 2015. 'Aerobic' and 'Anaerobic' terms used in exercise physiology: a critical terminology reflection. *In:* PADULO (ed.). Sports Medicine Open.
CHEN, L. K., LIU, L. K., WOO, J., ASSANTACHAI, P., AUYEUNG, T. W., BAHYAH, K. S., CHOU, M. Y., CHEN, L. Y., HSU, P. S., KRAIRIT, O., LEE, J. S. W., LEE, W. J., LEE, Y., LIANG, C. K., LIMPAWATTANA, P., LIN, C. S., PENG, L. N., SATAKE, S., SUZUKI, T., WON, C. W., WU, C. H., WU, S. N., ZHANG, T. M., ZENG, P., AKISHITA, M. & ARAI, H. 2014. Sarcopenia in Asia: Consensus Report of the Asian Working Group for Sarcopenia. *Journal of the American Medical Directors Association,* 15**,** 95-101.
CHO, H., YOSHIKAWA, T., OBA, M. S., HIRABAYASHI, N., SHIRAI, J., AOYAMA, T., HAYASHI, T., YAMADA, T., OBA, K., MORITA, S., SAKAMOTO, J. & TSUBURAYA, A. 2014. Matched Pair Analysis to Examine the Effects of a Planned Preoperative Exercise Program in Early Gastric Cancer Patients with Metabolic Syndrome to Reduce Operative Risk: The Adjuvant Exercise for General Elective Surgery (AEGES) Study Group. *Annals of Surgical Oncology,* 21**,** 2044-2050.

CLEGG, A., YOUNG, J., ILIFFE, S., RIKKERT, M. O. & ROCKWOOD, K. 2013. Frailty in elderly people. *Lancet,* 381**,** 752-762.
COLSON, M., BAGLIN, J., BOLSIN, S. & GROCOTT, M. P. W. 2012. Cardiopulmonary exercise testing predicts 5 yr survival after major surgery. *British Journal of Anaesthesia,* 109**,** 735-741.
COOPER, C., DERE, W., EVANS, W., KANIS, J. A., RIZZOLI, R., SAYER, A. A., SIEBER, C. C., KAUFMAN, J. M., VAN KAN, G. A., BOONEN, S., ADACHI, J., MITLAK, B., TSOUDEROS, Y., ROLLAND, Y. & REGINSTER, J. Y. L. 2012. Frailty and sarcopenia: definitions and outcome parameters. *Osteoporosis International,* 23**,** 1839-1848.
COPELAND, G. P., JONES, D. & WALTERS, M. 1991. POSSUM - A SCORING SYSTEM FOR SURGICAL AUDIT. *British Journal of Surgery,* 78**,** 356-360.
COURNEYA, K. S., FRIEDENREICH, C. M., QUINNEY, H. A., FIELDS, A. L. A., JONES, L. W. & FAIREY, A. S. 2003a. A randomized trial of exercise and quality of life in colorectal cancer survivors. *European Journal of Cancer Care,* 12**,** 347-357.
COURNEYA, K. S., MACKEY, J. R., BELL, G. J., JONES, L. W., FIELD, C. J. & FAIREY, A. S. 2003b. Randomized controlled trial of exercise training in postmenopausal breast cancer survivors: Cardiopulmonary and quality of life outcomes. *Journal of Clinical Oncology,* 21**,** 1660-1668.
CRAPO, R. O., CASABURI, R., COATES, A. L., ENRIGHT, P. L., MACINTYRE, N. R., MCKAY, R. T., JOHNSON, D., WANGER, J. S., ZEBALLOS, R. J., BITTNER, V., MOTTRAM, C. & COMM, A. T. S. 2002. ATS statement: Guidelines for the six-minute walk test. *American Journal of Respiratory and Critical Care Medicine,* 166**,** 111-117.
CRAWFORD, J. R., HENRY, J. D., CROMBIE, C. & TAYLOR, E. P. 2001. Normative data for the HADS from a large non-clinical sample. *British Journal of Clinical Psychology,* 40**,** 429-434.
CRUZ-JENTOFT 2019. Sarcopenia: revised European consensus on definition and diagnosis. Age and Ageing.
CRUZ-JENTOFT, A. J., BAEYENS, J. P., BAUER, J. M., BOIRIE, Y., CEDERHOLM, T., LANDI, F., MARTIN, F. C., MICHEL, J. P., ROLLAND, Y., SCHNEIDER, S. M., TOPINKOVA, E., VANDEWOUDE, M. & ZAMBONI, M. 2010. Sarcopenia: European consensus on definition and diagnosis. *Age and Ageing,* 39**,** 412-423.
DE TOLEDO PIZA SOARES, S. M., NUCCI, L. B., DE CARVALHO DA SILVA, M. M. & CAMPACCI, T. C. 2013. Pulmonary function and physical performance outcomes with preoperative physical therapy in upper abdominal surgery: a randomized controlled trial. *Clinical Rehabilitation,* 27**,** 616-627.
DETTLING, D. S., VAN DER SCHAAF, M., BLOM, R. L. G. M., NOLLET, F., BUSCH, O. R. C. & VAN BERGE

