Evaluation of the ‘Live Active’ exercise referral scheme

Faye Prior

School of Health and Society
University of Salford

Submitted in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

2019

Table of contents

Chapter 5 Study 1: Evaluation of uptake and adherence to the 'Live Active' exercise referral scheme 82

List of Tables

List of Figures

List of appendices

Acknowledgements

I am sincerely thankful to my supervisors, Margaret Coffey, Anna Robins and Penny Cook. This work would not have been possible without your time, guidance and support over the last four years.

Thank you to all of the staff, past and present, at Live Active. This work would also not have been possible without your daily effort and dedication to the scheme, your support, or endless supply of coffee and biscuits. Also, a special thanks to the participants of Live Active who agreed to participate in this research, your support was much appreciated.

Thank you to my friends and family who have encouraged and supported me through this PhD. Lastly, thank you to my dog Blake, who sacrificed several walks during this process, and snored the entire way through the write up.

List of frequently used abbreviations

BMI	Body mass index
BP	Blood pressure
CVD	Cardiovascular disease
DoH	Department of Health
EQ-5D-3L	EuroQol five dimensions three level questionnaire
EQ-5D VAS	EuroQol five dimensions visual analogue scale
ERS	Exercise referral scheme
HRQoL	Health-related quality of life
IPAQ	International physical activity questionnaire
IQR	Interquartile range
LTC	Long-term health condition
METs	Metabolic equivalents
MVPA	Moderate-vigorous intensity physical activity
NICE	National Institute for Health and Care Excellence
PA	Physical activity
PHE	Public Health England
SE	Socio-economic
SES	Socio-economic status
SEF	Standard evaluation framework
TTM	Transtheoretical model
WEMWBS	Warwick-Edinburgh mental wellbeing scale
WHO	World Health Organization

Abstract

Exercise referral schemes (ERSs) are a common approach to physical activity (PA) promotion in the United Kingdom (UK). They aim to reduce levels of physical inactivity amongst adults with long-term health conditions or risk factors.There are over 600 ERSs in the UK, but the evidence base for their impact is relatively small and equivocal, with further research required to understand if they are effective at increasing levels of health enhancing PA.

The overall aim of this thesis was to evaluate the impact of an existing ERS, the ‘Live Active’ scheme, at increasing PA levels and improving health outcomes in its participants. This was achieved through four studies, which i) explored engagement levels with the scheme, and if any characteristics could predict engagement; ii) identified the medium to long-term PA behaviour and health outcomes of participation, and if they differed by primary referral condition; iii) determined the perceived benefits of participation in the scheme, and; iv) identified short-term changes in accelerometer-measured PA and sedentary behaviour, and determined how these outcomes compared with self-report.

These studies demonstrated that the ‘Live Active’ ERS was successful at engaging participants, increasing long-term PA levels and improving health holistically. Firstly, uptake and adherence to the ERS were higher than expected, and were associated with a small number of factors such as season and smoking status. Analysis of secondary data identified long-term improvements in PA level and a range of health-related outcomes. However, the extent of improvement differed depending on a participant’s primary referral condition. Accelerometer-based measurement of PA demonstrated meaningful, but non-significant, increases in short-term PA. Discrepancies were also found between accelerometer-measured and self-reported PA and sedentary behaviour, the implications of which are discussed. Lastly, mixed methods exploration of the benefits of participation identified a wide range of holistic benefits. Many of the benefits reported qualitatively were not routinely evaluated by the scheme, or included in previous quantitative ERS evaluations. This highlighted that evaluating this scheme based solely on its routine quantitative outcomes would have underestimated the effectiveness of the intervention. Therefore, the findings of this thesis have identified methodological considerations for how ERSs are evaluated in the future, and whether the content of ERS evaluation frameworks provide a framework for a balanced appraisal of schemes.

Chapter 1 Introduction

1.1. Introduction

The aim of this chapter is to set for the scene for the role of physical activity (PA) in public health. It begins by discussing the burden of long-term health conditions (LTCs). The evidence base for PA as a risk factor for LTCs is summarised, and global and national levels of PA are discussed. A synthesis of recent local and national responses to physical inactivity is provided, in which the focus of this thesis is introduced. The chapter concludes with an overview of the context of the PhD, the overall aim of the research, and a structure of the thesis.

1.2. Long-term health conditions

Life expectancy in the UK is growing, albeit at a slower rate in recent years (Office for National Statistics [ONS], 2018a). However, the number of these extra years lived in good health is not growing at the same rate (ONS, 2017a). For the first time, individuals in the developing world are experiencing greater rises in years lived in good health than those in the developed world (Kyu et al., 2018; Murray et al., 2015). This can be attributed to changes in trends related to demography, epidemiology and socio-economic (SE) factors (Steel et al., 2018). There is, for example, a growing number of people in the United Kingdom (UK) living with a LTC, and an increase in the elderly population, who are likely to have age-related health conditions (Olshansky et al., 2005; Vos et al., 2012, 2015).

A LTC has been defined as “a condition that cannot, at present, be cured but is controlled by medication and/or other treatment/therapies” (Department of Health [DoH], 2012, p. 3), and includes conditions such as diabetes, hypertension, and arthritis (DoH, 2012). In England, the last available official figures indicate that 15 million people (a quarter of the population) have at least one LTC (DoH, 2012), with informal estimates suggesting this now stands at 26 million people (National Health Service [NHS], 2018; Social Care Institute for Excellence, 2018). These figures are of public health concern as in addition to the morbidity experienced, people with LTCs are frequent users of health care services, accounting for 50% of all general practitioner (GP) appointments, 70% of hospital bed days, and 70% of total health and social care expenditure (DoH, 2012).

The development of LTCs, and their prognosis once diagnosed are influenced by many factors, some of which are, to an extent, under the control of the individual. Some of the most common modifiable risk factors include physical inactivity, consuming a diet low in fruit and vegetables, smoking tobacco, alcohol and drug use, high body mass index (BMI), high blood pressure (BP), and high blood glucose levels (Forouzanfar et al., 2015; World Health Organization [WHO], 2011). However, the extent to which a person has control over these risk factors is variable, and dictated by broader determinants such as socio-economic status (SES) (Marchman Andersen, Dalton, Lynch, Johansen, & Holtug, 2013; Resnik, 2007, 2014). Health inequalities exist, for example, where LTCs are more prevalent in older age groups, and those who are more socio-economically deprived (Marmot et al., 2010; Newton et al., 2015; Public Health England [PHE], 2018a).

Physical inactivity as a risk factor has attracted increasing attention globally over the last decade with the emergence of strong evidence linking physical inactivity to the development of many major LTCs, and reduced life expectancy (Lee et al., 2012; Reiner, Niermann, Jekauc, & Woll, 2013; Warburton & Bredin, 2017).

1.3. Physical activity – definitions

There is a general consensus that physical activity can be defined as any bodily movement produced by skeletal muscle that results in energy expenditure (Caspersen, Powell, & Christenson, 1985; Chodzko-Zajko et al., 2009; DoH, 2011; PHE, 2014a). PA is a complex behaviour including many different forms of activity, which can be further classified into groups such as: occupational activity (e.g. carrying items at work), household activity (e.g. cleaning, gardening), leisure time activity (e.g. walking), active transport (e.g. cycling to work), sport (e.g. competitive football game), and exercise (e.g. planned gym activity) (Caspersen et al., 1985; DoH 2011).

The UK regularly publishes a set of evidence based PA guidelines for the population, reporting recommended levels of health enhancing PA for children, adults, and older adults (DoH, 2004, 2011). In the latest available guidelines, which were under consultation to be updated at the time of writing, it is recommended that adults and older adults accumulate at least: 150 minutes of moderate intensity PA per week, in bouts of at least 10 minutes; or at least 75 minutes of vigorous intensity PA per week; or a combination of both, also known as moderate-vigorous intensity PA (MVPA). In addition, adults should undertake activities which strengthen muscles twice per week and minimise time spent sedentary (DoH, 2011). Moderate intensity PA is defined as activities that require "an amount of effort and noticeably accelerates the heart rate, e.g. brisk walking… that is between 3 and 6 METs [metabolic equivalents]" (p. 54). Vigorous intensity activity is defined as activities that require a "large amount of effort, causes rapid breathing and a substantial increase in heart rate, e.g. running... that is above 6 METs" (p. 54). In this context, sedentary behaviour is defined as "activities that do not increase energy expenditure much above resting levels" such as "sitting, lying down and sleeping" (DoH, 2011, p. 54). These guidelines and applied definitions are the same as those recommended by other developed countries such as Australia (Brown, Bauman, Bull, & Burton, 2012), Canada (Canadian Society for Exercise Physiology, 2011), and the United States of America (US Department of Health and Human Services, 2008), all of which are based on the WHO's (2010) PA recommendations. These guidelines have been embedded into policy and PA promotion practices, to indicate that these recommendations are the minimum dose of PA required in order to improve health (Warburton & Bredin, 2017; Woods & Mutrie, 2012). However, they have not been uncritically accepted, especially as population PA levels have not substantially changed since their introduction (Weed, 2016). There is an argument for a greater public health impact by reducing the recommendations, for example, to 60 minutes of moderate intensity PA per week, based on evidence that this is a sufficient level of health enhancing PA and perhaps more achievable by the general population (Weed, 2016).

In contrast, there is no generally agreed upon definition of physical inactivity. The WHO (2010, p. 53) defines physical inactivity simply as "an absence of physical activity or exercise", whereas PHE (2014a, p. 4) defines it as "less than 30 minutes of physical activity a week". In epidemiological research, there is a consensus that physical inactivity is a PA level that is insufficient to meet the current PA guidelines, such as less than 30 minutes of moderate intensity PA, or 15 minutes of vigorous intensity PA per day (Dumith, Hallal, Reis, & Kohl III, 2011; Hallal et al., 2011; Lee et al., 2012).

Despite the lack of consensus on its definition, physical inactivity is distinct from sedentary behaviour, although the two terms have often been used interchangeably in the literature (Tremblay et al., 2017). In order to promote clarity, the Sedentary Behaviour Research Network (2012) have recommeded that future research adopts a consistent definition of sedentary behaviour, and defines it as "as any waking behaviour characterized by an energy expenditure ≤1.5 METs while in a sitting or reclining posture" (p. 540). This is distinctly different from their stance on the definition of physical inactivity, which they define as "an insufficient physical activity level to meet present physical activity recommendations" (Tremblay et al., 2017, p. 9).

1.4. Physical activity levels

Physical inactivity has become a major public health challenge worldwide (Guthold, Stevens, Riley, & Bull, 2018), with a global decline in all domains of PA (e.g. leisure, occupation, transportation, household PA) (Ng & Popkin, 2012). In the UK, between 1961 and 2005, total PA levels declined by 20% (Ng & Popkin, 2012). A similar pattern of transition has been observed in other developed countries such as China and the United States of America (Brownson, Boehmer, & Luke, 2004; Ng, Howard, Wang, Su, & Zhang, 2014; Ng & Popkin, 2012). This has been attributed to changes in activity at home and work, with the increasing availability of motor vehicles, occupational machinery, and home technologies that have replaced more labour intensive activities (Brownson et al., 2004; Ng et al., 2014; Ng & Popkin, 2012). Recent estimates confirmed that global physical inactivity is not improving, with more than a quarter of adults considered physically inactive (Guthold et al., 2018).

Considering the context of physical inactivity in England, results from the last Health Survey for England (2012) to measure PA reported that only 67% of men and 55% of women met the PA guidelines, with the percentage declining with increasing age, increasing BMI, and declining SES (Health and Social Care Information Centre [HSCIC], 2013). In addition, one in four adults were completely inactive, achieving less than 30 minutes of PA a week (HSCIC, 2013). However, these figures were based on self-reported PA. In a previous Health Survey for England (2008) that measured PA using accelerometers, it was estimated that only 6% of men and 4% of women met the recommended PA guidelines (HSCIC, 2009).

More recently, in 2017, Sport England published results of the first ‘Active Lives Survey’, a new survey to measure sport and PA across England, replacing the previous ‘Active People Survey’ (Sport England, 2017). The survey aims to provide self-reported PA data from a random sample of ~198,250 adults each year, distributed across English local authorities. The results of the first cohort (November 2015 – November 2016), which were generalised to the English population, reported that 25.6% were engaged in less than 30 minutes of moderate intensity PA per week (inactive), 13.7% were engaged in 30-149 minutes of moderate intensity PA per week (fairly active), and 60.7% were engaged in ≥150 minutes of moderate intensity PA per week (active). Further analysis showed gender, age, disability and SE differences, with those who were female, aged 75+ years, disabled, or from lower SE groups most likely to be inactive. Results from the most recent Active Lives Survey showed that national levels of PA had not changed, and that differences in PA were still apparent between groups (Sport England, 2018).

A limitation of the above evidence is that it is predominantly based on self-reported PA. The disparity, for example, between self-reported and accelerometer-measured PA in the Health Survey for England (2008) (HSCIC, 2009) raises the question of whether self-reported PA levels provide a true reflection of the impact of policy to improve population PA levels. It is also possible that self-reported population PA levels have incorrectly informed policy to increase PA levels. If PA levels are much lower than thought, significantly different approaches to PA promotion may have been required.

A further limitation is that the evidence is based on small survey samples representative of the general population (<200,000 adults), and then generalised to the adult population of England (>45 million adults). PA levels are also likely to be much lower in individuals with LTCs. However, compared to the general population, data on the PA levels of adults with LTCs is more limited (Brawner, Churilla, & Keteyian, 2016). Research has demonstrated low PA levels in specific disease groups such as arthritis (Dunlop et al., 2011; Kaptein & Bradley, 2012; Tierney, Fraser, & Kennedy, 2012), chronic obstructive pulmonary disease (COPD) (Saunders et al., 2016; Vorrink, Kort, Troosters, & Lammers, 2011), and cardiovascular disease (CVD) (Evenson, Butler, & Rosamond, 2014; Liu et al., 2000). However, few studies have analysed PA levels in samples representative of the general population, where direct comparisons of PA levels can be made between those with or without LTCs, between different LTCs, or the impact of having multiple LTCs.

In a recent study, Brawner et al. (2016) analysed responses to the National Health Interview Survey (n=36,697 adults), based on a representative sample of the general population of the United States of America. When considering the percentage of adults, with or without a LTC, meeting the recommended PA guidelines (≥150 minutes of moderate intensity PA per week), PA levels were lower in those with LTCs. The effect of having multiple health conditions was also explored, and compared to having no LTCs, the presence of each health condition was associated with a 17% lower odds of meeting the recommended PA guidelines.

These findings are similar to those of previous studies exploring the prevalence of meeting the recommended PA guidelines (Ashe et al., 2009), or being inactive, amongst adults with LTCs (Crespo et al., 1999; Kruger, Ham, & Sanker, 2008). Crespo et al. (1999) found a linear relationship between multiple health conditions and likelihood of being completely inactive, and all three studies found low prevalence of PA amongst similar LTCs (respiratory disease, diabetes, CVD) (Ashe et al., 2009; Crespo et al., 1999; Kruger et al., 2008).

1.5. Physical inactivity – the problem

In recent years, the evidence base for the impact of PA on health has grown to an extent that today, it is often uncritically accepted that PA is both a cause and cure for ill health (Henderson, Evans, Allen-Collinson, & Siriwardena, 2018; Williams & Gibson, 2018). Physical inactivity is now considered a major modifiable risk factor for the development and progression of multiple LTCs (Booth, Roberts, & Laye, 2012; Lee et al., 2012; Reiner et al., 2013). Globally, physical inactivity has been estimated to cause the development of 3.8% of dementia, 6% of coronary heart disease, 7% of type 2 diabetes, 10% of breast cancer, and 10% of colon cancer (Lee et al., 2012). Insufficient PA is also the fourth leading risk factor for global mortality after hypertension, tobacco smoking and high blood glucose (WHO, 2009). Therefore, PA as a modifiable risk factor is as important as other traditional risk factors such as hypertension and tobacco smoking (Lee et al., 2012; Stanaway et al., 2018; WHO, 2009).

There is a general consensus that any system in the body can benefit from PA, with additional benefits to independence, quality of life, and life expectancy (Chodzko-Zajko et al., 2009; Garber et al., 2011; O'Donovan et al., 2010; WHO, 2010). In 1994, Morris reviewed population studies of PA and coronary heart disease risk reduction, concluding that PA was “the best buy in public health” (Morris, 1994, p. 813). More than a decade later this message still appears to stand, with continued widely reported benefits such as: a 7% reduced risk of any cancer (Moore et al., 2016); 12-27% reduced risk of coronary heart disease (Sofi, Capalbo, Cesari, Abbate, & Gensini, 2008); 18% reduced risk of dementia (Blondell, Hammersley-Mather, &

Veerman, 2014); 20-30% reduced risk of type 2 diabetes (Aune, Norat, Leitzmann, Tonstad, & Vatten, 2015; Gill & Cooper, 2008) and 20-30% reduced risk of all-cause mortality (Ekelund et al., 2015; Samitz, Egger, & Zwahlen, 2011).

Furthermore, for those with existing LTCs, PA is important for the secondary prevention of further events and additional health conditions. Regular PA, for example, has been associated with lower mortality rates and complications associated with conditions such as CVD and diabetes (Colberg et al., 2016; Darden, Richardson, & Jackson, 2013).

The UK level of PA is therefore not sufficient to maintain good long-term health, placing individuals at heightened risk of developing LTCs that could have otherwise been potentially prevented by PA. Insufficient PA is also an economic burden in the UK, accounting annually for an estimated £7.42 billion towards health care expenditure and indirect costs to society (Centre for Economics and Business Research, 2015; Scarborough et al., 2011). If national levels of physical inactivity continue to rise, the burden on health and social care systems will become unsustainable, with reduced quality of life for individuals (PHE, 2014a).

1.6. Physical inactivity – global and national responses

Over the last decade, there has been an increasing effort to promote PA by both local and national government through to third sector organisations. Efforts have been strengthened due to trends indicating a progressive decline in population health, and the strength of the PA evidence base growing, but also by global initiatives to improve health and PA levels.

In 2010, after world-wide consultation with stakeholders, the Toronto charter for PA was published by the Global Advocacy Council for PA (GAPA) and the International Society for PA and Health (ISPAH). This was a global call for nations to commit to increasing levels of PA, by implementing four key actions based upon nine guiding principles (e.g. introduce and implement new policy, reorient services and funding in favour of PA). A complementary document was published, detailing seven suggested best investments for PA, such as transport policies, urban design, and PA integrated into health care systems (GAPA & ISPAH, 2011). The Toronto charter for PA was followed by the Bangkok declaration on PA for global health and sustainable development in 2016, a renewed call by ISPAH for nations to take action on physical inactivity (ISPAH, 2016).

In England, in 2014, PHE published a national PA framework titled ‘Everybody Active, Every Day’ (PHE, 2014a). It emphasised the need to take more action in four areas: 1) creating an active society through change in attitude; 2) activating networks of public-facing professionals to promote PA; 3) creating active environments, and; 4) scaling up interventions that work. Action in these areas was deemed necessary to create a cultural change in society towards becoming physically active, and to create new resources whilst making use of existing ones, to make PA accessible to all individuals. Two years later, revisiting the framework, PHE (2017) reported that the proportion of the population achieving ≥150 minutes of moderate intensity PA had increased by 1%, equivalent to half a million people.

In 2015, this national framework was followed by the publication of ‘Sporting Future’ by the Department for Culture, Media & Sport, the first national strategy for sport and PA published in over a decade. In 2016, Sport England then published ‘Towards an Active Nation’, a strategy for 2016-21 to deliver the five outcomes set out in ‘Sporting Future’: improved physical health, mental health, individual development, social and community development and economic development (Sport England, 2016).

In 2018, the WHO published a global action plan for PA, to reduce the global prevalence of physical inactivity by 15% by 2030. It set out four strategic objectives, achievable through 20 recommended policy actions. These objectives overlapped with PHE's national framework for PA (PHE, 2014a), to create active societies, active environments, active people, and active systems.

Throughout these policy developments, guidance has been published to support implementation of policy actions, on areas such as active travel (PHE, 2016), workplace activity (PHE, 2015a), the built environment (National Institute for Clinical Excellence [NICE], 2008), and primary care based interventions such as brief advice (NICE, 2013) and exercise referral schemes (ERSs) (NICE, 2014a).

The latter, ERSs, are the focus of this thesis and will be described in more detail in section 2.2. As an intervention that uses health care professionals to promote PA, by a system of referral to a community-based PA provider, ERSs fall under global and national recommended policy actions (PHE, 2014a; WHO, 2018).

1.7. Context of the PhD

This PhD was an evaluation of an existing ERS in Tameside, Greater Manchester, known as the 'Live Active' ERS. Tameside is an area with evident health inequalities (PHE, 2018b). It resides in the top 20% of most deprived districts in England, with higher than national average levels of smoking, physical inactivity, and excess weight (PHE, 2018b). At the time of writing, approximately 30.8% of adults in Tameside were considered inactive, achieving less than 30 minutes of moderate intensity PA per week (Sport England, 2018). The 'Live Active' scheme commenced in July 2015, the evaluation began shortly after, and concluded in 2018. 'Live Active' is a 6-month ERS offering dedicated contact time and specialist support from exercise professionals, to facilitate PA behaviour change in Tameside residents who are inactive and/or sedentary, and have a LTC or significant health risk factors. The scheme offers support for 6-months, and subsidised access to leisure facilities for 12-months. An overview and critical discussion of the ‘Live Active’ scheme, including the context of its location, is provided in chapter 3.