HENEGOUWEN, M. I. 2013. Feasibility and effectiveness of pre-operative inspiratory muscle training in patients undergoing oesophagectomy: a pilot study. *Physiotherapy research international : the journal for researchers and clinicians in physical therapy,* 18**,** 16-26.
DEVLIN, N. J. & KRABBE, P. F. M. 2013. The development of new research methods for the valuation of EQ-5D-5L. *European Journal of Health Economics,* 14**,** S1-S3.
DIAZ, K. M., KRUPKA, D. J., CHANG, M. J., PEACOCK, J., MA, Y., GOLDSMITH, J., SCHWARTZ, J. E. & DAVIDSON, K. W. 2015. Fitbit (R): An accurate and reliable device for wireless physical activity tracking. *International Journal of Cardiology,* 185**,** 138-140.
DRONKERS, J., VELDMAN, A., HOBERG, E., VAN DER WAAL, C. & VAN MEETEREN, N. 2008. Prevention of pulmonary complications after upper abdominal surgery by preoperative intensive inspiratory muscle training: a randomized controlled pilot study. *Clinical Rehabilitation,* 22**,** 134-142.
DRONKERS, J. J., LAMBERTS, H., REUTELINGSPERGER, I. M. M. D., NABER, R. H., DRONKERS-LANDMAN, C. M., VELDMAN, A. & VAN MEETEREN, N. L. U. 2010. Preoperative therapeutic programme for elderly patients scheduled for elective abdominal oncological surgery: a randomized controlled pilot study. *Clinical Rehabilitation,* 24**,** 614-622.
DU, Y., KARVELLAS, C. J., BARACOS, V., WILLIAMS, D. C., KHADAROO, R. G. & ACUTE CARE EMERGENCY SURG, A. G. 2014. Sarcopenia is a predictor of outcomes in very elderly patients undergoing emergency surgery. *Surgery,* 156**,** 521-527.
DUMBRILL, J. L., CONWAY, D. H., ATKINSON, D., BRYAN, A. & POWELL, R. 2019. The International Physical Activity Questionnaire and its relationship to performance at cardiopulmonary exercise testing in preoperative patients. *British Journal of Anaesthesia,* 122**,** E43-E43.
DUNNE, D. F. J., JACK, S., JONES, R. P., JONES, L., LYTHGOE, D. T., MALIK, H. Z., POSTON, G. J., PALMER, D. H. & FENWICK, S. W. 2016. Randomized clinical trial of prehabilitation before planned liver resection. *British Journal of Surgery,* 103**,** 504-512.
ELLIS, G., GARDNER, M., TSIACHRISTAS, A., LANGHORNE, P., BURKE, O., HARWOOD, R. H., CONROY, S. P., KIRCHER, T., SOMME, D., SALTVEDT, I., WALD, H., O'NEILL, D., ROBINSON, D. & SHEPPERD, S. 2017. Comprehensive geriatric assessment for older adults admitted to hospital. *Cochrane Database of Systematic Reviews*.
EVENSON, K. R., WEN, F. & FURBERG, R. D. 2017. Assessing Validity of the Fitbit Indicators for US Public Health Surveillance. *American Journal of Preventive Medicine,* 53**,** 931-932.

FEARON, K. C. H., LJUNGQVIST, O., VON MEYENFELDT, M., REVHAUG, A., DEJONG, C. H. C., LASSEN, K., NYGREN, J., HAUSEL, J., SOOP, M., ANDERSEN, J. & KEHLET, H. 2005. Enhanced recovery after surgery: A consensus review of clinical care for patients undergoing colonic resection. *Clinical Nutrition,* 24**,** 466-477.

FERREIRA, V., AGNIHOTRAM, R. V., BERGDAHL, A., VAN ROOIJEN, S. J., AWASTHI, R., CARLI, F. & SCHEEDE-BERGDAHL, C. 2018. Maximizing patient adherence to prehabilitation: what do the patients say? *Supportive Care in Cancer,* 26**,** 2717-2723.

FIELDING, R. A., VELLAS, B., EVANS, W. J., BHASIN, S., MORLEY, J. E., NEWMAN, A. B., VAN KAN, G. A., ANDRIEU, S., BAUER, J., BREUILLE, D., CEDERHOLM, T., CHANDLER, J., DE MEYNARD, C., DONINI, L., HARRIS, T., KANNT, A., GUIBERT, F. K., ONDER, G., PAPANICOLAOU, D., ROLLAND, Y., ROOKS, D., SIEBER, C., SOUHAMI, E., VERLAAN, S. & ZAMBONI, M. 2011. Sarcopenia: An Undiagnosed Condition in Older Adults. Current Consensus Definition: Prevalence, Etiology, and Consequences. International Working Group on Sarcopenia. *Journal of the American Medical Directors Association,* 12**,** 249-256.

FRIED, L. P., TANGEN, C. M., WALSTON, J., NEWMAN, A. B., HIRSCH, C., GOTTDIENER, J., SEEMAN, T., TRACY, R., KOP, W. J., BURKE, G., MCBURNIE, M. A. & CARDIOVASCULAR HLTH STUDY, C. 2001. Frailty in older adults: Evidence for a phenotype. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences,* 56**,** M146-M156.

GATT, M., ANDERSON, A. D. G., REDDY, B. S., HAYWARD-SAMPSON, P., TRING, I. C. & MACFIE, J. 2005. Randomized clinical trial of multimodal optimization of surgical care in patients undergoing major colonic resection. *British Journal of Surgery,* 92**,** 1354-1362.

GILLIS, C., LI, C., LEE, L., AWASTHI, R., AUGUSTIN, B., GAMSA, A., LIBERMAN, A. S., STEIN, B., CHARLEBOIS, P., FELDMAN, L. S. & CARLI, F. 2014. Prehabilitation versus Rehabilitation A Randomized Control Trial in Patients Undergoing Colorectal Resection for Cancer. *Anesthesiology,* 121**,** 937-947.

GRANT, S., AITCHISON, T., HENDERSON, E., CHRISTIE, J., ZARE, S., MCMURRAY, J. & DARGIE, H. 1999. A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest,* 116**,** 1208-1217.

GUYATT, G. H., TOWNSEND, M., BERMAN, L. B. & KELLER, J. L. 1987. A COMPARISON OF LIKERT AND VISUAL ANALOG SCALES FOR MEASURING CHANGE IN FUNCTION. *Journal of Chronic Diseases,* 40**,** 1129-1133.

HAMAGUCHI, Y., KAIDO, T., OKUMURA, S., KOBAYASHI, A., FUJIMOTO, Y., OGAWA, K., MORI, A., HAMMAD, A., HATANO, E. & UEMOTO, S. 2016. Muscle Steatosis is an Independent Predictor of Postoperative Complications in Patients with Hepatocellular Carcinoma. *World Journal of Surgery,* 40**,** 1959-1968.

HARADA, N. D., CHIU, V. & STEWART, A. L. 1999. Mobility-related function in older adults: Assessment with a 6-minute walk test. *Archives of Physical Medicine and Rehabilitation,* 80**,** 837-841.

HERDMAN, M., GUDEX, C., LLOYD, A., JANSSEN, M. F., KIND, P., PARKIN, D., BONSEL, G. & BADIA, X. 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research,* 20**,** 1727-1736.

HERNON, J. 2018. *SupPoRtive Exercise Programmes for Accelerating REcovery after major ABdominal Cancer surgery (PREPARE-ABC)* [Online]. https://www.acpgbi.org.uk/content/uploads/2017/10/prepare-abc.pdf: Association of Coloproctology Great Britain and Ireland. [Accessed 27/05/2019 2019].

HEYMSFIELD, S. B., WANG, Z. M., BAUMGARTNER, R. N. & ROSS, R. 1997. Human body composition: Advances in models and methods. *Annual Review of Nutrition,* 17**,** 527-558.

HOLDEN, D. A., RICE, T. W., STELMACH, K. & MEEKER, D. P. 1992. EXERCISE TESTING, 6-MIN WALK, AND STAIR CLIMB IN THE EVALUATION OF PATIENTS AT HIGH-RISK FOR PULMONARY RESECTION. *Chest,* 102**,** 1774-1779.

HUGHES, M. J., HACKNEY, R. J., LAMB, P. J., WIGMORE, S. J., CHRISTOPHER DEANS, D. A. & SKIPWORTH, R. J. E. 2019. Prehabilitation Before Major Abdominal Surgery: A Systematic Review and Meta-analysis. *World journal of surgery*.

INOUE, J., ONO, R., MAKIURA, D., KASHIWA-MOTOYAMA, M., MIURA, Y., USAMI, M., NAKAMURA, T., IMANISHI, T. & KURODA, D. 2013. Prevention of postoperative pulmonary complications through intensive preoperative respiratory rehabilitation in patients with esophageal cancer. *Diseases of the Esophagus,* 26**,** 68-74.

IRVING, B. A., RUTKOWSKI, J., BROCK, D. W., DAVIS, C. K., BARRETT, E. J., GAESSER, G. A. & WELTMAN, A. 2006. Comparison of Borg- and OMNI-RPE as markers of the blood lactate response to exercise. *Medicine and Science in Sports and Exercise,* 38**,** 1348-1352.

JANSSEN, I., HEYMSFIELD, S. B. & ROSS, R. 2002. Low relative skeletal muscle mass (sarcopenia) in older persons is associated with functional impairment and physical disability. *Journal of the American Geriatrics Society,* 50**,** 889-896.

JENSEN, B. T., PETERSEN, A. K., JENSEN, J. B., LAUSTSEN, S. & BORRE, M. 2015. Efficacy of a multiprofessional rehabilitation programme in radical cystectomy pathways: A prospective

randomized controlled trial. *Scandinavian Journal of Urology,* 49**,** 133-141.
JOGLEKAR, S., ASGHAR, A., MOTT, S. L., JOHNSON, B. E., BUTTON, A. M., CLARK, E. & MEZHIR, J. J. 2015. Sarcopenia Is an Independent Predictor of Complications Following Pancreatectomy for Adenocarcinoma. *Journal of Surgical Oncology,* 111**,** 771-775.
JOLLY, K., TAYLOR, R., LIP, G. Y. H., GREENFIELD, S., RAFTERY, J., MANT, J., LANE, D., JONES, M., LEE, K. W. & STEVENS, A. 2007. The Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Home-based compared with hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence. *Health Technology Assessment,* 11**,** 1-+.
JONES, L. W., COURNEYA, K. S., FAIREY, A. S. & MACKEY, J. R. 2004. Effects of an oncologist's recommendation to exercise on self-reported exercise behavior in newly diagnosed breast cancer survivors: A single-blind, randomized controlled trial. *Annals of Behavioral Medicine,* 28**,** 105-113.
JOSE TARAZONA-SANTABALBINA, F., CARMEN GOMEZ-CABRERA, M., PEREZ-ROS, P., MIGUEL MARTINEZ-ARNAU, F., CABO, H., TSAPARAS, K., SALVADOR-PASCUAL, A., RODRIGUEZ-MANAS, L. & VINA, J. 2016. A Multicomponent Exercise Intervention that Reverses Frailty and Improves Cognition, Emotion, and Social Networking in the Community-Dwelling Frail Elderly: A Randomized Clinical Trial. *Journal of the American Medical Directors Association,* 17**,** 426-433.
KAIBORI, M., ISHIZAKI, M., MATSUI, K., NAKATAKE, R., YOSHIUCHI, S., KIMURA, Y. & KWON, A. H. 2013. Perioperative exercise for chronic liver injury patients with hepatocellular carcinoma undergoing hepatectomy. *American Journal of Surgery,* 206**,** 202-209.
KEHLET, H. & WILMORE, D. W. 2005. Fast-track surgery. *British Journal of Surgery,* 92**,** 3-4.
KEHLET, H. & WILMORE, D. W. 2008. Evidence-based surgical care and the evolution of fast-track surgery. *Annals of Surgery,* 248**,** 189-198.
KIM, D. J., MAYO, N. E., CARLI, F., MONTGOMERY, D. L. & ZAVORSKY, G. S. 2009. Responsive Measures to Prehabilitation in Patients Undergoing Bowel Resection Surgery. *Tohoku Journal of Experimental Medicine,* 217**,** 109-115.
KING, M. T. 1996. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Quality of Life Research,* 5**,** 555-567.
KOBAYASHI, A., KAIDO, T., HAMAGUCHI, Y., OKUMURA, S., TAURA, K., HATANO, E., OKAJIMA, H. & UEMOTO, S. 2016. Impact of postoperative changes in sarcopenic factors on