1.8. Aims of the PhD

The overall aim of this research was to evaluate the effectiveness of the ‘Live Active’ scheme at increasing PA levels and improving health outcomes in its participants. It was conducted in two phases. The first phase involved: identifying the requirements of ‘Live Active’ stakeholders; a review of the literature; observation of the scheme, and; quantitative analysis of secondary data collected during the first year of the scheme. Collectively, the outcomes of these tasks informed the development of four research studies that were conducted in phase two, the primary objectives of which are listed below. The objectives of studies 1 and 2 were based on the requests of ‘Live Active’ stakeholders as part of this evaluation, and the results of the literature review. Studies 3 and 4 examined novel research questions that emerged from the findings of the literature review, observations of the scheme, and the results of secondary analyses of data conducted during phase one.

The key research objectives for the PhD were:

- What is the level of uptake and adherence to the scheme, what individual characteristics can predict these outcomes, and why do participants non-uptake or dropout?
- What are the medium to long-term PA behaviour and health outcomes of participation, and do they differ by primary referral condition?
- What are the wider benefits of participation in the scheme?
- What are the short-term changes in PA and sedentary behaviour when measured using accelerometers, and how do these outcomes compare with

self-reported PA and sedentary behaviour?

1.9. Structure of the thesis

An overview of the structure of the thesis is provided below on page 14. Chapter 2 begins with a literature review of ERSs, exploring the history of ERSs, the evidence for ERSs as a public health intervention, and the limitations of ERS evaluations to date. Chapter 3 provides an overview and critical discussion of the ‘Live Active’ scheme. Chapter 4 is a methodology chapter, and will detail the overall research framework for this thesis, the aims and objectives of each research study, and a critical discussion of the research approach taken. Chapters 5, 6, 7 and 8 are dedicated to each individual research study. Chapter 9 is a discussion, which summarises the findings of each study, discusses the strengths and limitations of the research, and the implications for future research, practice and policy, with recommendations for each.

An overview of the structure of the thesis, with a summary of each chapter:

Chapter 2. Literature review of ERSs
A review and critical discussion of the literature surrounding the history of ERSs, the evidence for the impact of ERSs, and the limitations of the literature.
Chapter 3. Overview of the ‘Live Active’ ERS
An overview and critical discussion of the research setting, including the rationale for the scheme, a summary of the referral process, and limitations of the outcomes collected.
Chapter 4. Methodology
A summary of the methodology used for this thesis, including the overall research framework, and critical discussion of the research approach adopted.
Chapter 5. Quantitative study to evaluate uptake and adherence to the ‘Live Active’ ERS
Aim: Evaluate uptake and adherence to the ERS. Method: Quantitative analysis of uptake and adherence for the period July 2015-July 2017, including multivariate analyses to identify predictors of these outcomes.
Chapter 6. Quantitative study to evaluate the routine outcomes of the ‘Live Active’ ERS
Aim: Evaluate the impact of participation in the ERS on PA behaviour, and a range of health-related outcomes. Method: Quantitative analysis of secondary data for the period July 2015-July 2017, using inferential statistics.
Chapter 7. Mixed methods study to explore the perceived benefits of participation in the ‘Live Active’ ERS
Aim: Understand the wider benefits of participation in the scheme in a more holistic manner. Method: Mixed methods using quantitative secondary data (feedback questionnaire), and qualitative investigation of the perceived benefits to participation, using semi-structured interviews.
Chapter 8. Quantitative study to identify accelerometer derived changes in PA and sedentary behaviour, and to compare agreement with self-reported outcomes
Aim: Explore the PA and sedentary behaviour of participants using accelerometer-based measurement, and compare agreement with self-report. Method: Quantitative analysis of change in accelerometer-measured PA and sedentary behaviour between weeks 1 and 12, and comparison with self-reported outcomes in a sample of participants.
Chapter 9. A discussion of the research undertaken in chapters 5-8
Synthesis of findings, strengths and limitations of the research, implications and recommendations for future research, practice and policy, and final conclusions.

Chapter 2 Literature review of ERSs

2.1. Introduction

The aim of this chapter is to set the scene for the historical development of ERSs as a public health intervention, and current issues in evaluating such schemes. It begins by discussing how the literature was searched, followed by discussions of what an ERS is, the historical development of ERSs and the evidence base, and methodological issues in the evaluation of such complex interventions. A synthesised review of the current evidence for ERSs is then provided, followed by a review of current issues in the evaluation of ERSs. Key themes are discussed such as: the range of routine outcomes currently collected by schemes; the diversity of ERS populations and the implications for an evaluation; the potential benefits to participation in an ERS, and; methods of PA measurement. A synthesised review of these themes is provided, with an in-depth review provided later throughout chapters 5-8.

The synthesised review provided in this chapter, and the in-depth reviews provided in chapters 5-8, are based on a review of published research, focusing on interventions aligned with the ERS model. The search strategy used for this review is further detailed in appendix 1, which included a search of several electronic databases using key words to identify published and peer-reviewed research studies or reviews, editorials, responses, or conference-proceedings available as abstracts or short papers. Relevant grey literature was also sourced. Research studies of ERS interventions were included in the review if they met the following criteria, using the NICE (2014a) definition of an ERS: 1) referral of an individual who is sedentary and/or inactive with a LTC or significant health risk factor, by a healthcare professional, to a third-party PA service provider with the aim of achieving an increase in PA; 2) initial assessment of the individual by ERS staff and continued monitoring throughout the programme, and; 3) provision of a PA programme by the third party PA provider, tailored to the individual’s needs.

2.2. What is an ERS?

ERSs are a form of non-medical intervention or ‘social prescription’ (Thomson, Camic, & Chatterjee, 2015), based on a system of referral by a health care professional to a third party PA provider. They have been defined as interventions that “seek to increase someone's physical activity levels on the basis that physical activity has a range of positive health benefits” (NICE, 2014a, p. 7). They are usually commissioned by local authorities in the UK, for adults who are inactive and/or sedentary, and have existing health conditions or other risk factors (NICE, 2014a). ERSs orginated in the UK, and are still very much a UK phenomenon (Sowden & Raine, 2008). Despite the ERS model being implemented abroad in recent years (Arsenijevic & Groot, 2017), albeit with variations, the evidence base is laregely derived from UK ERS evaluations (Campbell et al., 2015; Pavey et al., 2011).

Since their development in the UK in the early 1990s (Pavey et al., 2011), ERSs have been known by various other names, including but not limited to GP ERSs, exercise on prescription schemes, PA referral schemes, and active lifestyle schemes (see for e.g. James et al., 2009; Sørensen, Sørensen, Skovgaard, Bredahl, & Puggaard, 2011; Wormald et al., 2006). The emergence of PA referral schemes reflects a transition from structured exercise class-based schemes rooted in exercise science paradigms, towards individualised PA programmes rooted in PA promotion paradigms (Dugdill, Graham, & McNair, 2005).

Guidelines for the delivery of ERSs in the UK are provided by NICE, an independent organisation responsible for producing evidence based guidance on health and social care to commissioners, practitioners and managers (NICE, 2018). These guidelines are largely ambiguous as to how an ERS should be designed and delivered, but state that a model of ERS should include: i) identification of eligible adults who have a motivation to become physically active, by healthcare professions, and invitation to be referred to an ERS; ii) invitation to the individual, by the ERS, to attend an initial consultation (usually in a local leisure centre), where a PA specialist makes an assessment of their needs and preferences, and; iii) provision of opportunities to become active, based on this assessment. This is usually followed by a number of follow-up consultations, where the PA specialist can assess change in PA and health status, and provide motivational support and adjustments to the PA programme where necessary (NICE, 2014a).

Such ambiguity in these guidelines has led to variations in programme structure and delivery between providers (Campbell et al., 2015; Oliver, Hanson, Lindsey, & Dodd-Reynolds, 2016; Pavey et al., 2011). ERSs therefore tend to differ on key characteristics such as: eligibility criteria – which health conditions are included or excluded; whether referrals are accepted from primary care organisations, secondary care organisations, or both; whether or not the scheme accepts self-referrals, and therefore whether or not referral depends upon coming into contact with a health care professional; the scheme's duration – usually between 8 weeks to 24 weeks of participation; the number of follow-up consultations provided; the degree to which PA options are subsidised, or are provided for free, and; the degree of tailoring of the PA programme to the individual. Therefore, to-date, over 600 interventions have operated under the ERS model in the UK (Pavey et al., 2011), but they have been highly heterogenous and therefore not strictly comparable.

Differences in characteristics such as the eligibility criteria (e.g. which health conditions are accepted), can create a post-code lottery in equality of access, given the usual commissioning of one ERS per local authority area. A person with osteoarthritis, for example, may not be eligible for the ERS in their locality, but could be eligible for the ERS in the next neighbouring locality. However, even in an area where an individual is eligible, referral depends on the opportunistic identification of the individual as inactive by a health care professional, and a partnership being in place between the health care organisation and the local ERS to allow the referral to be made (Campbell et al., 2015). Recent evidence suggests that knowledge of PA guidelines is low amongst GPs, with low levels of confidence in discussing PA with patients (Chatterjee, Chapman, Brannan, & Varney, 2017). However, it is also known that where health care professionals are aware of PA initiatives, disparities exist in referral practices, with some more likely to make referrals than others, reporting barriers such as time, uncertainty over eligibility, and preference for other health promotion services such as smoking cessation (Din, Moore, Murphy, Wilkinson, & Williams, 2015). Therefore, not all adults who could benefit from an ERS are referred, whether that be due to the characteristics of the ERS such as eligibility criteria and referral routes, the practices of the health care professional, or indeed first coming in to contact with a health care professional.

2.3. The history of ERSs

ERSs first emerged in the UK in the early 1990s, and were rapidly implemented at scale across the nation without an established evidence base for their effectiveness (Fox, Biddle, Edmunds, Bowler, & Killoran, 1997; Iliffe, Tai, Gould, Thorogood, & Hillsdon, 1994). Concerns about the lack of evidence were first published in 1994 (Iliffe et al., 1994), and again in subsequent years (see for e.g. Dugdill et al., 2005;

Fox et al., 1997; Riddoch, Puig-Ribera, & Cooper, 1998), but expansion of schemes continued, facilitated by government endorsement (Dugdill et al., 2005).

By the late 1990s, the evidence base for ERSs was small and equivocal about effectiveness (Stevens, Hillsdon, Thorogood, & McArdle, 1998; Taylor, Doust, & Webborn, 1998). However, in a need to be seen as actively reducing national levels of physical inactivity, the DoH described ERSs as “a cost-effective alternative to prescribing long-term medication” (DoH, 1999, p. 24), despite no apparent evidence base for overall effectiveness, or published research comparing the cost-effectiveness between ERSs and medication prescription. By the mid 2000s, the evidence base for ERSs had become no more conclusive (Dugdill et al., 2005; Morgan, 2005).

In the same year that NICE published an evidence review commissioned by the DoH, which concluded that there was insufficient evidence to recommend ERSs (NICE, 2006), the DoH reported, in a white paper, that ERSs had “often been very successful”, calling for continued expansion of schemes (DoH, 2006, p. 51). The review by NICE recommended that schemes should only be endorsed if they included a well-designed and controlled research evaluation, to improve the ERS evidence base (NICE, 2006). However, contradiction by the DoH continued, when they again called for the continuation of ERSs, clarifying that only ERSs that accepted referrals from individuals without any underlying health condition or risk factors should include a controlled research evaluation (DoH, 2007). This therefore made the evaluation recommendations by NICE (2006) irrelevant, since the eligibility criteria for ERSs dictates the presence of a health condition or risk factor.

At the time of writing, it was estimated that over 600 ERSs are in operation in the UK (Pavey et al., 2011), with the ERS model now being implemented abroad (see for e.g. Gallegos-Carrillo, García-Peña, Salmerón, Salgado-de-Snyder, & Lobelo, 2017; Pedersen, Helmer-Nielsen, Dieperink, & Østergaard, 2016a; Sørensen et al., 2011), despite an evidence base that is still equivocal (Campbell et al., 2015; Pavey et al., 2011). Upon review of the history of ERSs, Sowden and Raine (2008) have argued that evaluation of the effectiveness of ERSs in England is now unrealistic, because schemes have already been widely established in numbers, and the results of any evaluation are unlikely to cause ERSs to be withdrawn, or indeed change how they operate.

2.4. Evaluation frameworks

When evaluating public health programmes, it is necessary to take an evidence based approach to the evaluation (Hanson & Jones, 2017). This necessitates tools to facilitate evaluations that balance scientific rigour with the practicalities of public health interventions that operate in open and unpredictable environments (Hanson & Jones, 2017). Over recent decades, various guidance documents have been established to facilitate the development and evaluation of ERSs. However, they have been ambiguous in their content, and therefore open to interpretation, leading to variation between ERSs in models of delivery and evaluation (Oliver et al., 2016).

In 2001, the DoH published the National Quality Assurance Framework (NQAF) for ERSs (DoH, 2001). Providing guidance and recommendations for the running and evaluation of schemes, the guidance provided around the evaluation of schemes was ambiguous, with no specific recommendations for the selection of an evaluation design. This was in contrast to the Medical Research Council's (MRC) detailed framework for the design and evaluation of complex interventions, published in the previous year (MRC, 2000). This framework was in-depth, recommending a phasic approach to evaluation that takes in to account the nature of complex interventions, which are often not fully defined from the beginning, and may evolve over time through experience, making a randomised controlled trial (RCT) unsuitable at the outset, but perhaps more feasible at a later stage (MRC, 2000). However, this approach does not account for complex interventions such as ERSs, which may still continue to adapt and evolve beyond their initiation period, through mechanisms such as feedback loops (Moore et al., 2018; Shiell, Hawe, & Gold, 2008).

Several years later, the British Heart Foundation National Centre (2010) published guidance to be used in addition to the NQAF for ERSs (DoH, 2001). It included more in-depth discussion around different types of experimental and non-experimental evaluation designs, and recommended the use of a logic model to plan an evaluation. A pragmatic framework was presented, that advocated both process and outcome evaluation over a 12-month period. This framework was in alignment with updated guidelines from the MRC (2008) for the development and evaluation of complex interventions, which recommended process evaluation as a complementary addition to outcome evaluation.

Cavill, Roberts, & Rutter (2012) then published the standard evaluation framework (SEF) for PA interventions, the most comprehensive guidelines for the evaluation of ERSs to date. Although not a framework designed specifically to evaluate ERSs, but rather any intervention that aimed to increase PA, the NICE (2014a) ERS guidelines recommended that ERSs should not be commissioned unless they collected data in line with this framework. The framework contained a list of criteria considered necessary to conduct a comprehensive and robust evaluation. It contained both process and outcomes measures, as recommended by the MRC's guidelines at the time (2008), in addition to the most recent guidelines (MRC, 2015). The guidance for this framework discussed the importance of understanding the objectives of an intervention, which would determine the outcomes that should be measured. In addition to the mandatory measurement of PA, it was suggested that these outcomes could include physiological outcomes (e.g. BMI), and health-related quality of life (HRQoL).

2.5. The evaluation of complex interventions

The development of the ERS evidence base has to an extent been hindered by ambiguous evaluation guidelines, and a preference by researchers and policy makers for positivist methods of evaluation, such as RCTs, which fail to reflect the complex nature of exercise referral interventions (Dugdill et al., 2005; Oliver et al., 2016).

A complex intervention has been defined as “an intervention comprising multiple components which interact to produce change” (MRC, 2015, p. 8), “characterised by unpredictability, emergence and non-linear outcomes” (MRC, 2015, p. 19). This complexity extends to the number and difficulty of behaviour changes being targeted, the number of groups and organisational levels being targeted, the range of intended outcomes, the required behaviours of those delivering the intervention, the degree of flexibility or tailoring of the intervention that is allowed, and the impact of context on implementation of the intervention (MRC, 2015). Therefore, ERSs can be considered to be complex interventions (Dugdill et al., 2005), as they have: heterogeneous delivery models that are open to change; heterogeneous participants that are referred to them; the involvement of multiple stakeholders from different organisations such as primary and secondary care organisations and local authorities; non-linear and unpredictable changes in behaviour and outcomes, and; they are often designed according to the local context.

Positivist methods such as RCTs are traditionally used to understand cause and effect in clinical interventions, hence their popularity in the medical and pharmaceutical literature (Deaton & Cartwright, 2018; Rutter et al., 2017). They aim to maximise internal validity through randomisation and control, although this is often at the expense of ecological validity, which is the extent to which the research represents real world practice (Denscombe, 2010; Gidlow, Johnston, Crone, & James, 2008; Ng & Colombani, 2015). Historically, they have ranked high up the evidence hierarchy, considered by many to be the gold standard for producing robust evidence about an intervention (Akobeng, 2005; Moore et al., 2015; Ng & Colombani, 2015).

The role of a method used to understand cause and effect in controlled clinical interventions, within evaluations of public health interventions, has been questioned (Pettman et al., 2012; Rutter et al., 2017). RCTs attempt to factor out the complex elements of an intervention, stripping away all of its layers, to identify its impact under controlled conditions (Dugdill et al., 2005; Rutter et al., 2017). However, public health interventions usually operate in open systems that can be resistant to control (Allmark, 2015), where multiple contextual factors may interact with the intervention (Craig et al., 2008; Victora, Habicht, & Bryce, 2004). Therefore, RCTs can show that an intervention works in a controlled environment, but not in the real world in which it will be delivered (Allmark, 2015; Pettman et al., 2012). Furthermore, RCTs can only provide simplistic conclusions about whether an intervention is effective or not based on the observed results, with limited opportunity for learning (Deaton & Cartwright, 2018; Dugdill et al., 2005). There is no actionable information for stakeholders, to understand why an intervention works, or why it does not work and how it can be adapted (Deaton & Cartwright, 2018; Dugdill et al., 2005).

The place of RCTs in the evaluation of ERSs has been argued based on these criticisms (e.g. Dugdill et al., 2005; Oliver et al., 2016). Such designs require manipulation of referral processes, providing an evaluation of controlled practice as opposed to real life practice (Gidlow et al., 2008). Furthermore, a control group is usually formed by placing a proportion of participants on a waiting list for the ERS (up to a year), which has ethical considerations as this restricts access to a service that would potentially be beneficial to their health status (Dugdill et al., 2005; Gidlow et al., 2008). This is also an important consideration because health care professionals have reported that they would be less likely to refer a person to an ERS that was part of a RCT, reporting discomfort at the withholding of a service from half of referrals (Din et al., 2015). Even when participants are referred and assigned to a control group as part of a RCT, they may still be referred to alternative services by healthcare professionals during the study period, such as physiotherapy, which may lead to PA and invalidate the original study (Sowden & Raine, 2008). This is one factor that can contribute to control group contamination, whereby meaningful increases in PA in the control group are observed (Hertogh, Schuit, Peeters, & Monninkhof, 2010; Waters, Reeves, Fjeldsoe, & Eakin, 2012). Furthermore, in PA interventions, patient satisfaction with randomisation to a control group can be low, with participation seen as less useful and rewarding to the participant, which is a further factor that can contribute to control group contamination (Courneya et al., 2013).

Despite these challenges, promotion of RCTs to evaluate ERSs is exacerbated (Oliver et al., 2016) by PHE’s adoption of ‘NESTA’ standards (Puttick & Ludlow, 2013), to identify promising community-based PA interventions that can be scaled across the nation (PHE, 2014b). To-date, two open calls have been made for organisations, groups or individuals delivering PA interventions, to submit details on delivery and outcomes (PHE, 2014b; PHE, 2018c). The interventions have been rated as emerging, developing, promising or proven, based on NESTA's academic standards of evidence (Puttick & Ludlow, 2013). According to these standards, only interventions that have included a control group can qualify as proven practice, irrespective of their outcomes (PHE, 2014b).

This growing pressure for community-based PA interventions to conduct RCTs in order to demonstrate their impact overlooks the skills, logistics, and finances required to carry out such an evaluation (Gidlow et al., 2008). Furthermore, many community-based PA interventions such as ERSs are designed according to the local context (Oliver et al., 2016). Therefore, it cannot be assumed that a successful ERS will have the same success in a different community when scaled-up (Campbell et al., 2015). This raises the question of whether identifying and replicating successful, complex interventions, which are locally tailored, is always a relevant objective.

Criticisms of methodology are common across several fields of public health, where rigorous systematic reviews almost exclusively select RCTs for inclusion, to generate consensus on the effectiveness of interventions, and subsequently inform policy (Pettman et al., 2012; Rutter et al., 2017). Rigorous criteria for inclusion often leads to systematic reviews excluding high quality yet uncontrolled interventions, with uncontrolled evaluations often failing to be published or widely disseminated (Egger, Dickersin, & Smith, 2008; Pettman et al., 2012). Whilst rigour in methodology of studies is necessary, strict criteria for inclusion in systematic reviews increases the likelihood that important information is not disseminated if it is not based on a RCT (Gidlow, Johnston, Crone, & James, 2005; Gidlow et al., 2008; Pettman et al., 2012). In the field of ERSs, many studies have been excluded from systematic reviews, with the latest systematic review of clinical effectiveness of ERSs based on only eight

RCTs (Campbell et al., 2015). This precludes any insight to be gained from observational studies examining the impact of ERSs in real world practice settings, and contributes to the lack of consensus on the impact of ERSs (Gidlow et al., 2008; Oliver et al., 2016).

Within the field of ERSs, choice of research methodology has rarely been debated (see for e.g. Dugdill et al., 2005; Gidlow et al., 2008). However, arguments exist for use of a greater range of research methods, including not just RCTs, but high quality observational and qualitative studies in real world settings, to enhance understanding of ERSs, and to ensure that policy is informed by the best available evidence (Gidlow et al., 2005, 2008). In defence of RCTs, Moore et al. (2015) argue that complex interventions such as ERSs can include a process evaluation to interpret the findings of the RCT. However, it is also argued that high quality, uncontrolled quantitative approaches, can provide insight in to a range of areas including patterns in referral, attendance and outcomes, whilst qualitative approaches can offer insights in to why these patterns exist, and why certain ERSs are successful (Dugdill et al., 2005; Gidlow et al., 2008).