outcomes after hepatectomy for hepatocellular carcinoma. *Journal of Hepato-Biliary-Pancreatic Sciences,* 23**,** 57-64.
KULKARNI, S. R., FLETCHER, E., MCCONNELL, A. K., POSKITT, K. R. & WHYMAN, M. R. 2010. Pre-operative inspiratory muscle training preserves postoperative inspiratory muscle strength following major abdominal surgery - a randomised pilot study. *Annals of the Royal College of Surgeons of England,* 92**,** 700-705.
LATIMER-CHEUNG, A. E., PILUTTI, L. A., HICKS, A. L., GINIS, K. A. M., FENUTA, A. M., MACKIBBON, K. A. & MOTL, R. W. 2013. Effects of Exercise Training on Fitness, Mobility, Fatigue, and Health-Related Quality of Life Among Adults With Multiple Sclerosis: A Systematic Review to Inform Guideline Development. *Archives of Physical Medicine and Rehabilitation,* 94**,** 1800-1828.
LEE, L., SCHWARTZMAN, K., CARLI, F., ZAVORSKY, G. S., LI, C., CHARLEBOIS, P., STEIN, B., LIBERMAN, A. S., FRIED, G. M. & FELDMAN, L. S. 2013. The association of the distance walked in 6 min with pre-operative peak oxygen consumption and complications 1 month after colorectal resection. *Anaesthesia,* 68**,** 811-816.
LENNON, O., CAREY, A., GAFFNEY, N., STEPHENSON, J. & BLAKE, C. 2008. A pilot randomized controlled trial to evaluate the benefit of the cardiac rehabilitation paradigm for the non-acute ischaemic stroke population. *Clinical Rehabilitation,* 22**,** 125-133.
LEVETT, D. Z. H. & GROCOTT, M. P. W. 2015. Cardiopulmonary exercise testing, prehabilitation, and Enhanced Recovery After Surgery (ERAS). *Canadian Journal of Anesthesia-Journal Canadien D Anesthesie,* 62**,** 131-142.
LIEFFERS, J. R., BATHE, O. F., FASSBENDER, K., WINGET, M. & BARACOS, V. E. 2012. Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery. *British Journal of Cancer,* 107**,** 931-936.
LIPKIN, D. P., SCRIVEN, A. J., CRAKE, T. & POOLEWILSON, P. A. 1986. 6 MINUTE WALKING TEST FOR ASSESSING EXERCISE CAPACITY IN CHRONIC HEART-FAILURE. *British Medical Journal,* 292**,** 653-655.
LLORENS, J., ROVIRA, L., BALLESTER, M., MORENO, J., HERNANDEZ-LAFORET, J., SANTONJA, F.-J., CASSINELLO, N. & ORTEGA, J. 2015. Preoperative Inspiratory Muscular Training to Prevent Postoperative Hypoxemia in Morbidly Obese Patients Undergoing Laparoscopic Bariatric Surgery. A Randomized Clinical Trial. *Obesity Surgery,* 25**,** 1003-1009.
LUTHER, A., GABRIEL, J., WATSON, R. P. & FRANCIS, N. K. 2018. The Impact of Total Body Prehabilitation on Post-Operative Outcomes After Major Abdominal Surgery: A Systematic Review. *World Journal of Surgery,* 42**,** 2781-2791.

MALIETZIS, G., JOHNS, N., AL-HASSI, H. O., KNIGHT, S. C., KENNEDY, R. H., FEARON, K. C. H., AZIZ, O. & JENKINS, J. T. 2016. Low Muscularity and Myosteatosis Is Related to the Host Systemic Inflammatory Response in Patients Undergoing Surgery for Colorectal Cancer. *Annals of Surgery,* 263**,** 320-325.

MAYO, N. E., ANDERSON, S., BARCLAY, R., CAMERON, J. I., DESROSIERS, J., ENG, J. J., HUIJBREGTS, M., KAGAN, A., MACKAY-LYONS, M., MORIELLO, C., RICHARDS, C. L., SALBACH, N. M., SCOTT, S. C., TEASELL, R. & BAYLEY, M. 2015. Getting on with the rest of your life following stroke: a randomized trial of a complex intervention aimed at enhancing life participation post stroke. *Clinical Rehabilitation,* 29**,** 1198-1211.

MAYO, N. E., FELDMAN, L., SCOTT, S., ZAVORSKY, G., KIM, D. J., CHARLEBOIS, P., STEIN, B. & CAR, F. 2011. Impact of preoperative change in physical function on postoperative recovery: Argument supporting prehabilitation for colorectal surgery. *Surgery,* 150**,** 505-514.

MCDONALD, C. M., HENRICSON, E. K., HAN, J. J., ABRESCH, R. T., NICORICI, A., ELFRING, G. L., ATKINSON, L., REHA, A., HIRAWAT, S. & MILLER, L. L. 2010. THE 6-MINUTE WALK TEST AS A NEW OUTCOME MEASURE IN DUCHENNE MUSCULAR DYSTROPHY. *Muscle & Nerve,* 41**,** 500-510.

MCHORNEY, C. A., WARE, J. E. & RACZEK, A. E. 1993. THE MOS 36-ITEM SHORT-FORM HEALTH SURVEY (SF-36) .2. PSYCHOMETRIC AND CLINICAL-TESTS OF VALIDITY IN MEASURING PHYSICAL AND MENTAL-HEALTH CONSTRUCTS. *Medical Care,* 31**,** 247-263.

MCLENNAN, E., OLIPHANT, R. & MOUG, S. J. 2019. Limited preoperative physical capacity continues to be associated with poor postoperative outcomes within a colorectal ERAS programme. *Annals of the Royal College of Surgeons of England,* 101**,** 261-267.

MCNAUGHT, C. E. & MACFIE, J. 2002. Optimization of surgical care. *Nutrition,* 18**,** 194-196.

MICHELSON, H., BOLUND, C., NILSSON, B. & BRANDBERG, Y. 2000. Health-related quality of life measured by the EORTC QLQ-C30 - Reference values from a large sample of the Swedish population. *Acta Oncologica,* 39**,** 477-484.