2.6. Current evidence base for ERSs

It is anticipated that through ERS participation, individuals will increase their PA levels and maintain this in the long term beyond the scheme’s duration, and as a consequence obtain health benefits (NICE, 2014a). However, systematic reviews of ERS evaluations do not show conclusive evidence that these objectives are met (Campbell et al., 2015; Pavey et al., 2011).

Thus far, change in self-reported PA has been the primary outcome of importance for evaluations of clinical-effectiveness (Campbell et al., 2015; Pavey et al., 2011; Williams, Hendry, France, Lewis, & Wilkinson, 2007) and cost-effectiveness (Anokye et al., 2011; Trueman & Anokye, 2013). To date, systematic reviews have found no strong evidence for an improvement in the proportion of people achieving ≥90-150 minutes of PA per week following ERS participation (Campbell et al., 2015; Pavey et al., 2011). Physical health outcomes have also been a primary indicator of success, with inconsistent evidence for an improvement in outcomes such as BMI and BP (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). Few studies examine parameters beyond this such as glycaemic control or physical fitness (for exceptions see: Isaacs et al., 2007; Sørensen, Kragstrup, Skovgaard, & Puggaard, 2008). There are fewer studies still that measure changes in psychological health and wellbeing, and where these exist they have tended not to show any significant impact of the ERS intervention (Campbell et al., 2015; Pavey et al., 2011) with the exception of a weak short-term reduction in depression level (Pavey et al., 2011).

However, these conclusions are based on the findings of systematic reviews, the limitations of which have been discussed in section 2.5. Furthermore, the latest available systematic review of ERS effectiveness was based on eight RCTs (Campbell et al., 2015), representing a small fraction of the 600+ schemes in operation across the UK at the time (Pavey et al., 2011). At the time of the last systematic review (Campbell et al., 2015), it is estimated that there were an additional ≥10 quantitative evaluations of ERS impact on PA and health (e.g. Lee et al., 2009; Lord & Green, 1995; Ward et al., 2010). However, these are also schemes that are highly heterogenous in their design and delivery (NICE, 2014a), as previously discussed, leading systematic reviews to provide a consensus on the impact of interventions which are not strictly comparable. It has therefore been argued that systematic reviews have provided an unfair assessment of the potential of ERSs (Beck et al., 2016).

2.7. Current issues in the evaluation of ERSs

Generation of a consensus on the impact of ERSs has been limited by ambiguous ERS guidelines and therefore heterogenous models of delivery, as well as limitations in how evidence of impact is selected and disseminated through systematic reviews. However, the evidence base is also limited by the complexity of evaluating an ERS.

PA is a complex behaviour (Adams & White, 2003), and therefore the evaluation of any intervention to change this behaviour can also be complex (Dugdill et al., 2005). Previous evaluations have been limited by not recognising this complexity, using methodologies that cannot fully answer the question of which schemes work, and for which particular groups (Oliver et al., 2016). Furthermore, evaluation guidelines to date have been ambiguous (Campbell et al., 2015; Pavey et al., 2011), with no strong evaluation framework available until more recently (the SEF – Cavill et al., 2012). Therefore, those evaluating ERSs have not been guided to consider the complexities of evaluation, or which methodologies may be appropriate. Some of the complexities underlying ERS evaluations, and their impact on reported outcomes, will be summarised in the following sections.

2.8. Evaluations of ERS – outcomes and problems

2.8.1. Process measures – evidence and problems

Poor uptake and adherence to ERSs is an established problem, evidenced by reoccurring low figures and a growing body of literature aimed at understanding how to facilitate better uptake and adherence (Morgan et al., 2016; Pavey et al., 2012). Uptake has been defined consistently across the literature as the proportion of individuals referred to an ERS who attend an initial consultation (Campbell et al., 2015; Pavey et al., 2012). Adherence, however, has not been consistently defined, and has therefore been expressed using different definitions, as will be discussed further in chapter 5. Variation in applied definitions has made interpretation of adherence figures from different evaluations difficult (Pavey et al., 2012; Tobi, Estacio, Yu, Renton, & Foster, 2012).

To date, the evidence regarding the predictors of, and reasons for engagement is equivocal with gaps in knowledge remaining (NICE, 2014a). Previous quantitative research has focused on predictors of uptake and adherence (Pavey et al., 2012), limited by the inclusion of a small number of characteristics such as age and gender (Pavey et al., 2012), with less known about the impact of variables such as SES (Hanson, Allin, Ellis, & Dodd-Reynolds, 2013), ethnicity (Mills, Crone, James, & Johnston, 2012) and disability status (Kelly et al., 2017). This, however, is partly explained by lack of available data due to poor participant profiling by ERSs (Hanson et al., 2013; Sowden & Raine, 2008; Tobi et al., 2012). Qualitative research has also sought to understand the reasons for uptake and adherence, exploring factors such as the quality of leisure facilities and the approach of ERS staff (Gidlow et al., 2005; Morgan et al., 2016). However, these approaches focus on factors associated with uptake and adherence, with limited knowledge of the reasons for non-uptake and dropout (Morgan et al., 2016).

Awareness of engagement levels, and reasons for non-uptake and dropout is important for stakeholders, to highlight areas of practice that may be adapted to facilitate higher engagement (Morgan et al., 2016; NICE, 2014a). Based on current evidence, in order to improve practices that facilitate engagement there is a need for future research to: generate a consensus on the definitions of uptake and adherence; improve participant profiling in ERSs in order to broaden the range of data available for prediction models of uptake and adherence, and; to understand the reasons why participants do not uptake, or dropout from an ERS. Uptake and adherence to ERSs is further discussed in chapter 5, including further discussion of uptake and adherence levels, factors associated with these outcomes, and the gaps in understanding.

2.8.2. Outcome measures – analysing a diverse population

Substantial heterogeneity exists in the delivery of ERSs, and the individuals that are referred to them, which may contribute to the equivocal findings of previous research (Campbell et al., 2015; Pavey et al., 2011). ERSs vary in their eligibility criteria, but if one considers how diverse the population can be within a single ERS, such as the variety of age groups and medical conditions of participants, finding non-significant outcomes at the population level can be reasonably anticipated, with improvements masked by population heterogeneity (Pavey et al., 2011; Rowley, Mann, Steele, Horton, & Jimenez, 2018). Whilst population heterogeneity has been acknowledged by systematic reviews, such sub-group analysis is rare (see for e.g. Murphy et al., 2012), leaving researchers unable to draw any consensus on whether certain groups benefit in different ways (Campbell et al., 2015; Pavey et al., 2011). Emerging research that has included sub-group analyses does suggest that participants referred for different health conditions may benefit from participation to different extents (Murphy et al., 2012; Rowley et al., 2018). Therefore, it is important for future evaluations to take population heterogeneity into account, by using methods such as sub-group analysis, to further understand which groups an ERS works for, and in what way (Anokye et al., 2011; Pavey et al., 2011). In doing so, areas for improvement in practice for particular groups may be identified, or referrals targetted for groups in which the benefits of participation are shown to be greatest (Anokye et al., 2011). The reported outcomes of ERSs are further discussed in chapter 6, including a discussion of population heterogeneity and gaps in knowledge.

2.8.3. Outcome measures – validity of physical activity measurement

Change in PA has been considered the primary outcome of ERS clinical and cost-effectiveness (Campbell et al., 2015; Pavey et al., 2011). Therefore, accurate measurement of PA is fundamental. However, there appears to have been no critical discussion about how PA should be measured in ERSs. This may have contributed to the variation in PA measures used in ERSs, and therefore how PA is reported, causing difficulty for researchers who wish to compare the effectiveness of different ERSs (Campbell et al., 2015; Pavey et al., 2011).

Historically, PA levels in ERS participants have been measured using self-report methods such as questionnaires (Campbell et al., 2015; Pavey et al., 2011). These methods have advantages, in that they are a low-cost and time efficient solution to gaining relatively in-depth data from participants (Haskell, 2012; Strath et al., 2013). However, as a self-report instrument, they are also subject to bias, which reduces their validity (Haskell, 2012; Helmerhorst, Brage, Warren, Besson, & Ekelund, 2012; Strath et al., 2013). A growing number of researchers have recommended that future ERS evaluations also measure PA using device-based methods such as accelerometers (Campbell et al., 2015; Hanson et al., 2013; Isaacs et al., 2007; Pavey et al., 2011), which are argued to measure PA more accurately than self-report (Hills, Mokhtar, & Byrne, 2014; Reilly et al., 2008; Trost & O'Neil, 2014). It has also been suggested that both methods are used where feasible to provide a complete measurement of PA (Haskell, 2012; Pedišić & Bauman, 2015). Therefore, with little research having focused on device-based measurement of PA in ERSs, there is an opportunity to implement accelerometers as a complementary measure to PA questionnaires in ERSs. A more in-depth critical discussion of PA measurement is presented in chapter 8, including methods of PA measurement in ERSs, and the advantages and limitations of different methods.

2.8.4. Outcome measures – what should be measured and reported?

Researchers have argued that ERS evaluations have been predominated by positivist research methods, which cannot embrace the complexity of real world practice (Oliver et al., 2016), or capture the holistic worth of participation (Crone, Smith, & Gough, 2005; Dugdill et al., 2005). The aim of an ERS is “to increase someone's physical activity levels on the basis that physical activity has a range of positive health benefits” (NICE, 2014a, p. 7). In this context, no definition of health is provided. However, it is known that interventions targeting behaviours such as PA are likely to have unintentional ripple effects on multiple outcomes beyond their specific target, spanning physical, mental, and social health (Smith & Petticrew, 2010; Wilson, 2015). The effectiveness of an intervention and its return on investment can therefore be easily underestimated if there is no attempt to capture these impacts (Wilson, 2015). However, ERS evaluations to date have predominantly focused on measuring change in PA and a small number of physiological outcomes such as BMI and BP (Campbell et al., 2015; Pavey et al., 2011). If these variables are not improved then the scheme is deemed ineffective, when in fact participants may have experienced other benefits not captured by the evaluation that would represent effectiveness. Qualitative research, for example, does suggest that participants benefit in a variety of ways, reporting important benefits such as changes to medication use (Queen, Crone, & Parker, 2015), healthcare utilisation (Queen et al., 2015), and other lifestyle behaviours such as dietary improvement (Wormald et al., 2006).

Failure to appreciate the full spectrum of potential benefits to participation, by continuing to conduct quantitative evaluations with limited outcome measures, may prevent researchers from conducting representative evaluations (Dugdill et al., 2005; Mills et al., 2012) and underestimate impact and return on investment (Wilson, 2015). Therefore, for more than decade, researchers have called for more pluralistic evaluations of ERSs, using methods such as qualitative enquiry, to understand how individuals benefit from participation (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998). However, there has been little progress to date, with few qualitative explorations published (e.g. Mills et al., 2012; Queen et al., 2015), and this is therefore an area for continued research. A more in-depth critical discussion of the potential benefits to participation in an ERS is presented in chapter 7, including a discussion of the alternative benefits reported through qualitative research, and the relevance of these benefits.

2.9. Summary

Physical inactivity is a major modifiable risk factor for the development and progression of many LTCs (Booth et al., 2012; Lee et al., 2012; Reiner et al., 2013). This is an important public health concern, as PA levels in the UK are insufficient to maintain good long-term health (HSCIC, 2013; Sport England, 2018). In the UK, ERSs have become one of the most common approaches to increasing PA levels in inactive or sedentary adults with health conditions or risk factors (Campbell et al., 2015; Pavey et al., 2011).

Systematic reviews have found no strong evidence for an improvement in long-term PA levels, physical health, or mental wellbeing after participation in an ERS (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). Engagement with ERSs is also often reported to be low (Morgan et al., 2016; Pavey et al., 2012).

However, a myriad of limitations can be encountered when attempting to evaluate an

ERS, which may explain the equivocal findings to-date. Substantial limitations of evaluations include, but are not limited to: heterogeneity in the nature of ERSs offered, and the individuals that are referred to them; predominant use of self-reporting instruments to capture key outcomes such as PA levels, and; a narrow focus of evaluations on a select few physiological and psychological outcomes, overlooking other adjunct but important benefits to participation that may represent effectiveness (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). These limitations will be further discussed in chapters 5-8.

Chapter 3 Overview and critical discussion of the ‘Live Active’ scheme

3.1. Introduction

The purpose of this chapter is to provide an overview of the setting in which this research took place, including the rationale for the scheme, and its delivery processes. This transparency should provide an exact understanding of what was delivered, and can be taken in to account when comparing this intervention with any other ERS intervention.

3.2. The ‘Live Active’ scheme

‘Live Active’ was a 6-month ERS, designed to provide strong and sustainable pathways into PA for Tameside residents who were sedentary or physically inactive and had a LTC or significant health risk factor. Once referred by a health care professional, participants were supported by an exercise professional through a PA programme tailored to their needs and preferences, with four routine contact points for support. The scheme aimed to engage 3500 participants over a three-year period, and offered participants subsidised access to its leisure facilities for 12 months. It was funded by Sport England and Tameside Public Health as part of the ‘Get Healthy, Get Active’ project; a range of research pilot projects developed to improve the evidence base around interventions that reduce levels of physical inactivity. The ‘Live Active’ scheme continues to operate at the time of writing.

‘Live Active’ was one of several schemes run by Active Tameside at the time of writing, to increase the PA levels of the local population. Active Tameside, formerly Tameside Sports Trust, is a ‘not for profit’ charity which operates Tameside Council’s sport and leisure facilities.

3.2.1. Rationale for the scheme

Tameside is an area in the North-West of England with evident health inequalities (PHE, 2018b). At the time of the scheme's development (2015), Tameside had the 10th highest rate of premature mortality in England (Steel et al., 2018). Life expectancy at birth was lower than average, at 76.9 years for males and 80.3 years for females, compared to a national average of 79.4 years for males and 83.1 years for females (PHE, 2015b). The life expectancy gap for the most deprived areas of Tameside were 10.3 years lower for men and 9.3 years lower for women, compared to the least deprived areas of Tameside (PHE, 2015b). There were also high levels of LTCs such as CVD, respiratory disease, and diabetes; conditions that could be improved by an increase in PA (PHE, 2015b).

Considering levels of physical inactivity, Tameside was ranked 118/150 of local authorities in England (1: Highest PA level, 150 lowest PA level), with 32.81% of its adults physically inactive (0 minutes of MVPA per week) (UKactive, 2014). At this time, the 'Active People Survey' (Sport England, 2015) reported that 52.3% of inactive adults in Tameside would like to take part in sport and become more active. Therefore, Active Tameside proposed the development of a PA service that would provide an opportunity for inactive individuals with LTCs to become active, whilst improving their overall health and wellbeing, and outcomes associated with LTCs. There were no other ERSs operating in Tameside at this time.

The scheme proposed a particular focus on the wider determinants of poor health, with the aim of targeting referrals from hard to reach groups such as the socio-economically disadvantaged, given the higher than national average level of SE deprivation in Tameside. At the time of the scheme's development, Tameside was one of the 20% most deprived districts in England, and over 60% of residents resided in the two most deprived quintiles of deprivation (PHE, 2015b). Approximately 74.8% of residents were employed, earning £82.30 less per week than the national average. The most frequent reason for not working was long-term sickness (33.9%), which was significantly higher than the national average (21.8%) (ONS, 2016a).

Based on Tameside's healthy life expectancy figures of 57.4 years for males and 56.6 years for females (ONS, 2015), the scheme predicted that the majority of referrals would be for individuals aged between 26-65 years. In addition, Tameside had a higher than average proportion of people aged 40-69 years, with the proportion of people aged 55-64 years expected to grow by 2020 (ONS, 2016b).

Lastly, the scheme aimed for 7% of its engagement to be with black and minority ethnic (BME) groups. At the time of the scheme’s development, Tameside had a higher than average white British population of 88.5% compared to a national average of 79.8%. The largest ethnic groups were Asian (6.6%), Mixed (1.4%), and Black/African/Caribbean (0.8%) (ONS, 2011).

3.2.2. Inclusion criteria for referral

Inclusion criteria for referral to the scheme was based on the presence of a LTC or significant health risk factor, and was developed in conjunction with a lead GP in Tameside. The condition specific referral pathways were: established CVD; high risk of CVD; metabolic syndrome; diabetes; mental health issues; respiratory diseases; musculoskeletal conditions, and; BMI ≥30 kg/m^2. These pathways were selected based upon population prevalence in Tameside, and the evidence of the benefits of PA for these conditions. Participants referred for health conditions not under these pathways, but for whom PA support would be deemed beneficial and could be provided safely, were accepted and categorised under ‘other’. Cancer, present or in-remission, was introduced as a new referral pathway in April 2017. Exclusion criteria included contraindications to PA (American College of Sports Medicine, 2017). A full list of the specific inclusion pathways and criteria, and exclusion criteria are documented below in Table 3.1.

3.2.3. Staff

The scheme was overseen by a lead referral officer, with several exercise professionals, and an administrator. Exercise professionals held a minimum of a Register of Exercise Professionals (REPs) level 3 qualification in exercise referral. Staff delivering specialist classes such as back pain or cardiac classes were level 4 qualified in the respective specialities.

Table 3.1. Inclusion and exclusion criteria for the 'Live Active' scheme

Inclusion Criteria		Exclusion Criteria
• CVD ➢ Previous event, valve surgery, angioplasty/bypass graft ➢ Stable angina ➢ Stable, controlled heart failure ➢ Symptomatic peripheral vascular disease • Cerebrovascular disease ➢ Stroke ➢ Transient ischaemic attack • High risk of CVD ➢ QRISK2 score ≥10% (or equivalent) ➢ Elevated BP (≥140 systolic, and/or ≥90 diastolic) ➢ Hyperlipidaemia • Metabolic syndrome (≥3 of the following) ➢ Waist circumference >102 cm for men, >88 cm for women ➢ Fasting glucose ≥5.6 (mmol/L) ➢ HDL cholesterol <1.03 for men, <1.29 for women (mmol/L) ➢ Triglycerides ≥1.7 (mmol/L) ➢ BP (≥135 systolic, and/or ≥85 diastolic (mmHg))	• Diabetes ➢ Type 1, Type 2 or Pre-diabetes • Mental Health Issues ➢ I.e. Stress, anxiety, depression, bipolar, schizophrenia (stable) • Respiratory Disease ➢ Asthma ➢ Chronic obstructive pulmonary disease ➢ Emphysema ➢ Bronchiectasis • Obesity ➢ BMI ≥30 kg/m^2 • Musculoskeletal Health ➢ Chronic in nature (i.e. osteoporosis, arthritis, fibromyalgia, MSK pain, poor mobility or falls risk) ➢ Previous joint replacement • Cancer ➢ During or after treatment • Other ➢ I.e. Neurological (multiple sclerosis, epilepsy, Parkinson's disease), chronic fatigue syndrome, controlled endocrine disorder	• Physically inactive without a long term condition or significant risk factor. • Aged <18 years • Recent cardiovascular or cerebral event (<6 weeks) • Unstable or uncontrolled angina, hypertension, heart failure, tachycardia, arrhythmia • BP: systolic ≥180 and/or diastolic ≥100 (mmHg) • Acute myocarditis, pericarditis, endocarditis, pulmonary embolus or pulmonary infarction • Suspected or known dissecting aneurysm • Symptomatic severe aortic stenosis • Uncontrolled or poorly controlled severe asthma • Third trimester pregnancy

3.2.4. Referral process

Referrals to the scheme were made by health care professionals from a variety of primary and secondary care settings including general practices, physiotherapy clinics, weight management services, and mental health services (Figure 3.1, below). Therefore, referrals were made from a variety of health care professionals such as GPs, nurses, surgeons, physiotherapists, health trainers, dieticians, podiatrists and psychiatrists. The scheme did not accept self-referrals.

Prior to making a referral, referrers were asked to follow ERS guidelines (NICE, 2014a) by: identifying the individual as physically inactive; discussing the reasons for increasing PA, and; asking the individual if they are willing to make a PA behaviour change and be referred to the programme. If the individual agreed, a referral was made to the 'Live Active' scheme by submission of a referral form.

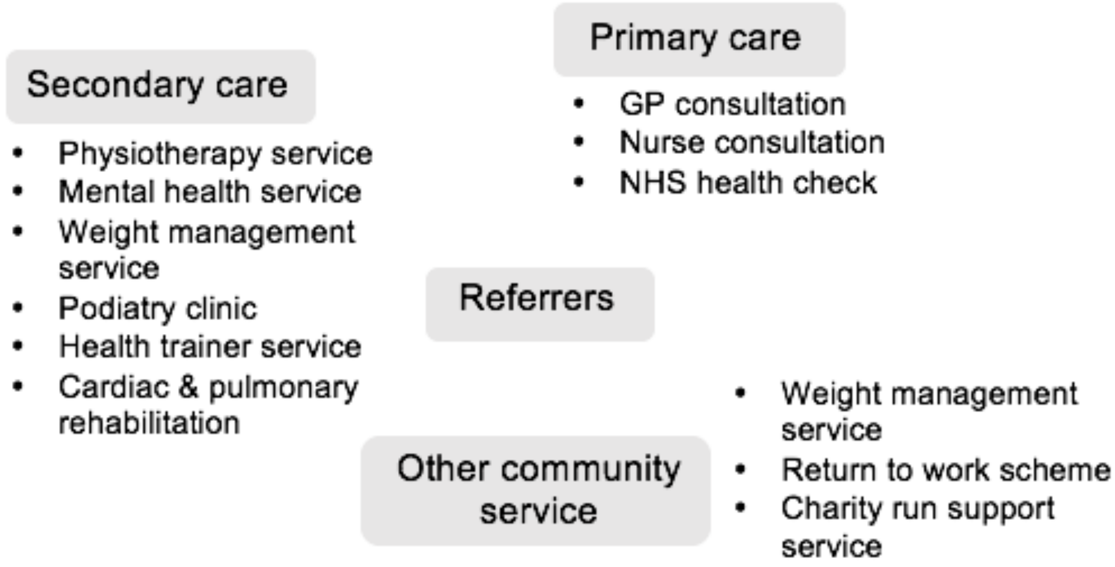


Figure 3.1. Examples of sources of referrals to the 'Live Active' scheme

Upon receipt of a referral, individuals were contacted by the 'Live Active' scheme by telephone to schedule an initial consultation. This telephone call also included an explanation of the referral pathway, and associated costs. If, after several attempts, the individual could not be contacted by telephone or text message, the contact details were checked with the referrer, and an invitation letter posted to the individual. In the event of no response, the individual was listed on the ERS database as ‘not participating’, indicating that the person did not book an initial consultation. If an individual declined the referral, or booked an initial consultation but cancelled this consultation and chose not to participate in the scheme, or did not attend multiple initial consultations, they were also listed as ‘not participating’ and the individual’s reason for this decision was recorded. When an individual attended the initial consultation, they were listed as ‘participating’, indicating that they were actively participating in the scheme. If an individual then chose to leave the scheme before its completion, they were listed as ‘left early’ and the reason for this decision was recorded.