MINNELLA, E. M., BOUSQUET-DION, G., AWASTHI, R., SCHEEDE-BERGDAHL, C. & CARLI, F. 2017. Multimodal prehabilitation improves functional capacity before and after colorectal surgery for cancer: a five-year research experience. *Acta Oncologica,* 56**,** 295-300.

MITNITSKI, A. 2001. Rockwood Accumulation of deficits as a proxy measure of aging. *In:* MOGILNER, A. (ed.). Scientific World Journal.

MOALLA, W., GAUTHIER, R., MAINGOURD, Y. & AHMAIDI, S. 2005. Six-minute walking test to assess exercise tolerance

and cardiorespiratory responses during training program in children with congenital heart disease. *International Journal of Sports Medicine,* 26**,** 756-762.

MOCK, V., FRANGAKIS, C., DAVIDSON, N. E., ROPKA, M. E., PICKETT, M., PONIATOWSKI, B., STEWART, K. J., CAMERON, L., ZAWACKI, K., PODEWILS, L. J., COHEN, G. & MCCORKLE, R. 2005. Exercise manages fatigue during breast cancer treatment: A randomized controlled trial. *Psycho-Oncology,* 14**,** 464-477.

MOK, M., ALLENDE, R., LEIPSIC, J., ALTISENT, O. A. J., DEL TRIGO, M., CAMPELO-PARADA, F., DELAROCHELLIERE, R., DUMONT, E., DOYLE, D., COTE, M., FREEMAN, M., WEBB, J. & RODES-CABAU, J. 2016. Prognostic Value of Fat Mass and Skeletal Muscle Mass Determined by Computed Tomography in Patients Who Underwent Transcatheter Aortic Valve Implantation. *American Journal of Cardiology,* 117**,** 828-833.

MORLEY, J. E., ABBATECOLA, A. M., ARGILES, J. M., BARACOS, V., BAUER, J., BHASIN, S., CEDERHOLM, T., COATS, A. J. S., CUMMINGS, S. R., EVANS, W. J., FEARON, K., FERRUCCI, L., FIELDING, R. A., GURALNIK, J. M., HARRIS, T. B., INUI, A., KALANTAR-ZADEH, K., KIRWAN, B. A., MANTOVANI, G., MUSCARITOLI, M., NEWMAN, A. B., ROSSI-FANELLI, F., ROSANO, G. M. C., ROUBENOFF, R., SCHAMBELAN, M., SOKOL, G. H., STORER, T. W., VELLAS, B., VON HAEHLING, S., YEH, S. S., ANKER, S. D. & SOC SARCOPENIA CACHEXIA, W. 2011. Sarcopenia With Limited Mobility: An International Consensus. *Journal of the American Medical Directors Association,* 12**,** 403-409.

MOUG, S. J., MUTRIE, N., BARRY, S. J. E., MACKAY, G., STEELE, R. J. C., BOACHIE, C., BUCHAN, C. & ANDERSON, A. S. 2019. Prehabilitation is feasible in patients with rectal cancer undergoing neoadjuvant chemoradiotherapy and may minimize physical deterioration: results from the REx trial. *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland*.

NBOCA. 2017. *National Bowel Cancer Audit. 2017 Annual Report* [Online]. https://www.nboca.org.uk/content/uploads/2017/12/NBOCA-annual-report-2017-v2.pdf. [Accessed].

NELA. 2017. *Third Patient Report of the National Emergency Laparotomy Audit (NELA) December 2015 to November 2016 EXECUTIVE SUMMARY* [Online]. https://www.nela.org.uk/reports: Royal Collage of Anaesthetists. [Accessed September 15th 2018].

NINOMIYA, G., FUJII, T., YAMADA, S., YABUSAKI, N., SUZUKI, K., IWATA, N., KANDA, M., HAYASHI, M., TANAKA, C., NAKAYAMA, G., SUGIMOTO, H., KOIKE, M., FUJIWARA, M. & KODERA, Y. 2017. Clinical impact of sarcopenia on prognosis in pancreatic ductal adenocarcinoma: A

retrospective cohort study. *International Journal of Surgery,* 39**,** 45-51.
NISHIGORI, T., TSUNODA, S., OKABE, H., TANAKA, E., HISAMORI, S., HOSOGI, H., SHINOHARA, H. & SAKAI, Y. 2016. Impact of Sarcopenic Obesity on Surgical Site Infection after Laparoscopic Total Gastrectomy. *Annals of Surgical Oncology,* 23**,** 524-531.
O'DONNELL, D. E., FLUGE, T., GERKEN, F., HAMILTON, A., WEBB, K., AGUILANIU, B., MAKE, B. & MAGNUSSEN, H. 2004. Effects of tiotropium on lung hyperinflation, dyspnoea and exercise tolerance in COPD. *European Respiratory Journal,* 23**,** 832-840.
O'NEILL, L. M., GUINAN, E., DOYLE, S. L., BENNETT, A. E., MURPHY, C., ELLIOTT, J. A., O'SULLIVAN, J., REYNOLDS, J. V. & HUSSEY, J. 2018. The RESTORE Randomized Controlled Trial Impact of a Multidisciplinary Rehabilitative Program on Cardiorespiratory Fitness in Esophagogastric cancer Survivorship. *Annals of Surgery,* 268**,** 747-755.
OKUMURA, S., KAIDO, T., HAMAGUCHI, Y., FUJIMOTO, Y., KOBAYASHI, A., IIDA, T., YAGI, S., TAURA, K., HATANO, E. & UEMOTO, S. 2016. Impact of the preoperative quantity and quality of skeletal muscle on outcomes after resection of extrahepatic biliary malignancies. *Surgery,* 159**,** 821-833.
OLDER, P., SMITH, R., COURTNEY, P. & HONE, R. 1993. PREOPERATIVE EVALUATION OF CARDIAC-FAILURE AND ISCHEMIA IN ELDERLY PATIENTS BY CARDIOPULMONARY EXERCISE TESTING. *Chest,* 104**,** 701-704.
PALANGE, P., WARD, S. A., CARLSEN, K. H., CASABURI, R., GALLAGHER, C. G., GOSSELINK, R., O'DONNELL, D. E., PUENTE-MAESTU, L., SCHOLS, A. M., SINGH, S. & WHIPP, B. J. 2007. Recommendations on the use of exercise testing in clinical practice. *European Respiratory Journal,* 29**,** 185-209.
PARMAR, K., LAW, J., CARTER, B., HEWITT, J., BOYLE, J., CASEY, P., MAITRA, I., FARRELL, I., PEARCE, L. & MOUG, S. 2019. Frailty in Older Patients Undergoing Emergency Laparotomy: Results From the UK Observational Emergency Laparotomy and Frailty (ELF) Study. Ann. Sure.
PARTRIDGE, J. S. L., HARARI, D., MARTIN, F. C., PEACOCK, J. L., BELL, R., MOHAMMED, A. & DHESI, J. K. 2017. Randomized clinical trial of comprehensive geriatric assessment and optimization in vascular surgery. *British Journal of Surgery,* 104**,** 679-687.
PATEL, M. S., ASCH, D. A. & VOLPP, K. G. 2015. Wearable Devices as Facilitators, Not Drivers, of Health Behavior Change. *Jama-Journal of the American Medical Association,* 313**,** 459-460.
PEARSE, R. M., HARRISON, D. A., JAMES, P., WATSON, D., HINDS, C., RHODES, A., GROUNDS, R. M. & BENNETT, E.

D. 2006. Identification and characterisation of the high-risk surgical population in the United Kingdom. *Critical Care,* 10.
PECORELLI, N., CARRARA, G., DE COBELLI, F., CRISTEL, G., DAMASCELLI, A., BALZANO, G., BERETTA, L. & BRAGA, M. 2016a. Effect of sarcopenia and visceral obesity on mortality and pancreatic fistula following pancreatic cancer surgery. *British Journal of Surgery,* 103**,** 434-442.
PECORELLI, N., FIORE, J. F., GILLIS, C., AWASTHI, R., MAPPIN-KASIRER, B., NICULISEANU, P., FRIED, G. M., CARLI, F. & FELDMAN, L. S. 2016b. The six-minute walk test as a measure of postoperative recovery after colorectal resection: further examination of its measurement properties. *Surgical Endoscopy and Other Interventional Techniques,* 30**,** 2199-2206.
PENEDO, F. J. & DAHN, J. R. 2005. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current Opinion in Psychiatry,* 18**,** 189-193.
PHAN, T. L. T., BARNINI, N., XIE, S., MARTINEZ, A., FALINI, L., ABATEMARCO, A., WALDRON, M., BENTON, J. M., FRANKENBERRY, S., COLEMAN, C., NGUYEN, L., BO, C., DATTO, G. A. & WERK, L. N. 2018. Feasibility of Using a Commercial Fitness Tracker as an Adjunct to Family-Based Weight Management Treatment: Pilot Randomized Trial. *Jmir Mhealth and Uhealth,* 6.
PRENTIS, J. M., MANAS, D. M. D., TRENELL, M. I., HUDSON, M., JONES, D. J. & SNOWDEN, C. P. 2012. Submaximal cardiopulmonary exercise testing predicts 90-day survival after liver transplantation. *Liver Transplantation,* 18**,** 152-159.
PRETTY, J., PEACOCK, J., SELLENS, M. & GRIFFIN, M. 2005. The mental and physical health outcomes of green exercise. *International Journal of Environmental Health Research,* 15**,** 319-337.
RABIN, R. & DE CHARRO, F. 2001. EQ-5D: a measure of health status from the EuroQol Group. *Annals of Medicine,* 33**,** 337-343.
ROCKWOOD, K., HOWLETT, S. E., MACKNIGHT, C., BEATTIE, B. L., BERGMAN, H., HEBERT, R., HOGAN, D. B., WOLFSON, C. & MCDOWELL, I. 2004. Prevalence, attributes, and outcomes of fitness and frailty in community-dwelling older adults: Report from the Canadian Study of Health and Aging. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences,* 59**,** 1310-1317.
ROCKWOOD, K., SONG, X., MACKNIGHT, C., BERGMAN, H., HOGAN, D., MCDOWELL, I. & MITNITSKI, A. 2005. A global clinical measure of fitness and frailty in elderly people. CMAJ.
RODRIGUEZ-MANAS, L., FEART, C., MANN, G., VINA, J., CHATTERJI, S., CHODZKO-ZAJKO, W., HARMAND, M. G. C., BERGMAN, H., CARCAILLON, L., NICHOLSON, C., SCUTERI, A., SINCLAIR, A., PELAEZ, M., VAN DER CAMMEN, T., BELAND, F., BICKENBACH, J.,

DELAMARCHE, P., FERRUCCI, L., FRIED, L. P., GUTIERREZ-ROBLEDO, L. M., ROCKWOOD, K., ARTALEJO, F. R., SERVIDDIO, G., VEGA, E. & GRP, F.-C. 2013. Searching for an Operational Definition of Frailty: A Delphi Method Based Consensus Statement. The Frailty Operative Definition-Consensus Conference Project. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences,* 68**,** 62-67.

ROLLAND, Y., CZERWINSKI, S., VAN KAN, G. A., MORLEY, J. E., CESARI, M., ONDER, G., WOO, J., BAUMGARTNER, R., PILLARD, F., BOIRIE, Y., CHUMLEA, W. M. C. & VELLAS, B. 2008. Sarcopenia: Its assessment, etiology, pathogenesis, consequences and future perspectives. *Journal of Nutrition Health & Aging,* 12**,** 433-450.

ROSENBERG 1989. Summary comments: epidemiological and methodological problems in determining nutritional status of older persons Am J Clin Nutr

ROSTAGNO, C. & GENSINI, G. F. 2008. Six minute walk test: a simple and useful test to evaluate functional capacity in patients with heart failure. *Internal and Emergency Medicine,* 3**,** 205-212.

ROUTLEDGE, H. C. & TOWNEND, J. N. 2006. Why does the heart rate response to exercise predict adverse cardiac events? *Heart,* 92**,** 577-578.