3.2.5. Referral pathway

An overview of the referral pathway is illustrated below in Figure 3.2. At week 1, participants attended an initial consultation lasting one hour, where an overview of the scheme was provided. Exercise professionals discussed the reason for referral with the participant, and the individual benefits of increasing PA. Participants were risk assessed using the Physical Activity Readiness Questionnaire (PAR-Q), and a baseline evaluation was completed, including: readiness to change (Transtheoretical Model (TTM) Stage of Change) (Prochaska & DiClemente, 1982), current PA level and sedentary behaviour (International Physical Activity Questionnaire Short Form (IPAQ-SF)) (IPAQ, 2005), HRQoL (EuroQol five dimensions three level questionnaire (EQ-5D-3L) and EuroQol five dimensions visual analogue scale (EQ-5D VAS)) (EuroQol Group, 1990), mental wellbeing (Warwick-Edinburgh Mental Well-being

Scale (WEMWBS)) (Tennant et al., 2007), BMI, BP, alcohol consumption, and smoking status. These outcomes are critically discussed in section 3.5. Barriers to PA were identified, and solutions put in place to overcome these barriers. A range of short (12 weeks), medium (24 weeks) and long-term (52 weeks) goals were also agreed between the exercise professional and participant. In partnership, the exercise professional and participant agreed upon a programme of weekly physical activities based upon the participant's needs and preferences, taking in to account their short, medium and long-term goals.

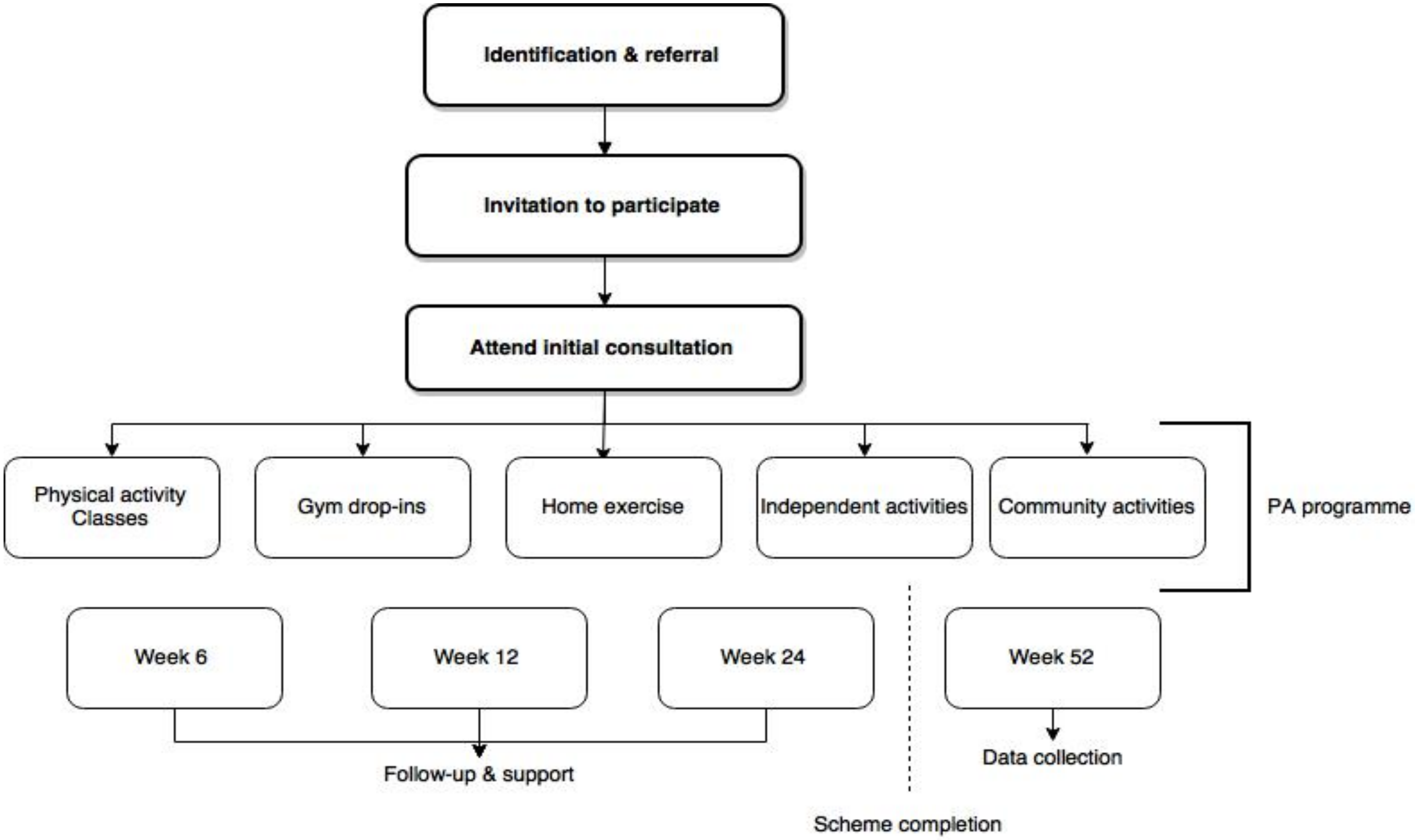


Figure 3.2. An overview of the referral pathway

A range of PA options were available for offer including but not limited to: swim and aqua based classes; specialist classes including circuit classes, chair based classes and clinical Pilates classes, which are provided exclusively for ‘Live Active’ participants; gym drop-ins, a designated two-hour slot where an ERS exercise professional was on the gym floor to assist ‘Live Active’ participants who wished to have supported gym-based PA; home work out programmes, with booklets and online videos to assist, and; community-based activities such as walking groups, tai chi and walking football. The scheme provided 12 months of subsidised access to Active Tameside’s leisure facilities and classes. The majority of community-based activities were free, or incurred a small pay-as-you-go fee (e.g. £2).

Participants then attended a follow-up consultation at weeks 6, 12, and 24 to: evaluate progression in PA behaviour; make appropriate changes to the PA programme if required, and; to offer support and motivation. These follow up consultations were also offered as telephone consultations, for those who did not wish to, or could not attend the consultation in person. At weeks 12 and 24, the baseline evaluation described above was also repeated. Participants also completed a ‘Live Active’ ERS feedback questionnaire (appendix 2), at weeks 12 or 24, which evaluated their satisfaction with the scheme thus far. At week 24, participants were discharged from the scheme, and no additional consultations were offered. At week 52, the baseline evaluation was repeated, and the costs for activities reverted to general population prices, with access to classes exclusive to ‘Live Active’ participants withdrawn.

The content of the routine data collected throughout the scheme, which was stored in the ERSs database and provided as secondary data for this research, was selected by the developers of the scheme, and was based upon the requirements of stakeholders such as the funding organisations, and the recommendations of the SEF (Cavill et al., 2012).

3.3. Alignment of the ERS with policy recommendations

The ‘Live Active’ scheme was delivered in alignment with relevant policy such as the ERS guidelines published by NICE (2014a). The eligibility criteria were in alignment with these guidelines, as the scheme did not accept referrals from individuals without a known LTC or risk factor for ill health. In addition, as recommended by NICE (2014a), the scheme collected data from participants in line with the essential criteria documented in the SEF for PA interventions (Cavill et al., 2012).

The scheme incorporated the four mandatory components of an ERS: i) assessment of PA level by a health care professional; ii) referral by a health care professional to the ‘Live Active’ scheme; iii) an individual assessment with an exercise professional to determine personal requirements; and; iv) an opportunity to participate in a PA programme (NICE, 2014a). Furthermore, throughout supervised PA sessions and one-to-one consultations, the scheme incorporated the core behaviour change techniques outlined by NICE (2014a; 2014b): i) recognising when a person is open to behaviour change; ii) agreeing behaviour change goals and relevant action plans; iii) advising upon social support; iv) tailoring behaviour change techniques and interventions to the individual; v) monitoring progress and providing feedback; and; vi) creating coping strategies to prevent relapse (see section 3.4, for further discussion).

The scheme recommended PA levels in accordance with the DoH’s (2011) PA guidelines (i.e. 150 minutes of moderate intensity PA, or 75 minutes of vigorous intensity PA per week). This was with the exception of health conditions for which these PA levels are not recommended, upon which condition-specific PA guidelines were followed.

3.4. Short summary of behaviour change theories in ERSs

There are a wealth of behaviour change theories relevant to behaviours such as PA (Michie, West, Campbell, Brown, & Gainforth, 2004). The core behaviour change techniques recommended by NICE (2014b) (section 3.3) were drawn from a range of theories. These included, for example, The Theory of Planned Behaviour (Ajzen 1991) and Self-Efficacy Theory (Bandura, 1997). Recommendations were also drawn from theories associating behaviour change with habits (Bordieu, 1977), social capital (Bordieu, 1986), and coping and resilience (Lazarus, 1976).

The Theory of Planned Behaviour (Ajzen 1991) and Self-Efficacy Theory (Bandura, 1997) are two of the most tested and utilised models of behaviour change applied to PA (Buchan, Ollis, Thomas, & Baker, 2012). Evidence supports self-efficacy as a predictor of behaviour change (Ashford, Edmunds, & French, 2010; McAuley, Szabo, Gothe, & Olson, 2011; Williams & French, 2011), and the model proposes guidelines for its development (Bandura, 1997). The Theory of Planned Behaviour proposes that behavioural intentions are influenced by attitudes, subjective norms and perceived control (Ajzen 1991). Evidence supports the role of this theory in predicting behavioural intentions (Hagger, Chatzisarantis, & Biddle, 2002; Hausenblas, Carron, & Mack, 1997). However, these theories have not been accepted without criticism. It has been argued that these theories assume that behaviour change is linear, predictable and controllable (Buchan et al., 2012; Kelly & Barker, 2016). They also do not account for environmental, social and political influences on behaviour change (Buchan et al., 2012; Davis, Campbell, Hildon, Hobbs, & Michie, 2015; McLaren & Hawe, 2005), which may be better explained by socio-ecological models (Sallis, Owen, & Fisher, 2008).

Very few published evaluations of ERSs have described the behaviour change theories or techniques implemented to promote PA behaviour change. Wormald et al. (2006) described offering motivational support and behaviour change strategies such as goal setting, identifying barriers to PA and developing coping strategies, providing diaries for self-monitoring of PA and diet, and prescription of group-based activities. Murphy et al. (2012) briefly noted that consultations were based on motivational interviewing (Miller, 1983), and included goal setting and relapse-prevention strategies. However, in practice, implementation of motivational interviewing was inconsistent (Moore, Moore, & Murphy, 2012). Lastly, Hanson et al. (2013) noted that consultations were based on the Transtheoretical Model (Prochaska & DiClemente, 1982), and involved assessment of stage of change and identification of barriers to increasing PA. Systematic assessment of ERS consultations, though, indicates that behaviour change techniques are not always delivered as described. In a content analysis of an ERS, Beck et al. (2016) identified that only 63.6% of the intended elements of a consultation were delivered. PA assessment and goal setting, for example, were not consistently delivered.

There does not appear, however, to be any best theory of behaviour change for PA at present (Michie et al., 2014). It is not surprising, therefore, that researchers have chosen to explore the role of other behaviour change theories in ERSs, to those recommended by NICE (2014b). In fact, the majority of identified research exploring behaviour change theories in ERSs have focused on Deci and Ryan's (1985) Self-Determination Theory. This theory proposes that motivation to change is influenced by three basic needs: perceived autonomy, perceived competence, and; perceived relatedness. Several researchers have explored whether the components of Self-Determination Theory can predict adherence to ERSs (Edmunds, Ntoumanis, &

Duda, 2007; Eynon, O'Donnell, & Williams, 2017; Morton, Biddle, & Beauchamp, 2008). Others have explored whether these components influence PA intentions in an ERS (Rouse, Ntoumanis, Duda, Jolly, & Williams, 2011), if participation improves these components (Littlecott, Moore, Moore, & Murphy, 2014), or if changes in the components can predict changes in PA (Littlecott et al., 2014; Rahman, Thogersen-Ntoumani, Thatcher, & Doust, 2011). Lastly, Duda et al. (2014) explored whether an autonomy supportive ERS could improve PA and health-related outcomes to a greater extent than a standard provision ERS. The findings showed that both ERSs were effective, but that compliance with delivery of an autonomy supportive ERS may have been reduced by competing work demands.

Therefore, it is apparent that different ERSs may be underpinned by different behaviour change theories. There are, though, commonalities between schemes in behaviour change techniques implemented, such as goal setting and identification of barriers. However, implementation of these behaviour change techniques does not appear to be consistent and as described (Beck et al., 2016; Duda et al., 2014; Moore et al., 2012). These findings indicate that ERS staff may not be sufficiently trained to implement behaviour change techniques (Beck et al., 2016), or that following training they view these techniques as unnecessary (Moore et al., 2012), or face competing time demands for data collection (Duda et al., 2014; Moore et al., 2012).

3.5. Critical discussion of routine outcomes

This section will provide a critical discussion of the routine measures collected during 'Live Active' consultations, including an overview of the measures and their limitations. These outcomes were those collected during routine consultations at weeks 1, 12, 24, and 52 of the scheme. These outcomes were not selected by the researcher, but by stakeholders, based upon the requirements of funding organisations, and recommendations from the SEF (Cavill et al., 2012). They included readiness to change (TTM stage of Change), current PA level and sedentary behaviour (IPAQ-SF), HRQoL (EQ-5D-3L and EQ-5D VAS), mental wellbeing (WEMWBS), BMI, BP, alcohol consumption, and smoking status.

3.5.1. Stage of change

In order to determine a participant's intentions and readiness to change their PA behaviour, the PA stage of change questionnaire (Prochaska & DiClemente, 1982) was administered face-to-face during consultations, or via telephone consultation (appendix 3). Although the stage of change questionnaire was developed as a self-administration questionnaire (Prochaska & DiClemente, 1982), it has often been administered face-to-face (Carey, Purnine, Maisto, & Carey, 1999; Marttila & Nupponen, 2003; Norcross, Krebs, & Prochaska, 2011).

The stage of change questionnaire is based on Prochaska & DiClemente's (1982) TTM of behaviour change. The model proposes that behaviour change is a process rather than a single event, and that individuals move through a series of stages when making health behaviour changes. These stages are: pre-contemplation; contemplation; preparation; action, and; maintenance (Prochaska & DiClemente, 1982). Progression through the stages is sequential, but not one-way. Individuals may progress, for example, from contemplation to action, and then return to contemplation. There is also no time limit or expected duration for which a person may spend in one stage (Prochaska & DiClemente, 1982). The model proposes that behaviour change interventions should be matched against these stages. Therefore, the goals of the individual, and the intervention that follows, are dependent upon the individual's stage of change (Prochaska & DiClemente, 1982).

The TTM has been a behaviour change theory of much contention, attracting both praise and deep criticism. An apparent advantage of the model has been the ability to categorise individuals into stages, which are aligned with specific processes of change that are hypothesised to progress an individual to the next stage (Buchan et al., 2012). However, it has been argued that this approach may present a false sense of behaviour change (Adams & White, 2003). Previous research indicates that progression through the stages is not always accompanied by changes in PA behaviour (Norris, Grothaus, Buchner, & Pratt, 2000). Therefore, a person may progress through stages without any significant change in PA level, presenting a false sense of change in PA behaviour (Norris et al., 2000).

The TTM was originally developed from the study of addictive behaviours such as smoking (DiClemente & Prochaska, 1982; Prochaska & DiClemente, 1983). However, there are fundamental differences between singular addictive behaviours such as smoking, and complex behaviours such as PA, which are characterised by multiple action types (e.g. leisure time activity, household activity, occupational activity) (Adams & White, 2003). Therefore, the TTM model is not directly generalisable to PA behaviour (Nigg et al., 2011). It has been argued that behaviours such as PA are not suited to stage based interventions (Bridle et al., 2005), as the focus of the model upon the individual underestimates the complexities of PA behaviour change (Buchan et al., 2012). In particular, PA is known to be influenced by many factors not considered by the TTM, such as age and SES, which are better explained by socio-ecological models of behaviour change (McLeroy, Bibeau, Steckler, & Glanz, 1988; Sallis et al., 2008; Stokols, 2000). This criticism is supported by systematic reviews that demonstrate mixed evidence for the impact of TTM based interventions for increasing PA, with the majority of studies finding no significant effect of a TTM based intervention on PA levels compared to a control group. The reviews concluded that the TTM approach is unsuitable for changing PA behaviour (Adams & White, 2004; Bridle et al., 2005; Mastellos, Gunn, Felix, Car, & Majeed, 2014; Riemsma et al., 2002).

It has been suggested that an advantage of the TTM is that it is intuitive and easy to use for practitioners (Nigg et al., 2011). However, there appears to be imprecision in applying the model correctly. Hutchison, Breckon, and Johnston (2009) reviewed the characteristics of interventions applying the TTM to PA behaviour change, and found that most interventions did not accurately represent the TTM. This is similar to the conclusions of other reviews, finding inconsistency in applying the principles of the TTM, and inconsistency between interventions (Mastellos et al., 2014; Riemsma et al., 2002). There appears to be great variation between interventions in the method of classifying individuals, with varying degrees of accuracy (Adams & White, 2003; Riemsma et al., 2002). Such imprecision impacts the methodological quality of interventions, and may have influenced the poor outcomes of systematic reviews.

3.5.2. Physical activity

The IPAQ-SF (IPAQ, 2005) was administered face-to-face during consultations, or via telephone consultation, to assess change in PA level (appendix 4). Although the IPAQ-SF was developed as a self-administration questionnaire (IPAQ, 2005), it has been administered through face-to-face (Hallal et al., 2010; Rosa, Gracia-Marco, Barker, Freitas Jr, & Monteiro, 2015; Van Dyck, Cardon, Deforche, & De Bourdeaudhuji, 2015) and telephone interviews (Craig et al., 2003; Hallal et al., 2010), with findings of at least moderate validity. An overview of the advantages and limitations associated with the IPAQ are discussed here, with a deeper discussion of the limitations of PA measurement presented in chapter 8.

There are two versions of the IPAQ, the IPAQ-SF and the IPAQ long form (IPAQ-LF) (IPAQ, 2005). The IPAQ-SF assesses three activity types: walking; moderate intensity PA, and; vigorous intensity PA. The IPAQ-LF is an extended questionnaire, assessing different types of activities (e.g. active transport, occupational PA) at different intensities (IPAQ, 2005). To-date, the IPAQ-SF has been preferred by researchers, due to its shorter length (Craig et al., 2003). However, its focus on PA intensity, rather than PA type, means that less information is gathered, and can only estimate changes in PA according to intensity, rather than change in PA domains such as active transport or leisure time PA (Boon, Hamlin, Steel, & Ross, 2010).

The IPAQ is the most frequently used self-report PA questionnaire worldwide (Lee, Macfarlane, Lam, & Stewart, 2011; Silsbury, Goldsmith, & Rushton, 2015; van Poppel, Chinapaw, Mokkink, van Mechelen, & Terwee, 2010). It was developed for population surveillance of PA, for adults aged between 15-69 years, to allow for comparison of trends between countries (IPAQ, 2005). However, it is being routinely used for adults aged over 69 years, and as an evaluation tool in small scale intervention studies (IPAQ, 2005). Whilst systematic reviews find good reliability for the IPAQ, its validity is variable, with research finding poor criterion validity when compared with measures such as accelerometers (Cerin et al., 2016; Craig et al., 2003; Steene-Johannessen et al., 2016; van der Ploeg et al., 2010). There is a general consensus that individuals over-report time spent in moderate or vigorous intensity PA, and under-report time spent sedentary or in light intensity PA (Cerin et al., 2016; Chastin et al., 2018; Rosenberg, Bull, Marshall, Sallis, & Bauman, 2008; Steene-Johannessen et al., 2016; van der Ploeg et al., 2010). As a self-report instrument, the IPAQ is liable to bias that may contribute to this poor validity. Potential sources of bias include: poor memory and recall of information (Altschuler et al., 2009; Baranowski, 1988; Durante & Ainsworth, 1996; Heesch, van Uffelen, Hill, & Brown, 2010); misinterpretation of terminology such as exercise intensity (Altschuler et al., 2009; Finger et al., 2015; Sallis & Saelens, 2000; Scholes, Bridges, Ng Fat, & Mindell, 2016); difficulty classifying the intensity of non-aerobic activities such as resistance exercise (Pettee-Gabriel, Morrow Jr., & Woolsey, 2012); cultural differences (Wolin, Fagin, Ufere, Tuchman, & Bennett, 2010; Yates et al., 2015); social desirability (Adams et al., 2005; Motl, McAuley, & DiStefano, 2005), and; the effect of seasonality on PA level (Pettee-Gabriel et al., 2012). Researchers have therefore argued that PA questionnaires provide an estimate of perceived PA (Pettee-Gabriel et al., 2012), and that PA levels derived from self-report PA questionnaires should be interpreted with caution (Steene-Johannessen et al., 2016).