SEGAL, R., EVANS, W., JOHNSON, D., SMITH, J., COLLETTA, S., GAYTON, J., WOODARD, S., WELLS, G. & REID, R. 2001. Structured exercise improves physical functioning in women with stages I and II breast cancer: Results of a randomized controlled trial. *Journal of Clinical Oncology,* 19**,** 657-665.

SEGAL, R. J., REID, R. D., COURNEYA, K. S., MALONE, S. C., PARLIAMENT, M. B., SCOTT, C. G., VENNER, P. M., QUINNEY, H. A., JONES, L. W., D'ANGELO, M. E. S. & WELLS, G. A. 2003. Resistance exercise in men receiving androgen deprivation therapy for prostate cancer. *Journal of Clinical Oncology,* 21**,** 1653-1659.

SHEN, W., PUNYANITYA, M., WANG, Z. M., GALLAGHER, D., ST-ONGE, M. P., ALBU, J., HEYMSFIELD, S. B. & HESHKA, S. 2004. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *Journal of Applied Physiology,* 97**,** 2333-2338.

SHULMAN, M. A., CUTHBERTSON, B. H., WIJEYSUNDERA, D. N., PEARSE, R. M., THOMPSON, B., TORRES, E., AMBOSTA, A., WALLACE, S., FARRINGTON, C., MYLES, P. S. & MEASUREMENT EXERCISE TOLERANCE, S. 2019. Using the 6-minute walk test to predict disability-free survival after major surgery. *British Journal of Anaesthesia,* 122**,** 111-119.

SINGH, S. J., MORGAN, M. D. L., SCOTT, S., WALTERS, D. & HARDMAN, A. E. 1992. DEVELOPMENT OF A SHUTTLE WALKING TEST OF DISABILITY IN PATIENTS WITH

CHRONIC AIRWAYS OBSTRUCTION. *Thorax,* 47**,** 1019-1024.
SMITH, S. M., ZHANG, Y. Y., JENKINSON, M., CHEN, J., MATTHEWS, P. M., FEDERICO, A. & DE STEFANO, N. 2002. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage,* 17**,** 479-489.
SNOWDEN, C. P., PRENTIS, J. M., ANDERSON, H. L., ROBERTS, D. R., RANDLES, D., RENTON, M. & MANAS, D. M. 2010. Submaximal Cardiopulmonary Exercise Testing Predicts Complications and Hospital Length of Stay in Patients Undergoing Major Elective Surgery. *Annals of Surgery,* 251**,** 535-541.
SOLWAY, S., BROOKS, D., LACASSE, Y. & THOMAS, S. 2001. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest,* 119**,** 256-270.
SPINHOVEN, P., ORMEL, J., SLOEKERS, P. P. A., KEMPEN, G., SPECKENS, A. E. M. & VANHEMERT, A. M. 1997. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychological Medicine,* 27**,** 363-370.
SPORT-ENGLAND. 2017a. *Active Lives Adult Survey* [Online]. https://www.sportengland.org/media/13038/active-lives-adult-survey-nov-16-17-report.pdf. [Accessed].
SPORT-ENGLAND. 2017b. *Active Lives Adult Survey* [Online]. https://www.sportengland.org/media/13038/active-lives-adult-survey-nov-16-17-report.pdf. [Accessed 21 March 2018].
STATHOPOULOU, G., POWERS, M. B., BERRY, A. C., SMITS, J. A. J. & OTTO, M. W. 2006. Exercise interventions for mental health: A quantitative and qualitative review. *Clinical Psychology-Science and Practice,* 13**,** 179-193.
STERNBERG, S. A., SCHWARTZ, A. W., KARUNANANTHAN, S., BERGMAN, H. & CLARFIELD, A. M. 2011. The Identification of Frailty: A Systematic Literature Review. *Journal of the American Geriatrics Society,* 59**,** 2129-2138.
SUZUKI, Y., OKAMOTO, T., FUJISHITA, T., KATSURA, M., AKAMINE, T., TAKAMORI, S., MORODOMI, Y., TAGAWA, T., SHOJI, F. & MAEHARA, Y. 2016. Clinical implications of sarcopenia in patients undergoing complete resection for early non-small cell lung cancer. *Lung Cancer,* 101**,** 92-97.
TAKACS, J., POLLOCK, C. L., GUENTHER, J. R., BAHAR, M., NAPIER, C. & HUNT, M. A. 2014. Validation of the Fitbit One activity monitor device during treadmill walking. *Journal of Science and Medicine in Sport,* 17**,** 496-500.
TAMANDL, D., PAIREDER, M., ASARI, R., BALTZER, P. A., SCHOPPMANN, S. F. & BA-SSALAMAH, A. 2016. Markers of sarcopenia quantified by computed tomography predict adverse long-term outcome in patients with resected oesophageal or gastro-oesophageal junction cancer. *European Radiology,* 26**,** 1359-1367.

TEW, G. A., BATTERHAM, A. M., COLLING, K., GRAY, J., KERR, K., KOTHMANN, E., NAWAZ, S., WESTON, M., YATES, D. & DANJOUX, G. 2017. Randomized feasibility trial of high-intensity interval training before elective abdominal aortic aneurysm repair. *British Journal of Surgery,* 104**,** 1791-1801.
THEOU, O., STATHOKOSTAS, L., ROLAND, K. P., JAKOBI, J. M., PATTERSON, C., VANDERVOORT, A. A. & JONES, G. R. 2011. The effectiveness of exercise interventions for the management of frailty: a systematic review. *Journal of aging research,* 2011**,** 569194-569194.
THOMAS, K. S., MUIR, K. R., DOHERTY, M., JONES, A. C., O'REILLY, S. C., BASSEY, E. J. & COMMUNITY OSTEOARTHRITIS RES, G. 2002. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. *Bmj-British Medical Journal,* 325**,** 752-755.
THORSEN, L., SKOVLUND, E., STROMME, S. B., HORNSLIEN, K., DAHL, A. A. & FOSSA, S. D. 2005. Effectiveness of physical activity on cardiorespiratory fitness and health-related quality of life in young and middle-aged cancer patients shortly after chemotherapy. *Journal of Clinical Oncology,* 23**,** 2378-2388.
TROTTER, J., JOHNSTON, J., NG, A., GATT, M., MACFIE, J. & MCNAUGHT, C. 2018. Is sarcopenia a useful predictor of outcome in patients after emergency laparotomy? A study using the NELA database. *Annals of the Royal College of Surgeons of England,* 100**,** 377-381.
UMMELS, D., BEEKMAN, E., MOSER, A., BRAUN, S. M. & BEURSKENS, A. J. Patients' experiences with commercially available activity trackers embedded in physiotherapy treatment: a qualitative study. *Disability and Rehabilitation*.
VAES, A. W., DELBRESSINE, J. M. L., MESQUITA, R., GOERTZ, Y. M. J., JANSSEN, D. J. A., NAKKEN, N., FRANSSEN, F. M. E., VANFLETEREN, L. E. G. W., WOUTERS, E. F. M. & SPRUIT, M. A. 2019. Impact of pulmonary rehabilitation on activities of daily living in patients with chronic obstructive pulmonary disease. *Journal of Applied Physiology,* 126**,** 607-615.
VAN ADRICHEM, E. J., MEULENBROEK, R. L., PLUKKER, J. T. M., GROEN, H. & VAN WEERT, E. 2014. Comparison of Two Preoperative Inspiratory Muscle Training Programs to Prevent Pulmonary Complications in Patients Undergoing Esophagectomy: A Randomized Controlled Pilot Study. *Annals of Surgical Oncology,* 21**,** 2353-2360.
VAN WAART, H., STUIVER, M. M., VAN HARTEN, W. H., GELEIJN, E., KIEFFER, J. M., BUFFART, L. M., DE MAAKER-BERKHOF, M., BOVEN, E., SCHRAMA, J., GEENEN, M. M., TERWOGT, J. M. M., VAN BOCHOVE, A., LUSTIG, V., VAN DEN HEILIGENBERG, S. M., SMORENBURG, C. H., HELLENDOORN-VAN VREESWIJK, J., SONKE, G. S. & AARONSON, N. K. 2015. Effect of Low-Intensity Physical Activity and Moderate-to High-Intensity Physical Exercise

During Adjuvant Chemotherapy on Physical Fitness, Fatigue, and Chemotherapy Completion Rates: Results of the PACES Randomized Clinical Trial. *Journal of Clinical Oncology,* 33**,** 1918-U90.

WAGNER, D., BUETTNER, S., KIM, Y., GANI, F., XU, L., MARGONIS, G. A., AMINI, N., KAMEL, I. R. & PAWLIK, T. M. 2016. Clinical and morphometric parameters of frailty for prediction of mortality following hepatopancreaticobiliary surgery in the elderly. *British Journal of Surgery,* 103**,** E83-E92.

WARE, J. E. & SHERBOURNE, C. D. 1992. THE MOS 36-ITEM SHORT-FORM HEALTH SURVEY (SF-36) .1. CONCEPTUAL-FRAMEWORK AND ITEM SELECTION. *Medical Care,* 30**,** 473-483.

WERNER, D. A. T., GROTLE, M., GULATI, S., AUSTEVOLL, I. M., LONNE, G., NYGAARD, O. P. & SOLBERG, T. K. 2017. Criteria for failure and worsening after surgery for lumbar disc herniation: a multicenter observational study based on data from the Norwegian Registry for Spine Surgery. *European Spine Journal,* 26**,** 2650-2659.

WHO 2010. Global recommendations on physical activity for health. http://www.who.int: World Health Organisation.

WIGODSKI, S., CARRASCO, F., BUNOUT, D., BARRERA, G., HIRSCH, S. & DE LA MAZA, M. P. 2018. Sarcopenia: The need to establish different cutting points of fat-free mass for the Chilean population. *Nutrition (Burbank, Los Angeles County, Calif.),* 57**,** 217-224.

WILKINSON. 2010. *An Age Old Problem A review of the care received by elderly patients undergoing surgery* [Online]. Available: http://www.ncepod.org.uk/2010report3/downloads/EESE_fullReport.pdf: NCEPOD. [Accessed 27th March 2018].

WU, M., MCINNES, M., MACDONNALD, B., KIELAR, A. & DUIGENAN, S. 2013. CT in Adults: Systematic Review and Meta-Analysis of Interpretation Discrepancy Rates. Radiology.

YOOWANNAKUL, S., TANGVORAPHONKCHAI, K. & DAVENPORT, A. 2018a. The prevalence of muscle wasting (sarcopenia) in peritoneal dialysis patients varies with ethnicity due to differences in muscle mass measured by bioimpedance. *European Journal of Clinical Nutrition,* 72**,** 381-387.

YOOWANNAKUL, S., TANGVORAPHONKCHAI, K., VONGSANIM, S., MOHAMED, A. & DAVENPORT, A. 2018b. Differences in the prevalence of sarcopenia in haemodialysis patients: the effects of gender and ethnicity. *Journal of human nutrition and dietetics : the official journal of the British Dietetic Association*.

ZAMBONI, M., MAZZALI, G., FANTIN, F., ROSSI, A. & DI FRANCESCO, V. 2008. Sarcopenic obesity: A new category of obesity in the elderly. *Nutrition Metabolism and Cardiovascular Diseases,* 18**,** 388-395.

ZHONG, K., LEI, S.-F., FANG, Y. & CHEN, X.-D. 2012. The differences of sarcopenia-related phenotypes: effects of gender and population. <i data-test="journal-title" style="box-sizing: inherit; caret-color: rgb(51, 51, 51); color: rgb(51, 51, 51); font-family: "Trebuchet MS", -apple-system, BlinkMacSystemFont, "Segoe UI", Roboto, Oxygen-Sans, Ubuntu, Cantarell, "Helvetica Neue", sans-serif; font-size: 14px;">European Review of Aging and Physical Activity volume.