3.5.3. Health-related quality of life

In order to obtain a standardised measure of HRQoL, the EQ-5D-3L was administered (EuroQol Group, 1990) face-to-face during consultations, or via telephone consultation (appendix 5). The EQ-5D-3L was developed as a self-administration questionnaire (EuroQol Group, 1990). However, the developers have since developed interview scripts for telephone or face-to-face interviews for participants who are unable to read, write, or visit the study site. Previous research has also administered the EQ-5D-3L through face-to-face (Lozano et al., 2016; McPhail et al., 2009) and telephone interviews (Hays et al., 2009; McPhail et al., 2009).

The EQ-5D-3L questionnaire is composed of two components, the health profile and the visual analogue scale (EQ-5D VAS). Together, these components can provide information on the health status of individuals and populations, and the effectiveness of interventions on HRQoL (Parkin, Rice, & Devlin, 2010).

The EQ-5D-3L health profile is composed of five different questions, with a three point likert scale response (Reenen & Oppe, 2015). This distinguishes it from the EuroQol five dimensions five level questionnaire (EQ-5D-5L), which has a five-point likert scale response. The five questions evaluate five different health dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Individuals classify themselves for each dimension as one of three levels (1: no problems with this dimension, 2: some problems with this dimension, 3: extreme problems with this dimension). Recording of the responses results in a five-digit code (e.g. 11213), which has no arithmetic or scoring properties. This five-digit code is transformed in to a single summary score, ranging from <0 (worse than being dead) to 1 (full health) using a formula that assigns weights to each health dimension (EuroQol Group, n.d.). These weightings have been obtained from value sets of EQ-5D-3L data collected from the general public. There are different value sets for different countries; the value set for the UK is taken from 'Measuring and Valuing Health' study (Dolan, 1997). In this study, individuals were asked to consider if they had experienced any level of the five health dimensions, and to assign a value to each level using a visual analogue scale. This is known as Time Trade Off (Dolan, 1997), a method used to understand how much people value various health states (Torrance, Thomas, & Sackett, 1972).

The second part of the EQ-5D-3L, the EQ-VAS, is a quantitative measure of an individual's overall assessment of their health on a given day. Individuals are asked to rate their health on a scale from 0-100, with 0 corresponding with worst imaginable health, and 100 corresponding with best imaginable health (Reenen & Oppe, 2015).

The EQ-5D-3L has shown good reliability and validity in multiple populations (Dyer, Goldsmith, Sharples, & Buxton, 2010; Janssen, Lubetkin, Sekhobo, & Pickard, 2011;

Johnson & Coons, 1998; Pickard et al., 2008). However, there has been growing criticism surrounding methods of scoring, analysis and reporting of EQ-5D-3L health profile data (Devlin, Parkin, & Browne, 2010; Parkin et al., 2010; Parkin, Devlin, & Feng, 2014). The 3-point likert scale scoring function of the health profile has led to a ceiling effect in previous research, where limited response options have led participants to report the highest scores possible, questioning the extent to which true HRQoL was measured (e.g. Bharmal & Thomas, 2006; Craig, Pickard, & Lubetkin, 2014; Feng, Devlin, & Herdman, 2015; Sullivan, Lawrence, & Ghushchyan, 2005). Furthermore, the limited response options have left researchers unable to identify small to moderate changes in health status (Feng et al., 2015; Macran, Weatherly, & Kind, 2003; Myers & Wilks, 1999; Pan et al., 2015; van de Willige, Wiersma, Nienhuis, & Jenner, 2005). In contrast, the more recent EQ-5D-5L, which introduced the 5-point likert scale response, has reduced the ceiling effect and increased sensitivity to change (Craig et al., 2014; Feng et al., 2015; Janssen et al., 2013; Pan et al., 2015), due to the increased availability of responses that more accurately describe current health status (Craig et al., 2014).

The health profile contains detailed self-reported data about an individual (Parkin et al., 2010). However, categorical changes in the health profile dimensions following an intervention have rarely been reported, with much of this potentially insightful information lost when researchers opt to solely report the single summary score (Devlin et al., 2010). The single summary score has to date been the most popular method of reporting EQ-5D-3L health profile outcomes, due to the ease at which it can be analysed for statistical inference, and reported as a quantitative measure of HRQoL (Devlin et al., 2010). Upon recognition of trends in reporting of the health profile, the developers of the EQ-5D-3L, have recommended that health profile data is always reported in frequency tables, listing the number and percentage of individuals reporting each level of each health dimension (Reenen & Oppe, 2015). However, no recommendation have been provided for how this data should be analysed for statistical inference (Devlin et al., 2010; Parkin et al., 2010).

The formulation of the single summary score has also received criticism, in particular the weightings that are applied to the health profile data in order to formulate this score (Devlin et al., 2010). It has been argued that in clinical populations, single summary scores are invalid, since the weightings have been formulated using data obtained from the general population (Devlin et al., 2010). Value sets also introduce an exogenous source of variance to the EQ-5D-3L data, since the single summary scores are dependent upon which value set is used, and its properties (Parkin et al., 2010; Wilke et al., 2010). Any analysis of a single summary score is an analysis of not only the data from the given study, but also of the data from the value set used (Parkin et al., 2010). Therefore, the direction and magnitude of any result may reflect a variance in the value set rather than the given data (Devlin et al., 2010; Parkin et al., 2010). Furthermore, the applied weightings cause a non-normal distribution of the single summary score data, which is typically bi-modal (Parkin et al., 2014). This can be explained by the disproportionate decline in the single summary score when a health dimension response moves between levels 2 (some problems) and 3 (extreme problems), compared to levels 1 (no problems) and 2 (Fransen & Edmonds, 1999).

In light of these complexities, it has been argued that statistical testing of the single summary score is inappropriate, as it may provide misleading results (Parkin et al., 2010). When reporting EQ-5D-3L data, it has been recommended that the EQ-VAS single score is favoured over the health profile single summary score, with frequency tables for the dimensions of the health profile provided (Devlin et al., 2010; Parkin et al., 2010; Parkin et al., 2014).

3.5.4. Mental wellbeing

Mental wellbeing was assessed using the WEMWBS questionnaire (Tennant et al., 2007), which was administered face-to-face during consultations, or via telephone consultation (appendix 6). The WEMWBS was developed as a self-administration questionnaire, and no identified published studies have administered the questionnaire using alternative modes, such as face-to-face administration (Tennant et al., 2007).

To date there is no single, widely agreed upon definition of mental wellbeing. Ryan and Deci (2001) argued that mental wellbeing is defined by two concepts: feeling good – positive affect and life satisfaction (hedonic perspective) and; functioning well - positive psychological functioning and self-realisation (the eudaimonic perspective). This is the perspective of mental wellbeing from which the WEMWBS was developed (Taggart, Stewart-Brown, & Parkinson, 2015).

The WEMWBS is a 14-item scale for measuring mental wellbeing over the previous two weeks in the general population (Tennant et al., 2007). Distinct from measuring mental illness, the WEMWBS addresses subjective wellbeing and psychological functioning, using a 5-point likert scale, with responses ranging from ‘none of the time’ to ‘all of the time’. The possible resulting score ranges from 14 to 70. There are no definitive cut-off points to classify overall scores as low or high mental wellbeing, however, at the individual level a change in score of 3 or more points can be considered a significant change (Maheswaran, Weich, Powell, & Stewart-Brown, 2012).

The WEMWBS is considered to be a psychometrically robust scale for the evaluation of population level mental wellbeing. It has shown good validity and reliability for use with general and minority populations aged 13 years and above, in multiple languages including the English language (Stewart-Brown et al., 2011; Taggart et al., 2013; Tennant et al., 2007). To date, the WEMWBS has been tested as a self-administered questionnaire, however its robustness during a face-to-face interview format has not been explored (Taggart et al., 2015).

The WEMWBS is sensitive to change at both the individual and population level, in both general and psychiatric populations (Maheswaran et al., 2012). However, it was not developed for the context of individual level mental wellbeing, but rather population level mental wellbeing (Taggart et al., 2015). It was also not developed to screen or identify individuals with high or low mental wellbeing, therefore no cut-off points for these states were published (Taggart et al., 2015). Despite this, many researchers have used the WEMWBS as a screening tool (see for e.g. Gremigni & Bianco, 2012), and subsequently cut-off points have been suggested, such as: overall score ≤43.5 to identify individuals with symptoms of depression, or; a change in total score of ±3 to indicate a significant change in mental wellbeing (Taggart et al., 2015).

The WEMWBS tends to have a near normal distribution, without a ceiling effect (Bartram, Yadegarfar, Sinclaur, & Baldwin, 2011; Taggart et al., 2015; Tennant et al., 2007). It’s distribution in the general population can be influenced by several factors such as age, gender, SES, PA level, fruit and vegetable consumption, alcohol consumption, and smoking status (Stranges, Samaraweera, Taggart, Kandala, & Stewart-Brown 2014; Taggart et al., 2015; Wilson, Kellock, Adams, & Landsberg, 2015). WEMWBS scores, for example, tend to be higher for those: of older age; who are male; of higher SES; who are physically active, and; do not currently smoke (Stranges et al., 2014; Taggart et al., 2015; Wilson et al., 2015).

3.5.5. Body mass index

Body composition was evaluated by calculating BMI, which represents the relationship between body weight and height (Keys, Fidanza, Karvonen, Kimura, & Taylor, 1972; WHO, 2000). Where height and weight could not be measured objectively, approximate self-report measures were recorded. However, this is generally not an accepted method of measurement, as it is liable to under- or over-estimation (Connor Gorber, Tremblay, Moher, & Gorber, 2007; Maukonen, Männistö, & Tolonen, 2018).

A variety of methods exist for the measurement of body composition. When selecting a method, it is important to consider the evidence for different methods of body composition measurement, in addition to consideration of practicality and cost (Toomey, Cremona, Hughes, Norton, & Jakeman, 2015). Simple methods such as body weight and BMI measurement have become the standard measurement in public health settings, despite limitations affecting the specificity and sensitivity of measurement (Andreoli, Garaci, Cafarelli, & Guglielmi, 2016; Madden & Smith, 2016; Toomey et al., 2015). Using methods that have major limitations can lead to incorrect conclusions being made about the health status of an individual, or the impact of an intervention. Ultimately these conclusions may incorrectly inform policy development (Toomey et al., 2015).

Due to its simplicity and ease of administration, body weight has been the most widely used method for assessing body composition in applied settings (Andreoli et al., 2016; Madden & Smith, 2016). Body weight is the sum of all body components, including muscle, fat and water. However, it cannot discriminate between these components, and in which component the changes have occurred (Madden & Smith, 2016). Change in body weight is susceptible to diurnal variations due to food and water intake/excretion (Madden & Smith, 2016). For individuals with health conditions, body weight may also vary over time due to pathological changes, such as changes in skeletal muscle mass (Ribeiro & Kehayias, 2014) or fluid balance (Morgan, Madden, Jennings, Elia, & Fuller, 2006). Where actual measurement of body weight is not feasible, self-report or observer-reported body weight is not an accepted method of measurement, with systematic reviews noting significant underestimations of body weight using this method (Connor Gorber et al., 2007; Maukonen et al., 2018).

BMI represents the relationship between body weight and height (Keys et al., 1972; WHO, 2000). Both measures are relatively easily to collect using inexpensive equipment and standardised procedures, requiring minimal training (Tuttle, Montoye, & Kaminsky, 2016). BMI is also relatively well understood by the general population (Tuttle et al., 2016). As it is derived from body weight, BMI also cannot discriminate between different components of the body (Madden & Smith, 2016). Therefore, it is an imperfect measure of adiposity, and is prone to misclassifying individuals (Okorodudu et al., 2010). Approximately 50% of individuals with excess body fat are misclassified as normal weight when measured using BMI (Okorodudu et al., 2010).

Lack of discriminatory capacity to stratify individuals correctly, in addition to criticism of the capacity of BMI to predict adverse health outcomes (Neeland & de Lemos, 2016), has led researchers to consider alternative measurements such as waist circumference, skinfold thickness and dual-energy X-ray absorptiometry (Cornier et al., 2011; Neeland & de Lemos, 2016).

3.5.6. Blood pressure

BP was measured using automatic BP monitors, which are the most frequently used device in office-based settings, as they are quick and easy to use, and do not require training to operate (Palatini & Frick, 2012). They also eliminate the problem of observer error and bias, compared to manual methods of measurement (Palatini & Frick, 2012; Parati et al., 2008; Pickering et al., 2005). Measurement in settings such as offices assumes that the measurement produced is indicative of real life average BP (Pickering, Gerin, Schwartz, Spruill, & Davidson, 2008).

The accuracy of automatic BP measurement is limited by both technical and human errors. Monitors require regular maintenance and calibration to maintain accuracy of measurement to within ±3 millimetres of mercury (mmHg) (O’Brien et al., 2003; Pickering et al., 2005). However, research suggests that monitors in healthcare settings are not regularly calibrated, resulting in measurement errors (Ali & Rouse, 2002; Coleman, Steel, Ashworth, Vowler, & Shennan, 2005; Knight et al., 2001; Markandu, Whitcher, Arnold, & Carney 2000). Measurement must also be taken using an arm cuff that is appropriate for the size of the arm (O’Brien, 1996; O’Brien et al., 2003; Pickering et al., 2005). A cuff that is too small will produce an erroneously high measurement, whilst a cuff that is too large will produce an erroneously low measurement (Wofford et al., 2002). It is possible that an appropriate arm cuff may not be available for obese populations (Irving, Holden, Stevens, & McManus, 2016), where both the size and shape of the arm must be taken in to account (Palatini & Parati, 2011). BP measurement may also be influenced by seating position (Netea, Lenders, Smits, & Thien, 2003), position of the arm (Netea, Lenders, Smits, & Thien, 1999), and interaction between the practitioner and individual (O’Brien et al., 2003). Some individuals may also experience anxiety when having BP measured, more commonly known as “white coat syndrome” or “white coat hypertension” in clinical settings (Mancia et al., 1983a; Mancia et al., 2013). Lastly, BP is a variable outcome, influenced over a 24-hour period by many behavioural and environmental factors such as: sleep, temperature, respiration, emotions, recent PA, nutrition, and smoking (Brook, Weder, & Rajagopalan, 2011; Mancia et al., 1983b; Mancia, 2012; Parati et al., 1990).

3.6. Summary

This chapter has provided an overview of the research setting in which this research took place, including the rationale for the scheme, the eligibility criteria, the referral process, and the routine outcomes that were measured. This was detailed, so that readers of this thesis have an exact understanding of what was delivered. This should also facilitate the comparison of this ERS to others, for those who wish to determine the transferability of the results.

A critical discussion of the routine outcomes that were measured during consultations was also provided, given that these outcomes were selected by the developers of the ERS, and not the researcher. This identified that the majority of the routine outcomes collected have both advantages and disadvantages, which should be considered in the interpretation of the results presented in chapters six and eight.

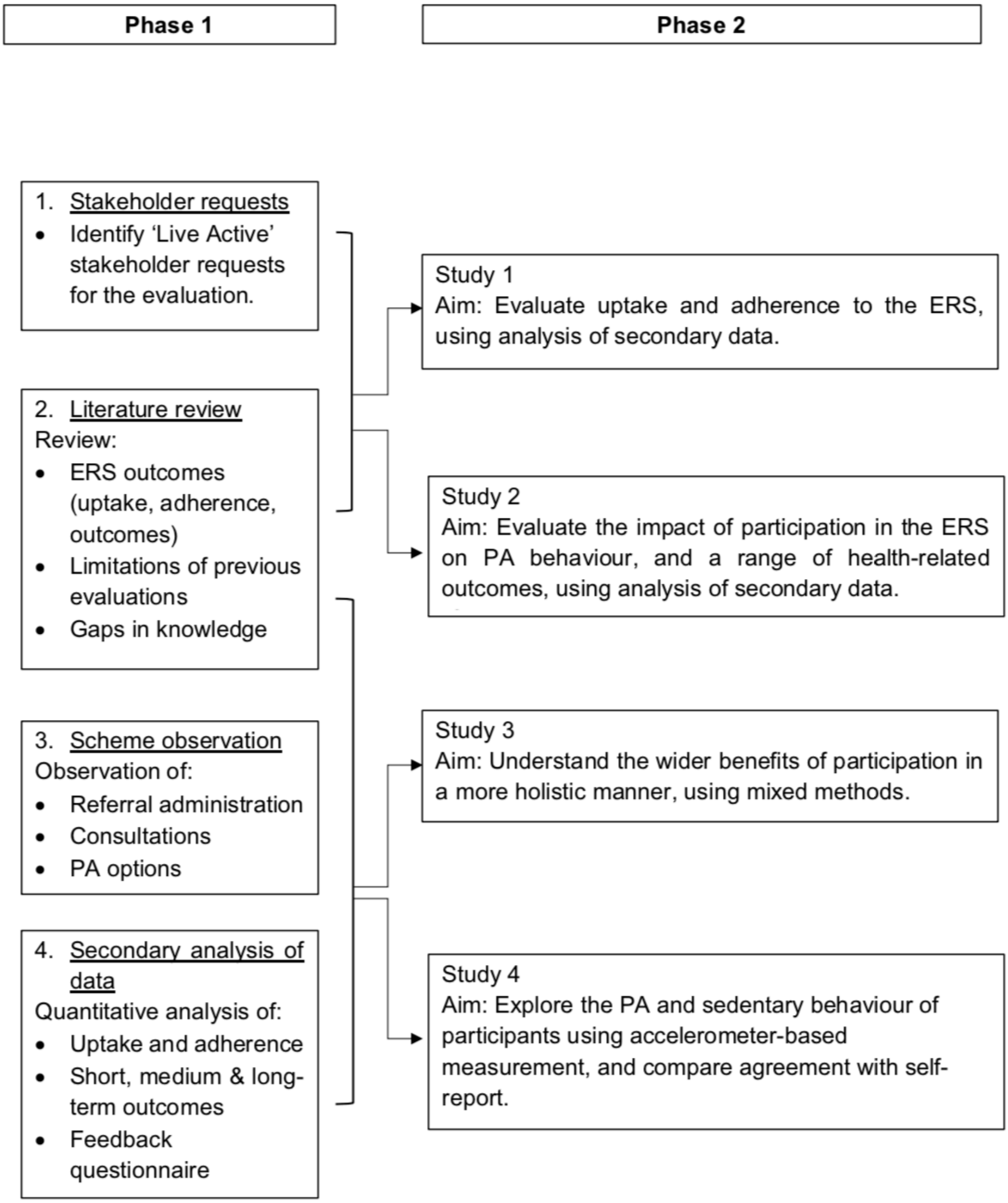
Chapter 4 Methodology

4.1. Introduction

The purpose of this chapter is to provide an overview of the methodology for this research study, including the overall research framework, and a critical discussion of the research approach that was taken. Relevant methodology for each research study will be discussed in the respective chapters.

4.2. Research framework

An overview of the research framework is provided in Figure 4.1, and is repeated at the beginning of chapters 5-8 to highlight which stage of the research the chapter relates to. The overall aim of this research was to evaluate the effectiveness of the 'Live Active' scheme at increasing PA levels and improving health outcomes in its participants, as per the aim of an ERS provided by NICE (2014a). To achieve this, a mixed methods study was carried out in two phases. The first phase, which was completed between October 2015 and December 2016, was exploratory in nature with the overall aim of developing a set of research questions to be addressed in phase 2. A critical discussion and rationale for using mixed methods in research can be found below in section 4.4.2 and 4.4.3 respectively.



1. Stakeholder requests
- Identify ‘Live Active’ stakeholder requests for the evaluation.

2. Literature review

Review:
- ERS outcomes (uptake, adherence, outcomes)
- Limitations of previous evaluations
- Gaps in knowledge

3. Scheme observation

Observation of:
- Referral administration
- Consultations
- PA options

4. Secondary analysis of data

Quantitative analysis of:
- Uptake and adherence
- Short, medium & long-term outcomes
- Feedback questionnaire

Study 1

Aim: Evaluate uptake and adherence to the ERS, using analysis of secondary data.

Study 2

Aim: Evaluate the impact of participation in the ERS on PA behaviour, and a range of health-related outcomes, using analysis of secondary data.

Study 3

Aim: Understand the wider benefits of participation in a more holistic manner, using mixed methods.

Study 4

Aim: Explore the PA and sedentary behaviour of participants using accelerometer-based measurement, and compare agreement with self-report.

Figure 4.1. Overview of the research framework, including the objectives of phase 1, and the research aims to be addressed in phase 2.

4.3. Overview of the two phases of research

This section provides an overview of the two phases of the research, including the objectives of each phase, the research aims and objectives to emerge, and a summary of the methods used to answer the research aims and objectives.

4.3.1. Phase 1

In phase 1, the overall aim was to identify the research aims and objectives to be addressed in phase two. To achieve this, four objectives were set out and accomplished: 1) identify the requirements of ‘Live Active’ stakeholders for this evaluation; 2) review the existing literature regarding ERSs; 3) observe all areas of the ‘Live Active’ scheme, and; 4) conduct a secondary analysis of routine outcomes collected by the ERS during the first year of operation.

To identify the requirements of ‘Live Active’ stakeholders for this evaluation a steering group was formed, which met quarterly. The group included representatives from the University of Salford, Tameside public health, ‘Live Active’, and Active Tameside. The primary areas of interest to arise from the group included: an understanding of rates of uptake and adherence to the scheme, reasons for non-uptake or dropout, and characteristics of individuals choosing not to uptake, or dropout, and; the medium to long-term PA behaviour and health outcomes of participation, with respect to the outcomes routinely collected by the scheme.

An observation of the scheme was conducted throughout the first year, involving areas such as referral administration, consultations, exercise classes, gym sessions, and community activities. These observations were made overtly, with the prior verbal consent of participants. This provided an opportunity to become familiarised with the ERS, and observe participants and settings first hand rather than basing understanding of the scheme on second-hand accounts (Cohen, Manion, & Morrison, 2011). It also provided an opportunity for everyday events and behaviours to be observed, that otherwise may have gone unnoticed or not reported (Cooper & Schindler, 2013).

Secondary analysis of data collected by the scheme between July 2015 and July 2016 was conducted on a quarterly basis, and presented in a quarterly and annual report. This was a descriptive analysis of: readiness to change (TTM Stage of Change); current PA level (IPAQ-SF); HRQoL (EQ-5D-3L and EQ-5D VAS); mental wellbeing (WEMWBS); BMI; BP; alcohol consumption; and smoking status, which were measured during consultations at weeks 1, 12, 24 and 52. The scheme's routine feedback questionnaire (appendix 2), completed at either the week 12 or 24 consultation, was also analysed, in addition to figures relating to total referrals received, referrals not up-taking, referrals dropping-out, and referrals completing. The objective was to compare the outcomes with those reported in the literature, in order to identify any typical or untypical findings that could be further explored in phase two of the research. It is noted that whilst this information was presented in a quarterly report to stakeholders, intended as feedback, no substantial changes to practice were introduced until the conclusion of this evaluation. Any changes introduced during the evaluation period were not judged to significantly change practice or outcomes. A critical discussion of using secondary data in research can be found below in section 4.4.1.

The combination of stakeholder requests and the findings of the literature review informed the development of research questions to be addressed in studies 1 and 2. The combination of the literature review, scheme observation and secondary analysis of data informed the development of research questions to be addressed in studies 3 and 4.

4.3.2. Phase 2

In phase 2, the research aims to emerge from phase 1 were addressed concurrently. This included:

Study 1: Evaluate uptake and adherence to the scheme, using analysis of secondary data.

Study 2: Evaluate the impact of participation in the ERS on PA behaviour, and a range of health-related outcomes, using analysis of secondary data.

Study 3: Understand the wider benefits of participation in a more holistic manner, using mixed methods.

Study 4: Explore the PA and sedentary behaviour of participants using accelerometer-based measurement, and compare agreement with self-report.

Studies 1 and 2 built on the descriptive analysis of data conducted in phase 1, using a two-year data set (July 2015 – July 2017), and inferential statistics to identify predictors of uptake and adherence, and evaluate health-related outcomes.

Study 1

The aim of study one was to evaluate uptake and adherence to the ERS, using analysis of secondary data. The research objectives were:

1.1. Identify rates of uptake, non-uptake, dropout and scheme completion.

1.2. Report the characteristics of participants who are referred to the scheme, uptake or do not uptake the referral, and dropout or complete the scheme (e.g. age, gender, SES, disability status, referral source, referral reason).

1.3. Identify at which time point the majority of dropout occurs.

1.4. Report the reasons for non-uptake and dropout.

1.5. Identify variables (e.g. age, gender, SES, source of referral, reason for referral) that predict uptake and adherence.

Objectives 1.1 - 1.4 emerged from stakeholder requests for this evaluation, to understand level of uptake and adherence to the scheme, the characteristics of participants who uptake and adhere or not, and why participants non-uptake or dropout. Objective 1.5 emerged from the literature review, which indicated that knowledge could be enhanced by the inclusion of further variables, such as SES and ethnicity, in prediction models (section 2.8.1). Novel aspects of this study, identified in the literature review, included: i) identifying referral rates in under-represented groups (e.g. disability status, ethnicity) (1.2); ii) identifying the time points at which dropout occurs (1.3); iii) identifying the reasons for non-engagement (1.4), and; iv) including a wider range of variables in prediction models of engagement (1.5). The rationale for this study and its methodology, including the original contribution to knowledge, are further discussed in chapter 5.

Study 2

The aim of study two was to evaluate the impact of participation in the ERS on PA behaviour, and a range of health-related outcomes, using analysis of secondary data. The research objectives were:

2.1. Identify changes in PA level, BMI, BP, HRQoL (EQ-5D-3L, EQ-5D VAS), mental wellbeing (WEMWBS), and smoking and alcohol consumption between weeks 1, 12, and 24 (medium-term), and weeks 1, 24 and 52 (long-term).

2.2. Investigate whether differences in PA level and health-related outcomes exist between primary referral condition groups.

Objective 2.1. emerged from stakeholder requests for this evaluation, to understand the impact of participation on the routinely measured outcomes. Objective 2.2. emerged from the literature review, which indicated that participants with different referral conditions may improve these outcomes to different extents (section 2.8.2). Novel aspects of this study included: i) reporting new outcomes such as the proportion of participants physically active <90-150 minutes per week, as well as ≥90-150 minutes per week (2.1); ii) identifying the long-term impact of participation on PA and health-related outcomes (2.1.), and; iii) identifying whether differences in outcomes exist between referral condition groups (2.2.). The rationale for this study and its methodology, including the original contribution to knowledge, are further discussed in chapter 6.

Study 3

The aim of study three was to understand the wider benefits of participation in a more holistic manner, using mixed methods. It set out to achieve the following research objectives:

3.1. Evaluate the medium to long-term (6-12 months) benefits of participation in the ERS, using semi-structured interviews and secondary analysis of the 'Live Active' feedback questionnaire.

3.2. Identify if these emerging benefits are outcomes measured by the 'Live Active' ERS, previous ERS evaluations, or ERS evaluation frameworks.

The objectives of this study emerged from i) the literature review, which identified a need for evaluations to consider the holistic benefits of participation, particularly through qualitative research (section 2.8.4); ii) observations of the scheme, which noted that participants discussed a wide range of benefits to participation in follow-up consultations, many of which were not routinely captured by the scheme, and; iii) secondary analysis of feedback questionnaire data, which identified frequent reporting of alternative benefits such as improved physical fitness and socialisation (appendix 7). Novel aspects of this study included: i) using a mixed methods approach to explore the benefits of participation (3.1); ii) identifying whether the benefits to emerge were routinely captured by the scheme, and therefore whether routinely collected quantitative outcomes estimate impact correctly (3.2), and; iii) exploring the significance of benefits to emerge, including potential cost savings. The rationale for this study and its methodology, including the original contribution to knowledge, are further discussed in chapter 7.

Study 4

The aim of study four was to explore the PA and sedentary behaviour of participants in the ERS using accelerometer-based measurement, and compare agreement with self-reported PA and sedentary behaviour. The research objectives were:

4.1. Evaluate accelerometer-measured PA and sedentary behaviour of participants in week 1, to determine patterns of behaviour upon entry to the scheme.

4.2. Evaluate accelerometer-measured changes in PA and sedentary behaviour between weeks 1 and 12.

4.3. Compare time spent sitting per day according to i) the IPAQ-SF and ii) the activPAL™ accelerometer, at weeks 1 and 12.

4.4. Compare time spent in light and moderate-vigorous intensity PA lasting 10 minutes or more according to i) the IPAQ-SF and ii) the activPAL™ accelerometer, at weeks 1 and 12.

The objectives of this study emerged from i) the literature review, which identified that to-date, all measurement of habitual PA in published evaluations of ERSs (UK based) was self-reported (section 2.8.3); ii) observations of the scheme, which noted that participants often experienced difficulty recalling PA and sedentary behaviour, and; iii) secondary analysis of data, which identified untypical increases in PA compared to previous literature (appendix 8). Novel aspects of this study included: i) using accelerometers to identify short-term changes in PA and sedentary behaviour (4.2.), and; ii) the comparison of accelerometer-measured and self-reported PA and sedentary behaviour, in ERS participants (4.3., 4.4). The rationale for this study and its methodology, including the original contribution to knowledge, are further discussed in chapter 8.

4.4. Critique of the methods

The following section will provide a critique of secondary data analysis, which was the basis of studies 1 and 2. This is followed by a critique of mixed methods research, and the rationale for using mixed methods in this thesis. It will finish with a reflexive statement about how my philosophical stance has potentially contributed to the methods used in this thesis.

4.4.1. Secondary data: a summary

Secondary analysis refers to the analysis of a data set that was not collected by the researcher, but another individual, group or organisation (Boslaugh, 2007; Dale, 2004; Vartanian, 2011). Using secondary data can be time and cost-effective, as the data are readily available to be analysed, avoiding time consuming research processes such as instrument development and data collection (Doolan & Froelicher, 2009; Smith, 2008; Vartanian, 2011). Often, the data set may be of larger quantity and depth, and higher quality than the researcher may have been able to otherwise accomplish (Boslaugh, 2007; Dale, 2004; Smith, 2008). This is beneficial since larger samples, for example, may be more representative of the population of interest, and may increase the validity and generalisation of the results (Smith, 2008; Smith et al., 2011). Furthermore, a secondary analysis of data can be exploratory, and highlight relationships and hypotheses of interest to be pursued in the future research (Dale, 2004).

Secondary analysis of data however has several limitations, which primarily arise from the researcher not being involved with the original data collection (Boslaugh, 2007; Dale, 2004; Vartanian, 2011). As the data were collected by other individuals, and often for different objectives, the data set may not always contain relevant or necessary outcomes for the researcher's objectives (Boslaugh, 2007; Doolan & Froelicher, 2009; Vartanian, 2011). Even where relevant outcomes have been collected, the researcher may have preferred these to be collected using alternative research instruments. The researcher is also unable to influence how the data is collected, or may not know how it was collected, such as through a face-to-face interview or self-administered questionnaire (McKnight & McKnight, 2011; Vartanian, 2011). There is also no influence on how data are recorded, such as using continuous or categorical variables, and the data may not be readily compatible with analysis software (Dale, 2004). Missing data may also be problematic, and the reasons for missing data may be unknown to the researcher (McKnight & McKnight, 2011; Vartanian, 2011). Therefore, secondary analysis of data can be considered a trade-off between accessibility and convenience, and control over the data collection, and is more suitable where the objective is exploration and hypothesis generation, rather than hypothesis testing (Vartanian, 2011).

4.4.2. Mixed methods research: a summary

Mixed methods research has become a widely accepted practice in many fields such as health care (Caracelli & Greene, 1993; Casebeer & Verhoef, 1997; Greene & Caracelli, 1997; Morse, 1991; Tashakkori & Teddlie, 1998), and has been described as the third major research approach, following quantitative and qualitative research

(Johnson, Onwuegbuzie, & Turner, 2007). Various definitions for mixed methods exist, but a central theme is that it brings together quantitative and qualitative research methods in a single study, or multi-phased study (Creswell, 2003; Johnson et al., 2007; Tashakkori & Teddlie, 1998).

The emergence of mixed methods as an accepted practice has not been without debate. Much of the argument has surrounded the opposing philosophical assumptions of quantitative and qualitative research, with some arguing that the two methods cannot be mixed due to the ontological and epistemological differences (Sale, Lohfeld, & Brazil, 2002). Quantitative research is rooted in positivism, and proposes that there is one objective reality that exists independent of human perception (Denzin & Lincoln, 1994). In contrast, qualitative research is rooted in interpretivism (Altheide & Johnson, 1994; Kuzel & Like, 1991; Secker, Wimbush, Watson, & Milburn, 1995), which posits that there are multiple realities that are based on one's construction of reality (Berger & Luckmann, 1966; Denzin & Lincoln, 2017).

Mixed methods research however does not root itself in either positivism or interpretivism, rather it is based on pragmatism, which offers a different world view where the focus of the researcher is on which methods can answer the research questions at hand (Brewer & Hunter, 1989; Creswell & Plano Clark, 2011; Miller, 2006; Tashakkori & Teddlie, 1998). This is summarised well by Powell (2001), "to a pragmatist, the mandate of science is not to find truth or reality, the existence of which are perpetually in dispute, but to facilitate human problem-solving" (p. 884). Others have argued that the two methods can be mixed because they share a unified logic (King, Keohane, & Verba, 1994), with a shared goal of understanding the world in which we live (Haase & Myers, 1988), and how the human condition can be improved (Reichardt & Rallis, 1994). Furthermore, both methods share a commitment to rigor, conscientiousness, and critique in the research process (Reichardt & Rallis, 1994). Reichardt and Cook (1979) also dismissed the idea that certain epistemological paradigms must be attached to certain methods. Miles and Huberman (1984) argue that researchers should not become pre-occupied with justifying the philosophical underpinnings of mixed methods research, since the paradigm debate will not be resolved soon, and that "epistemological purity doesn't get research done" (p. 21).

There are many reasons for a researcher to consider mixed methods. Firstly, many research problems encountered, particularly in public health, involve the study of complex phenomena. Complete understanding often requires data from a range of perspectives, which dictates the use of a range of quantitative and qualitative methods (Baum, 1995; Clarke & Yaros, 1988; Steckler, McLeroy, Goodman, Bird, & McCormick, 1992).

One of the most acknowledged reasons for using mixed methods is that using two methods can yield greater knowledge than using either method alone (Barbour, 2000; Bowling, 2014; Bryman, 2001). Singular methods inherently have both strengths and weaknesses, and combining both methods can be complementary and allow the weaknesses of one method to be overcome by the strengths of another (Bryman, 1984; Sieber, 1973). In example, quantitative methods typically provide numerical descriptives, estimates of size and distribution of effects, and allow for tests of statistical significance (Bamberger, Rao, & Woolcock, 2010). Qualitative approaches can progress upon these outcomes, with in-depth exploration, description and analysis of processes and patterns (Bamberger et al., 2010). Mixed methods can therefore be complementary, capturing the strengths of both methods (Kushman, 1992; Tashakkori & Teddlie, 2003), providing richer data and a greater depth in understanding of a subject (Collins, Onwuegbuzie, & Jiao, 2006; Greene, Caracelli, & Graham, 1989; Rossman & Wilson, 1985). However, in practical terms, combining two research methods can be much more difficult than using either alone, requiring different methods of data collection and data analysis, followed by triangulation of quantitative and qualitative findings (Hansen, 2006).

Further strengths of mixed methods research include the potential for triangulation of quantitative and qualitative data, to increase confidence in the conclusions reached about a research question, by either confirmation or corroboration (Brannen, 1992; Greene et al., 1989; Rossman & Wilson, 1985). Mixed methods can also increase the scope of research, for example the results of a quantitative study can inform the development of follow-up qualitative research (Barbour, 2000; Greene et al., 1989). It can also inform the development of research instruments, for example qualitative data may be used to inform the development of a quantitative survey (Gabriel & Bowling, 2004; Greene et al., 1989).

Mixed methods approaches have grown faster in the field of evaluation than any other, perhaps due to the practical nature of evaluations, and the necessity to consider multiple sources of evidence when drawing conclusions (Johnson et al., 2007). In 1979, Reichardt and Cook encouraged researchers to adopt mixed method for evaluation, stating “there is no reason for researchers to be constrained to either one of the traditional, though largely arbitrary, paradigms when they can have the best from both” (p. 18-19). Mixed methods has also been advocated for by the MRC's (2008) framework for the evaluation of complex interventions.

The use of mixed methods in evaluations has advantages that go beyond those discussed. Evaluations that are grounded solely in pre and post quantitative measures render the design inflexible. In practice, evaluations that involve stakeholders often require flexibility, to adapt to changes in circumstances, and meet the priority questions of stakeholders, which may change with time. Stakeholders also often require real-time feedback, to identify implementation issues and outcome forecasts, not all of which can be provided by pre and post measures (Bamberger et al., 2010).

Furthermore, different stakeholders will value different types of outcomes, and therefore research approaches (Bamberger et al., 2010; Farquhar, Ewing, & Booth, 2011). Mixed methods therefore provides a solution to the barriers that often cause low uptake of evaluations (Patton, 2008), with the collection of data that can adapt to changing requirements, and satisfy a range of stakeholders (Bamberger et al., 2010; Farquhar et al., 2011). In addition, mixed methods data allows the evaluation to be communicated in various styles to the needs of the audience, such as figures and tables to outcome focused stakeholders, or qualitative excerpts to those interested in participant experiences (Bamberger et al., 2010).

In summary, mixed methods research has attracted much criticism, mostly on a philosophical level. Despite this, it is becoming an increasingly popular research method, particularly in research evaluations, due to the advantages it can provide in solving research questions.

4.4.3. Rationale for using mixed methods

Having considered the strengths and limitations of mixed methods research, this section summarises the rationale for using mixed methods in this thesis.

The first phase of the research considered the emerging quantitative outcomes of the first year of the ‘Live Active’ scheme, using secondary data. A sustained quantitative approach to the evaluation would limit its usefulness (e.g. opportunities for learning), be insufficient to fully address the research aims of the study, and would almost certainly miss out on phenomena occurring (Johnson & Onwuegbuzie, 2004). Consequently, the impact of the ERS could be underestimated through this sustained approach (Wilson, 2015). Therefore, introducing a qualitative approach in phase 2 was appropriate, and allowed for a deeper exploration and discovery of the outcomes of participation (Creswell & Plano Clark, 2011; Johnson & Onwuegbuzie, 2004; Tashakkori & Teddlie, 2003), in order to make a balanced evaluation of the scheme's impact (Dugdill et al., 2005; Riddoch et al., 1998). Using a mixed methods approach was also appropriate to overcome the limitations of RCTs, which cannot comprehensively evaluate complex interventions such as ERSs (Thirsk & Clark, 2017).

A mixed methods approach was also appropriate due to the varying, and sometimes changing priorities of stakeholders associated with the 'Live Active' scheme (Bamberger et al., 2010). This approach satisfied the requirements of stakeholders interested in quantifiable changes to health outcomes and process measures. However, this approach also assured the relevance of the evaluation to stakeholders who were more interested in participant experiences and journeys, and mechanisms contributing to changes in PA behaviour.

4.4.4. Reflexive statement

Reflexivity has typically been associated with enhancing the credibility of qualitative research, due to the researcher's position as an instrument of inquiry, where personal perspectives may influence the research process (Mays & Pope, 2000; Patton, 2002). However, the principles of reflexivity transfer to mixed methods research, and evaluation, if the researcher shapes the inquiry and generation of knowledge (Walker, Read & Priest, 2013).

Therefore, having noted how the research processes in phase one contributed to the development of the four research studies, it is also important to reflect on how my philosophical stance has contributed to the rationale for the adoption of mixed methods research. My personal values align most with pragmatism, with a lean towards interpretivism. Exploration in the pursuit of solving real world problems is therefore characteristic of my research, but in a manner that recognises the contribution of human experience to knowledge where appropriate.

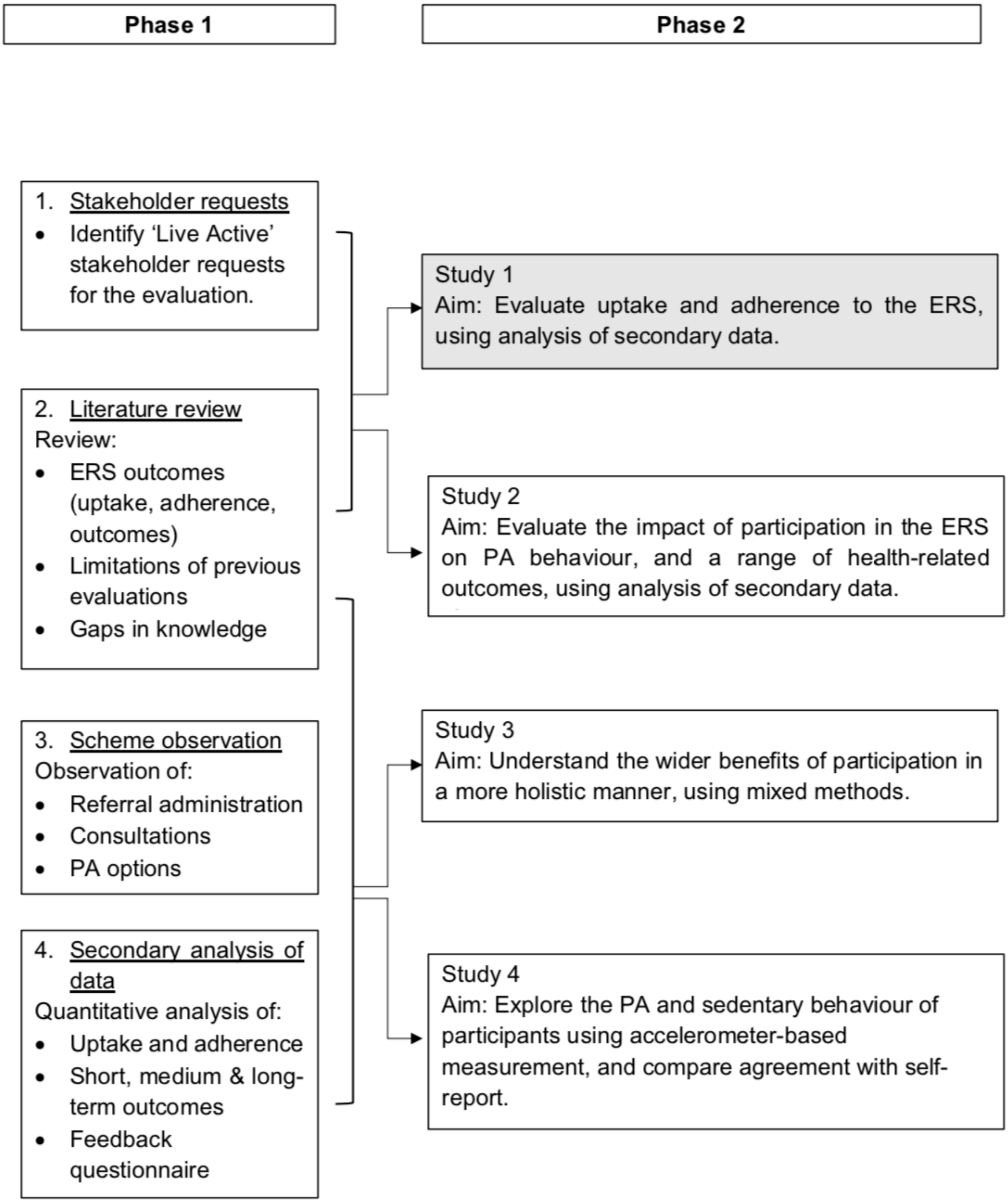
As a result of my values, I perceived that mixed methods research was the only viable method to comprehensively evaluate this ERS. It has been said that, for pragmatists, doubt or a sense that something is not right, can initiate research inquiry (Elkjaer & Simpson, 2011). It was my view that whilst quantitative analyses could provide important information to stakeholders, it would only answer the “what” questions, rather than the “how” and “why”. In turn, I felt that this approach would not provide a full picture of the impact of the ‘Live Active’ scheme. At the conclusion of my literature review, most likely due to my philosophical stance, I perceived that participants’ voices had been lost in the majority of previous ERS evaluations. As an ERS is a real world intervention developed for and participated in by members of the community, it was my view that I could not make an evaluation of this ERS without attempting to capture the experiences of these individuals.

For pragmatists, research starts with a problem, and ideally ends with practical solutions to inform future practice. Theories, hypotheses, and results are therefore considered in terms of their practical consequences (Saunders, Lewis, Thornhill, & Bristow, 2015). It was my view that a purely quantitative approach to this evaluation would limit the production of actionable information to inform future practice in ERSs. Furthermore, in the context of ERSs, numerous positivist evaluations have led researchers to question whether or not the funding of ERSs should be continued. It was as a result of my values, that I came to the view that positivist evaluations would rarely capture the true impact of a real world intervention, and continued predominance of these studies could lead to the demise of a potentially successful scheme. Interpretivism has therefore clearly contributed to the development of study three. However, from a pragmatic point of view, there was a challenge of working within the constraints of secondary data, stakeholder requests, and the need to produce a coherent evaluation that advances knowledge, whilst working within these philosophical values. Therefore, whilst the studies within this thesis are primarily quantitatively driven, and more congruent with positivist values, this was done so from a pragmatic stance. Aside from study three, interpretivism is alternatively present in the interpretation of the findings of these quantitative studies, which are carefully considered and reflected upon, and not merely accepted as observable, unambiguous facts.

Pragmatism and interpretivism have therefore influenced my research, by driving a research strategy that seeks to comprehensively evaluate the ‘Live Active’ scheme, whilst working within several constraints. Answers to real world problems are therefore reached using mixed methods, with considered reflection of the quantitative findings, acknowledging that the values and methodologies of researchers can unintentionally influence the future of ERSs.

4.5. Research Ethics Approval

Ethical approval for the evaluation of secondary data (HSCR15/124) was sought and granted by the University of Salford Research, Innovation and Academic Engagement Ethical Approval Panel in December 2015, with a minor amendment granted to the researcher in August 2016. This was for the purpose of studies 1 and 2. Further ethical approval was sought and approved by the panel for two further studies: the perceived benefits of participation in an ERS (HSR1617-67), granted March 2017, which is the basis of chapter 7, and; agreement between accelerometer assessed and self-reported PA and sedentary behaviour, in a sample of ERS participants (HSR1617-58), granted March 2017, which is the basis of chapter 8.



Overview of the research framework (study 1)

Chapter 5 Study 1: Evaluation of uptake and adherence to the 'Live Active' exercise referral scheme

5.1. Study 1 – Introduction, aims and objectives

An awareness of uptake and adherence to an ERS is important for stakeholders, in order to understand how well the scheme attracts and retains participants. Continued retention is also important to increase the potential for sustained behaviour change and positive health outcomes, and also for evaluators to be able to make firm conclusions about the effectiveness of schemes based on reasonable follow-up rates (Morgan et al., 2016; Pavey et al., 2012; Williams et al., 2007).

Low uptake and adherence to ERSs is a common problem, evidenced by reoccurring low figures and a growing body of literature aimed at understanding how to facilitate better outcomes (Morgan et al., 2016; Pavey et al., 2012; Williams et al., 2007). The most recent detailed figures from the systematic review of Pavey et al. (2012), analysing the uptake and adherence rates of published schemes, show an average uptake rate of 66-81%, and an average adherence rate of 43%-49%. These figures, however, are only a representation of ERSs with published evaluations, including controlled and observational studies, and not of real world ERSs across the nation.

In addition to reporting rates of uptake and adherence, systematic reviews have synthesised the evidence on predictors of uptake and adherence from both quantitative (Campbell et al., 2015; Pavey et al., 2012; Williams et al., 2007) and qualitative research (Morgan et al., 2016). Despite this previous research, continued exploration of uptake and adherence is important, as the existing evidence base is still relatively small and equivocal, with research gaps remaining and growth in knowledge hindered by the inconsistencies of previous research (NICE, 2014a).

The aim of study 1 was to evaluate uptake and adherence to the 'Live Active' ERS, whilst addressing some of the gaps and limitations of previous research. Firstly, the study reports rates of non-uptake, uptake, dropout and completion. Characteristics (e.g. demographics) are also reported for all participants referred to the service, those classified as uptake or non-uptake, and as having dropped out or completed the scheme. This builds upon previous research by detailing participation rates in under-represented groups such as ethnic minorities and disabled persons.

The reasons for non-uptake and dropout are reported, and any inappropriate referrals identified. The study also places an equal focus on both uptake and adherence, with exploration of which characteristics can predict these outcomes, furthering knowledge in this area by exploring a wider range of characteristics than previous evaluations, such as ethnicity, SES, and disability status. Furthermore, the study determines at which stage the highest rate of dropout occurred. Collectively, these objectives will identify areas for improvements in practice, and any further research required to do so.

Study 1 aim: Evaluate uptake and adherence to the scheme, using analysis of secondary data.

Study 1 objectives:

1.1. Identify rates of uptake, non-uptake, dropout and scheme completion.

1.2. Report the characteristics of participants who are referred to the scheme, uptake or do not uptake the referral, and dropout or complete the scheme (e.g. age, gender, SES, disability status, referral source, referral reason).

1.3. Identify at which time point the majority of dropout occurs.

1.4. Report the reasons for non-uptake and dropout.

1.5. Identify variables (e.g. age, gender, SES, source of referral, reason for referral) that can predict uptake and adherence.

5.2. Literature review of uptake and adherence to ERSs

5.2.1. Factors associated with uptake

There is a general consensus that uptake can be defined as the proportion of referred participants who accept their referral and initially attend the scheme (Campbell et al., 2015; Pavey et al., 2012). Several characteristics have been consistently implicated in the exploration of uptake to an ERS, which will be discussed below.

A considerable body of literature has recognised that females are more likely to uptake than males (Dugdill et al., 2005; Hanson et al., 2013; Lord & Green, 1995; Moore, Raisanen, Moore, Din, & Murphy, 2013; Sowden, Breeze, Barber, & Raine, 2008), with the exception of the findings by Harrison, McNair, and Dugdill (2005a) and Gidlow et al. (2007), who found no association between gender and uptake. To-date, there is no clear hypothesis to explain this association.

Increasing age has also been frequently associated with the likelihood of uptake (Dugdill et al., 2005; Gidlow et al., 2007; Hanson et al., 2013; Lord & Green, 1995; Moore et al., 2013; Sowden et al., 2008). However, this appears to be with the exception of the very old (≥75 years), who have uptake rates similar to younger aged participants (<35 years) (Hanson et al., 2013; Harrison et al., 2005a; Lord & Green, 1995; Sowden et al., 2008).

To date, studies exploring the association between the reason for referral and uptake have produced equivocal results, which is perhaps due to the variation in reporting of medical conditions, and varying inclusion criteria of different schemes (Pavey et al., 2012). The equivocal nature of findings can be seen in the example of medical conditions such as obesity (James et al., 2008; Lord & Green, 1995), musculoskeletal conditions (James et al., 2008; Lord & Green, 1995; Sowden et al., 2008), and mental health conditions (Crone, Johnston, Gidlow, Henley, & James, 2008; Harrison et al., 2005a; James et al., 2008; Lord & Green, 1995; Moore et al., 2013), which have been associated with uptake in several studies but non-uptake in others. The profession of the referrer has also been explored, although to a lesser extent, with James et al. (2008) finding higher uptake rates in those referred by a GP compared to an allied health professional (AHP), but Hanson et al. (2013) finding no relationship between referrer profession and uptake. A hypothesis for difference in uptake rates between referrer professions has not been explored to-date.

There is an interest in the association between uptake and under-represented groups using demographic variables such as ethnicity, SES and disability status, but these data are often not collected, and hence is an area for future research for ERSs that do collect this information (Hanson et al., 2013; Harrison et al., 2005a; Taylor et al., 1998). In an example using SES, one of the more frequently reported variables, Hanson et al. (2013) and Moore et al. (2013) reported lower uptake rates for participants from areas of higher deprivation, whilst Harrison et al. (2005a) and Sowden et al. (2008) found no relationship between uptake and SES. These conflicting results are interesting given the strong SE patterns in PA in England (Farrell, Hollingsworth, Propper, & Shields, 2014), which would lead one to expect a positive relationship between increasing SES and likelihood of uptake.

In comparison to quantitative investigation, there has been very little qualitative exploration of why participants uptake a referral to an ERS or not. Stathi, McKenna, and Fox (2004), for example, briefly noted that a GP’s recommendation to be referred was important to older persons’ decisions to uptake. Growth in knowledge has therefore been limited by a focus on quantitative investigation of uptake, and where this has been conducted, by the narrow focus on characteristics that can predict uptake, such as age and gender. Quantitative exploration of uptake, including a broader range of potential demographic factors, such as ethnicity and employment status, would identify any other significant variables that could help to explain uptake. This may also identify areas for follow-up qualitative research, which may also identify factors influencing uptake which are unrelated to these demographics. Collectively, these findings may identify areas for improvements in practice in ERSs, which can maximise uptake.

5.2.2. Factors associated with adherence

One hinderence to the evaluation of adherence to ERSs has been the inconsistent reporting of key terms. Definitions of 'adherence' vary between evaluations, causing difficulty in comparing adherence between schemes (Pavey et al., 2012). Some evaluations have understood adherence to be based on level of attendance at consultations, whilst others have considered it to be based on attendance at leisure centre-based exercise sessions (see Table 5.1 below). The latter fails to recognise that not all participants will wish to undertake exercise class-based activities, potentially mistaking those being independently active elsewhere as non-adherent. If adherence is to be based upon PA participation, it should therefore be specifically measured according to the PA prescription agreed between the participant and exercise professional. Further difficulty has been encountered in understanding the level of attendance required to qualify as an 'adherer', with a review suggesting that adherence is attendance at ≥75% of scheduled contact points (Pavey et al., 2012).

Table 5.1. Applied definitions of adherence.

Authors	Definition of adherence
Hanson et al., 2013	Attended the week 12 consultation
Isaacs et al., 2007; Taylor et al., 1998	Number of prescribed exercise sessions attended
James et al., 2009	Attended ≥80% of exercise sessions
Jones, Harris, Waller, & Coggins, 2005	Attended 24 exercise sessions
Lord & Green, 1995	Attended week 10 consultation and still exercising
Morton et al., 2008	Attended at least one exercise session per week for the duration of the study (10-12 weeks)
Tobi et al., 2012	Attendance at both the 13th and 20-26th week consultations

However, irrespective of definition, it is now well established from a variety of studies that, while females are more likely than males to uptake a referral (Dugdill et al., 2005; Hanson et al., 2013; Lord & Green, 1995; Moore et al., 2013; Sowden et al., 2008), males are more likely than females to adhere to and complete an ERS programme (Dugdill et al., 2005; Gidlow et al., 2007; James et al., 2009; Lee, Griffin, & Simmons, 2009). There is no clear hypothesis for the positive association between females and uptake, but negative association between females and adherence, and this is therefore an area for future research to further understanding.

There is also strong agreement that the likelihood of adherence and ERS completion rises with increasing age (Dugdill et al., 2005; Gidlow et al., 2007; Hanson et al., 2013; James et al., 2009; Kelly et al., 2017; Mills et al., 2012; Sowden et al., 2008; Tobi et al., 2012), with the exception of Lee et al. (2009) who found no association. Hypotheses to explain these quantitative findings suggest that older aged participants have more free time to incorporate PA in to daily life (Tobi et al., 2012), and value the opportunity for social interaction (James et al., 2009; Tobi et al., 2012).

However, a higher likelihood of adherence in older aged participants, may in some cases be due to the perception of safety when being active in a supervised environment, such as a PA class, for those who otherwise lack confidence to be active alone (James et al., 2009; Mills et al., 2012). This may be detrimental in the long term, causing a potential dependence upon the scheme and its supervised activities, rather than encouraging long-term independent PA (Mills et al., 2012).

As with the association with uptake, there is great variation in the relationship between referral conditions and likelihood of adherence. Emerging evidence suggests that the likelihood of dropout is higher for those referred for mental health conditions (Crone et al., 2008; Dugdill et al., 2005; Kelly et al., 2017; Tobi, Kemp, & Schmidt, 2017) and respiratory conditions (James et al., 2009; Mills et al., 2012). Several studies have reported lower likelihood of dropout amongst those referred for CVD (Dugdill et al., 2005; James et al., 2009; Kelly et al., 2017; Mills et al., 2012; Sowden et al., 2008), with the exception of Tobi et al. (2012). The evidence regarding musculoskeletal conditions (Kelly et al., 2017; Tobi et al., 2012), and diabetes (Sowden et al., 2008; Tobi et al., 2012) is equivocal between studies, and evidence for the association between obesity and dropout is minimal (Hanson et al., 2013). Hypotheses to explain to these findings suggest that the symptoms of health conditions, such as respiratory conditions, can create uncomfortable PA experiences, and slow down the attainment of improvements (James et al., 2009; Tobi et al., 2012). Qualitative investigation supports these hypotheses, and suggests that for people referred for respiratory or musculoskeletal conditions, dropout may be due to the varying and unpredictable nature of condition related symptoms, and perceived restriction of PA by symptoms such as breathlessness, pain or functional ability, and a feeling of being unable to progress whilst experiencing these symptoms (Mills et al., 2012). Furthermore, Tobi et al. (2017) suggest that low rates of adherence in those referred for mental health conditions may be due to the nature of the condition, with low motivation and a tendency to avoid social situations. However, the authors also suggest that analysis of the association between adherence and referral conditions has been difficult, due to low numbers of referrals received for groups such as mental health conditions.

There has been little exploration or analysis of how demographic variables, such as SES, and under-represented groups such as ethnic minorities and disabled persons, are associated with adherence, although this can be partly explained by lack of available data due to these variables not being routinely collected by schemes (Hanson et al., 2013; Sowden et al., 2008; Tobi et al., 2012). To date, evidence suggests that adherence is greater in those from less deprived areas (Gidlow et al., 2007; Hanson et al., 2013), with the exception of Sowden et al. (2008) who found no association. Furthermore, those of mixed ethnicity appear least likely to dropout (James et al., 2009; Mills et al., 2012), and thus far there appears to be no association between disability status and adherence (Kelly et al., 2017).

Few identified studies have explored the association between lifestyle behaviours, such as smoking and alcohol consumption, with adherence. Two identified studies reported a higher likelihood of dropout amongst smokers (Kelly et al., 2017; Ward, Phillips, Farr, & Harries, 2010), whilst one identified identified a lower likelihood of dropout amongst alcohol drinkers (Kelly et al., 2017). Negative health behaviours, such as smoking and physical inactivity, are known to cluster (Kaczynski, Manske, Mannell, & Grewal, 2008; Saint Onge & Krueger, 2017; Spring, Moller, & Coons, 2012). The physiological impacts of smoking can also impair physical capabilities, and therefore the likelihood of being active (Kaczynski et al., 2008). It is therefore important to establish if groups with negative health behaviours, such as current smokers, are less likely to adhere to an ERS, or require additional support to become physically active.

To summarise the available evidence that has been identified using quantitative approaches, growth in knowledge has been limited by the narrow focus on characteristics that can predict adherence such as age and gender, and to some extent referral condition. Further exploration of adherence, including a broader range of potential demographic factors, such as ethnicity and employment status, would identify any other significant variables that could help to explain engagement. These findings may identify areas for improvements in practice in ERSs, to make schemes more appropriate for different groups (Hanson et al., 2013), and to maximise adherence.

As noted, exploration of adherence has been explored using both quantitative and qualitative methods. Qualitative research supports a range of intrinsic and extrinsic factors that facilitate adherence (Morgan et al., 2016). Many of these factors relate to characteristics of the scheme, and therefore extrinsic factors, such as: the quality and location of leisure facilities; the gym environment; the range of PA options available, including their price and time, and; the professionalism, personality and support given by ERS professionals (Morgan et al., 2016).

Social features of schemes, such as socialisation and peer support, have not typically been considered to be important outcomes of participation. However, social support from friends, family and peers, and perceived social inclusion within schemes appears to facilitate adherence (Campbell et al., 2015; Morgan et al., 2016; Williams et al., 2007). Social support may contribute to motivation and confidence for PA (Crone et al., 2005), whilst the development of social networks within schemes may promote confidence and enjoyment that is conducive to adherence (Crone et al., 2005; Martin & Woolf-May, 1999; Wormald & Ingle, 2004).

The professionalism, support and supervision from ERS professionals is a frequently cited facilitator for adherence (Morgan et al., 2016). Functions of this facilitator include: support from professionals to master physical activities and equipment, and advice on suitability of exercises (Crone et al., 2005; Hardcastle & Taylor, 2001); providing encouragement to sustain the participant's motivation and increase confidence in a new environment (Crone et al., 2005; Hardcastle & Taylor, 2001; Martin & Woolf-May, 1999; Stathi et al., 2004; Wormald & Ingle, 2004), and; providing a sense of safety whilst being active in a supervised environment (Mills et al., 2012). The latter, in addition to a desire to sustain professional support beyond the scheme (Wormald et al., 2006), are facilitators that may create a potential dependency on the scheme and professionals, which cannot be sustained in the long term (Mills et al., 2012).

Intrinsic factors, such as perceived improvements in health and wellbeing appear to facilitate adherence (Queen et al., 2015), with exacerbation of health issues, or poor progress towards goals such as weight loss impeding adherence (Morgan et al., 2016). A study by Queen et al. (2015) found that participants who had improvements in self-reported health throughout the scheme adhered, whilst those who did not dropped out. Collectively, these findings highlight the need for ERS professionals to set realistic goals and expectations of what can be achieved during an ERS.

The predominant focus of existing literature relates to why people adhere to as opposed to why people dropout of an ERS (Gidlow et al., 2005; Morgan et al., 2016; Pavey et al., 2012). In order to inform practices which can increase adherence to an ERS, research must also identify factors leading to non-uptake and dropout, rather than inferring this information from the views of those that uptake and adhere (NICE, 2014a).

5.2.3. Limitations of research on uptake and adherence

To date, research exploring the factors leading to uptake and adherence has been limited by a focus on a small number of potential predictors, such as age and gender (Pavey et al., 2012). There has also been a predominance to explore adherence more so than uptake (Morgan et al., 2016). In comparison to adherence, there has been limited quantitative or qualitative exploration of non-uptake, including identification of the reasons for non-uptake, and follow up qualitative investigation to understand these reasons.

Where research has explored adherence, the predominant focus has been on the characteristics of the scheme that promote adherence, or demographics such as age and gender, with less attention paid to the effect of individual characteristics such as SES and ethnicity (Gidlow et al., 2005; Morgan et al., 2016; Pavey et al., 2012). However, this narrow focus is partly explained by a lack of data regarding individual characteristics, due to poor participant profiling by ERSs (Hanson et al., 2013; Sowden & Raine, 2008; Tobi et al., 2012). Where associations have been found, such as between young age and low adherence, there is insufficient exploration of why this association exists, such as through follow-up qualitative investigation; instead the conclusion is often made that ERSs are potentially not suitable for younger age groups (i.e. <35 years) (James et al., 2009). Further exploration of associations could facilitate improvements in ERS pathways that are more appropriate for these groups, as opposed to accepting that ERSs are unsuitable for them and potentially excluding them from participation (Hanson et al., 2013).

An area that has been highlighted for future research, is exploration of non-uptake and dropout, as opposed to uptake and adherence, including the reasons for these events (Morgan et al., 2016; NICE, 2014a; Pavey et al., 2012). Stakeholders would also benefit from knowledge of when the majority of dropout occurs, and why it occurs more frequently at certain time points if this is the case. Previous research, for example, indicates that the highest percentage of dropout occurs in the first six weeks, but the reasons for this are unknown (Kelly et al., 2017; Moore et al., 2013), Knowledge in this area is important, as it can inform the refinement of practice in existing, and future ERSs, to maximise participation (Morgan et al., 2016; NICE, 2014a; Pavey et al., 2012).

A further limitation of previous research is the tendency to assume that all non-uptake and dropouts are negative events, with these individuals remaining physically inactive (Pavey et al., 2012). Whilst there is in fact evidence that some participants who dropout of ERSs subsequently become independently active, unfortunately this is overshadowed by the masses that drop out for more typical reasons such as time and motivation (Harrison et al., 2005a; Pavey et al., 2012). Furthermore, there is little recognition that non-uptake and dropout figures can be negatively inflated by poor referral processes, such as referral without full explanation of what an ERS is (Pavey et al., 2012), or no explanation of the benefits of participation (Stathi et al., 2004), leading to loss of uptake. Inappropriate referrals have also been received for those with medical contraindications to PA (Johnston, Warwick, De Ste Croix, Crone, & Sldford, 2005; Pavey et al., 2012), and/or having no motivation to change (Dugdill et al., 2005; Moore, Moore, & Murphy, 2011). This highlights the value of understanding the reasons for non-uptake and dropout, when interpreting these outcomes.

Arsenijevic and Groot (2017) argue that substantial heterogeneity between scheme characteristics, such as duration, and target referral groups, may introduce a source of bias when analysing the effectiveness of ERSs, if variations in such factors are associated with better outcomes, including process measures such as uptake and adherence. The authors suggest that for comparison of outcomes and process measures, schemes should only be compared to those of a similar nature, such as programme length and inclusion criteria. Further to this, Tobi et al. (2017) argue that the local policy context in which an ERS is developed may influence process outcomes, such as a strategic aim to target referrals with health inequalities, resulting in higher proportions of participants from these groups, compared to other schemes with alternative local policy-related aims.

5.2.4. Study 1 – summary

Low uptake and adherence to ERSs is a common problem (Pavey et al., 2012; Williams et al., 2007). A growing body of literature has sought to understand which factors can predict uptake and adherence, and how to facilitate these outcomes (Morgan et al., 2016; Pavey et al., 2012; Williams et al., 2007). However, further research is required to understand participation rates in under-represented groups, which factors can predict uptake and adherence, and the reasons for non-uptake and dropout (Morgan et al., 2016; NICE, 2014a; Pavey et al., 2012).

5.3. Study 1 - Methods

This section provides a description of the methods used to address the aims and objectives detailed in section 5.1.

This study was a quantitative process evaluation of uptake and adherence to the ‘Live Active’ ERS. Process evaluations are concerned with understanding the functioning of an intervention, such as its reach, maintenance, implementation and context (Baranowski & Stables, 2000; Moore et al., 2015). Recommended outcomes of process evaluations in PA interventions include the number of participants recruited, the number of participants attending, and the reasons for opt-out (Cavill et al., 2012).

The process measures analysed in this study were:

- Number of referrals received.
- Number of referrals that did not uptake referral, did uptake a referral, had dropped out, or had completed.
- Frequency of dropout by time point (i.e. week 6, 12, 24).
- Reasons for non-uptake or dropout.
- Predictors of uptake and adherence.

In this study, uptake referred to initial attendance, and adherence referred to the duration of participation, described in terms of completion or dropout (Tobi et al., 2012). Therefore, referrals were classified as:

- Non-uptake: participant chose not to schedule an initial consultation, or scheduled an initial consultation but did not attend.
- Participating: attended an initial consultation, but had not yet reached the week 24 consultation.
- Dropped out: attended an initial consultation, but chose to leave the scheme before week 24.
- Completed: attended both an initial consultation and week 24 consultation.

5.3.1. Data management

Anonymised data for referrals made to the scheme between July 2015 - July 2017 were extracted from the ERS database for secondary analysis. The advantages and disadvantages of secondary data were previously discussed in section 4.4.1.

Exclusions were made for new referrals awaiting an initial consultation. From a total of 2104 referrals, after exclusions, this left 1987 participants eligible for analysis.

Referral and participant characteristics, along with outcomes collected during the initial consultation, were extracted from the ERS database for participants included in the analysis. Data were cleaned and checked for potential incorrect entries, with any missing information coded as missing. Potential incorrect entries were verified with the exercise professionals who had input the data.

For the purpose of Chi-square tests of association, and binary logistic regression, referral and participant characteristics were categorically coded, using categories defined by previous literature were possible. Age, for example, was coded in 10-year age bands for consistency with previous literature (e.g. Hanson et al., 2013; Harrison et al., 2005a; Kelly et al., 2017). Referrer organisation was categorised according to the scope of the organisation, i.e. primary care or secondary care/community service (e.g. Kelly et al., 2017). Primary referral condition was categorised according to the ERS's existing condition pathways (e.g. Hanson et al., 2013; Harrison et al., 2005a; Kelly et al., 2017), with the addition of the category 'multiple', for people referred specifically for 3 or more conditions. Categorical outcomes collected during the initial consultation (TTM stage of change, smoking status and alcohol consumption) were coded, whilst all other outcomes were retained as continuous outcomes. These outcomes were critically discussed in section 3.5.

5.3.2. Statistical analyses

Statistical analyses were performed using IBM SPSS version 23 (SPSS, Inc., Chicago, IL, USA). Attendance at each scheduled consultation (week 1, 6, 12, 24) was categorically coded, and used to identify at which time point participants dropped out. Reason for non-uptake or dropout, supplied by the participant at the time of the event and recorded in the ERS database, was categorised according to theme and explored using frequency counts. Association between time point of dropout and reason for dropout was explored using a chi-square test (χ^2).

Demographics of referrals were explored using descriptive statistics such as means and frequency counts. Differences in referral and demographic characteristics between individuals who were classified as uptake or non-uptake, and completers or dropouts were explored using a chi-square (χ^2) test for categorical variables, and a t-test or Mann-Whitney U test for continuous variables. Differences in baseline assessment outcomes, collected during the initial consultation, were also explored between completers and dropouts using the same tests.

Two binary logistic regressions were used to explore if any referral or demographic characteristics could predict uptake (regression 1), or if any referral or demographic characteristics, or any baseline assessment outcomes collected during the initial consultation, could predict adherence (regression 2). Logistic regression was used as it is a predictive analysis that allows the relationship between multiple continuous or categorical independent variables and one dichotomous dependent variable, such as participation status, to be explored (Field, 2017). Goodness of fit was examined using the Hosmer & Lemeshow test. Tolerance and variance inflation factors were inspected for multicollinearity.

5.4. Study 1 - Results

This section reports the results of data analyses, aiming to address the aims and objectives detailed in section 5.1. Uptake and adherence rates are reported, followed by the reasons for non-engagement, and predictors of uptake and adherence.

5.4.1. Number of referrals

Between July 2015 and July 2017, 2104 referrals were made to the scheme. Of these, 117 were new referrals awaiting an initial consultation, and have been excluded from further analysis. Of the remaining 1987 referrals, 1609 (81%) accepted their referral, whilst 378 (19%) did not uptake their invitation to participate, equivalent to an uptake rate of 81%. Of the 1609 participants who entered the scheme, 515 (32%) were still participating (between week 1 and 24), 623 (38.7%) had completed, and 471 (29.3%) had dropped out before week 24. When considering adherence in terms of completion and dropout, the equivalent adherence rate was 56.9% (n=623 completed, n=471 (43.1%) dropped out).

5.4.2. Demographics of referrals

Referrals were predominantly female (60.2%, n=1197), Caucasian (82.3%, n=1635), and from the most deprived quintile (Q1, 39.9%, n=793), with a mean (± standard deviation (SD)) age of 55.4 years (±14.8). A total of 128 (8%) referrals were known to have a disability. The three most common sources of referrals were medical practices (44.8%, n=890), physiotherapy services (17.1%, n=340), or an NHS health trainer service (15.5%, n=308). The most common reasons for referral were for having: multiple (≥3) health conditions (33.7%, n=669), musculoskeletal related conditions (23.5%, n=467), existing CVD or significant risk factors (15.1%, n=300), or for being overweight or obese (12.5%, n=248). Characteristics of participants who were referred to the scheme, did uptake the referral, did not uptake the referral, completed the scheme, or dropped out early are presented below in Table 5.2.

Table 5.2. Demographics of all referrals, and participants who did uptake the referral, did not uptake the referral, completed the scheme, or dropped out early from the scheme.

	All referrals		Uptake		Non-uptake		6-month completers		Dropouts	
	n	%	n	%	n	%	n	%	n	%
Gender										
Female	1197	60.2	959	59.6	238	37	344	55.2	286	39.3
Male	790	39.8	650	40.4	140	63	279	44.8	185	60.7
Age										
Mean (SD)	55.4 (14.8)		55.7 (14.8)		54 (15.1)		57.2 (14.4)		54.4 (14.8)	
<65 years	1345	67.7	1094	68	251	66.4	395	63.4	339	72
≥65 years	592	29.8	501	31.1	91	24.1	222	35.6	127	27
Missing	50	2.5	14	0.9	36	9.5	6	1	5	1.1
Ethnicity										
Caucasian	1635	82.3	1505	93.5	130	34.4	590	94.7	443	94.1
Asian	76	3.8	72	4.5	4	1.1	24	3.9	21	4.5
Mixed	6	0.3	6	0.4	0	0	3	0.5	1	0.2
Black	18	0.9	18	1.1	0	0	5	0.8	4	0.8
Other	8	0.4	8	0.5	0	0	1	0.2	2	0.4
Unknown	244	12.3	0	0	244	64.6	0	0	0	0
Employment status										
Full-time employed	373	18.8	369	22.9	4	1.1	132	21.2	119	25.3
Part-time employed	125	6.3	123	7.6	2	0.5	57	9.1	25	5.3
Unemployed	240	12.1	235	14.6	5	1.3	94	15.1	79	16.8
Retired/other	886	44.6	882	54.8	4	1.1	340	54.6	248	52.7
Unknown	363	18.3	0	0	363	96	0	0	0	0
Quintile of deprivation										
1 (most)	793	39.9	613	38.1	180	47.6	228	36.6	204	43.3
2	551	27.7	455	28.3	96	25.4	180	28.9	124	26.3
3	270	13.6	221	13.7	49	13	94	15.1	63	13.4
4	248	12.5	223	13.9	25	6.6	82	13.2	56	11.9
5 (least)	64	3.2	55	3.4	9	2.4	25	4	9	1.9
Unavailable	61	3.1	42	2.6	19	5	14	2.2	15	3.2
Has a disability										
Declared	-	-	128	8	-	-	48	7.7	36	7.6
Non-declared	-	-	1481	92	-	-	575	92.3	435	92.4
Source of referral										
Medical Practice	890	44.8	714	44.4	176	46.6	262	42.1	222	47.1
Health trainer service	308	15.5	264	16.4	44	11.6	117	18.8	79	16.8
Physiotherapy service	340	17.1	275	17.1	65	17.2	72	11.6	76	16.1
Weight management service	182	9.1	149	9.3	33	8.7	81	13	44	9.3
Other service	267	13.5	207	12.8	60	15.9	91	14.5	50	10.7

	All referrals		Uptake		Non-uptake		6-month completers		Dropouts	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Primary referral condition										
CVD/High risk of CVD	300	15.1	234	14.5	66	17.5	112	17.9	63	13.4
Diabetes/metabolic syndrome	76	3.8	58	3.6	18	4.8	27	4.4	16	3.4
Mental health issue	126	6.3	99	6.2	27	7.1	31	5	38	8.1
Respiratory condition	48	2.4	45	2.8	3	0.8	21	3.4	15	3.2
Musculoskeletal condition	467	23.5	392	24.4	75	19.8	112	18	106	22.5
≥Overweight	248	12.5	203	12.6	45	11.9	99	15.9	58	12.3
Other condition	53	2.7	45	2.8	8	2.1	15	2.4	13	2.8
Multiple conditions	669	33.7	533	33.1	136	36	206	33.1	162	34.4

5.4.3. Time of dropout

Of the 29.3% (n=471) of participants who dropped out, 52% left before week 6 (n=245), 33.8% left before week 12 (n=159), and 14.2% left before week 24 (n=67).

5.4.4. Reasons for non-uptake or dropout

Of the 19% (n=378) of participants who did not up-take a referral, 14 different reasons emerged for doing so. The most common reason for non-uptake, was that despite repeated attempts, the ERS could not establish contact with the referred individual (n=140, 37%). This included participants the ERS had initially established contact with, but requested a call-back, which was not answered.

Common reasons for non-uptake also included: the participant perceived they could not be active due to their current health status (n=51, 13.5%), and; the participant was not interested in participating (n=43, 11.4%). A full list of reasons for non-participation are presented below in Table 5.3.

Of the 29.3% (n=471) of participants who dropped out, 14 different reasons emerged for doing so. The most common reasons were: unable to establish contact with the participant, despite multiple attempts using multiple methods (phone call, text message, mail) over a 3-month period (n=182, 38.6%); participants perceived they could not be active due to their current health status (n=71, 15.1%), and; participants had a new contraindication to PA (n=54, 11.5%). A full list of reasons for dropout are presented below in Table 5.3.

Table 5.3. Reasons for classification as non-uptake or dropout from the 'Live Active' ERS.

	Non-uptake		**Dropout**	
Reason	**n**	**%**	**n**	**%**
Unable to establish contact with the participant	140	37	182	38.6
Perceived they could not be active due to their health status	51	13.5	71	15.1
Was not interested in participating	43	11.4	32	6.8
Had no time to participate	35	9.3	49	10.4
Had a contraindication to PA	27	7.1	54	11.5
Reported they were already physically active	27	7.1	-	-
DNA or cancelled ≥3 initial appointments	18	4.7	-	-
DNA or cancelled ≥3 appointments	-	-	11	2.3
Accessibility issues (e.g. transport, finance)	13	3.5	2	0.4
Wanted to become active independently	11	2.9	-	-
Relocated out of the area	3	0.8	7	1.5
Resided outside of the catchment area	3	0.8	-	-
Wanted to participate at a later date	3	0.8	1	0.2
Wanted to attend an alternative service (e.g. dietary)	3	0.8	2	0.4
Participant died	1	0.3	2	0.4
Became independently active	-	-	33	7
Participant requested to be discharged without reason	-	-	23	4.9
Did not perceive any benefits from increasing PA	-	-	2	0.4

DNA, did not attend appointment.

There was a significant association between time of dropout and reason for dropout ($\chi^2(26) = 41.2$, $p \leq 0.05$). The most frequent reasons for dropout were similar at each time point. However, participants who dropped out between weeks 6 and 12 left for a greater variety of reasons than those who left in the first 6 weeks, or between weeks 12 and 24 (Table 5.4, below). Inspection of the standardised residuals indicated that dropout before week six, due to lack of interest, significantly contributed to the overall chi-square statistic (z = 2.5).

Table 5.4. Distribution of reasons for classification as dropout from the 'Live Active' ERS.

	<Week 6		<Week 12		<Week 24	
Reason	**n**	**%**	**n**	**%**	**n**	**%**
Unable to establish contact with the participant	*85*	34.7	*65*	40.9	*32*	47.8
Perceived they could not be active due to their health status	*35*	14.3	*22*	13.8	*14*	20.9
Had a contraindication to PA	*25*	10.2	*21*	13.2	*8*	11.9
Had no time to participate	*24*	9.8	*21*	13.2	*4*	6
Became independently active	*24*	9.8	*7*	4.4	*2*	3
Was not interested in participating	*27*	11	*4*	2.5	*1*	1.5
Participant requested to be discharged without reason	*13*	5.3	*5*	3.1	*5*	7.5
DNA or cancelled ≥3 appointments	*6*	2.4	*4*	2.5	*1*	1.5
Relocated out of the area	*4*	1.6	*3*	1.9	-	-
Did not perceive any benefits from increasing PA	*1*	0.4	*1*	0.6	-	-
Accessibility issues (e.g. transport, finance)	*1*	0.4	*1*	0.6	-	-
Wanted to attend an alternative service (e.g. dietary)	-	-	*2*	1.3	-	-
Participant died	-	-	*2*	1.3	-	-
Wanted to participate at a later date	-	-	*1*	0.6	-	-

DNA, did not attend appointment.

5.4.5. Characteristics associated with uptake

There were significant associations between uptake and quintile of deprivation (χ^2(4, n=1926) = 22.1, *p* ≤0.001), and season of referral to the scheme (χ^2(3, n=1987) = 9.3, *p* ≤0.05). Uptake was higher amongst participants from less deprived quintiles (Q1 77.3%, Q2 82.6%, Q3 81.9%, Q4 89.9%, Q5 85.9), and those who were referred in the summer months (85.9%), compared to spring (81.3%), winter (79%) or autumn (78.8%). Inspection of the standardised residuals indicated that lower uptake by those residing in Q1 (z = -3.8), and higher uptake by those residing in Q4 (z = 3.7), significantly contributed to the overall chi-square statistic. Higher uptake by those referred in the summer months was also a significant contributor (z = 2.8). There were no significant associations between uptake and age group, gender, ethnicity, employment status, referrer organisation type, referrer profession or primary referral condition (*p* >0.05). Association between non-uptake and disability status could not be examined due to lack of data available from the referral form.

5.4.6. Predictors of uptake

Binary logistic regression was undertaken to explore the variation in participation status. Variables included in the binary logistic regression are presented below in Table 5.5. Ethnicity (missing n=244) and employment status (missing n=363) were excluded due to missing data. A test of the model including these variables was non-significant. A further 109 cases were excluded due to missing data in other fields, leaving 1878 cases included in the model. A test of the model was statistically significant (Hosmer-Lemeshow: $\chi^2(8) = 2.9$, $p > 0.05$), indicating that it was able to distinguish between participants who did or did not uptake a referral. However, inspection of the classification table indicated that inclusion of the variables did not improve upon the percentage of correctly classified cases observed in the null model (82.7%). Therefore, this combination of variables did not perform well at predicting uptake. However, as shown in Table 5.5, there was an association between uptake and two independent variables, quintile of deprivation and season of referral. Compared to the most deprived quintile, residing in areas of moderately higher (quintile 2) and moderately lower deprivation (quintile 4) was associated with uptake. Referral to the scheme in the summer months, when compared to winter, was also associated with uptake.

Table 5.5. Logistic regression to predict uptake (n=1878), and the % of participants per category who did uptake.

Predictor	% uptake	n	Adj OR	95% CI	*p*
Age					0.201
75+	83.7	172	1.000 (ref)		
65-74	85.4	405	1.135	0.687-1.876	0.621
55-64	84.9	471	1.139	0.696-1.864	0.604
45-54	80.7	389	0.833	0.505-1.374	0.474
35-44	78	241	0.678	0.394-1.168	0.161
25-34	81.5	173	0.862	0.474-1.570	0.628
18-24	77.8	27	0.676	0.239-1.914	0.461
Quintile of deprivation					0.003
Q1 (most deprived)	78.9	769	1.000 (ref)		
Q2	83.9	535	1.377	1.025-1.848	0.034
Q3	83.1	266	1.224	0.842-1.778	0.289
Q4	90.6	245	2.396	1.497-3.834	0.000
Q5 (least deprived)	87.3	63	1.788	0.824-3.880	0.141
Gender					
Female	81.6	1144	1.000 (ref)		
Male	84.5	734	1.231	0.944-1.605	0.125
Referrer organisation					
Other	83.7	1030	1.000 (ref)		
Medical practice	81.6	848	0.932	0.672-1.291	0.670
Referrer profession					0.385
AHP	83.8	1011	1.000 (ref)		
Nurse	80.7	456	0.810	0.587-1.118	0.200
Doctor	82.5	411	0.979	0.658-1.457	0.916
Primary referral condition					0.322
Multiple conditions	81.5	644	1.000 (ref)		
CVD or high risk of CVD	81.5	271	0.826	0.561-1.216	0.332
Diabetes/metabolic syndrome	79.7	69	0.870	0.458-1.652	0.671
Respiratory condition	95.6	45	4.361	1.032-18.427	0.045
Musculoskeletal condition	84.5	446	1.189	0.839-1.683	0.331
≥Overweight	83.2	232	1.251	0.815-1.922	0.306
Other condition	85.7	49	1.374	0.588-3.212	0.464
Mental health issue	80.3	122	1.106	0.659-1.854	0.704
Season of referral					0.001
Winter	80.9	450	1.000 (ref)		
Autumn	79.9	527	0.956	0.691-1.323	0.788
Spring	82.1	526	1.060	0.762-1.474	0.729
Summer	89.9	375	2.095	1.383-3.173	0.000

n, number of participants; Adj OR, adjusted odds ratio; CI, confidence interval; ref, reference category.

5.4.7. Characteristics associated with adherence

Analysis of adherence was conducted by comparing those who dropped out with those who completed, excluding participants who did not uptake or who were still participating. Participants who dropped out were younger than those who completed (54.4±14.8 vs 57.2±14.4, t(1081) = 3.176, p ≤0.01). There were significant associations between completion and employment status (χ^2(3, n=1094) = 7.9, p ≤0.05), referrer profession (χ^2 (2, n=1094) = 6.6, p ≤0.05), and season of entry to the scheme (χ^2(3, n=1094) = 15.7, p ≤0.01). Completion was higher amongst participants who were: part-time employed (69.5%), compared to those who were retired or not working for other reasons (57.8%), unemployed (55.4%), or full-time employed (52.6%); or were referred by a nurse (62.6%), compared to an AHP (56.2%) or GP (51.4%), or; started participating in summer (65.9%), compared to autumn (57.6%), spring (52.2%) or winter (50%). There were no significant associations between completion and age group, gender, ethnicity, quintile of deprivation, disability status, referrer organisation type or primary medical condition (p >0.05).

At week 1, there were significant differences between completers and dropouts for: HRQoL - EQ-5D-3L (median (Mdn)= .727 (interquartile range (IQR) .263) vs .725 (.434), U=126578.5, n=1070, p ≤0.01), EQ-5D VAS (50 (30) vs 50 (30), U=126714.5, n=1070, p ≤0.01); mental wellbeing - WEMWBS (47 (13) vs 46 (13), U=91990.5, n=922, p ≤0.01), and; smoking status (χ^2 (1, n=1070) = 20.6, p ≤0.001). Completers reported better HRQoL and mental wellbeing scores at week 1 compared to those who dropped out. Current smoking was more frequent amongst dropouts (60%), compared to completers (40%). There were no significant differences between dropouts and completers for readiness to change (TTM stage of change), PA and sedentary behaviour (MET minutes of PA per week, sitting time per day), BMI, systolic BP, diastolic BP, or alcohol consumption frequency (p >0.05).

5.4.8. Predictors of adherence

Binary logistic regression was undertaken to explore the variation in participation status. Variables included in the binary logistic regression are presented below in Table 5.6. A total of 802 cases were included in the model, with 292 cases excluded due to missing data (most frequently missing: n=172 WEMWBS, n=177 BP). Removal of variables with high levels of missing data such as WEMWBS and BP did not improve the fit of the model, or impact the variables making a statistically significant contribution. A test of the model was statistically significant (Hosmer-Lemeshow: $\chi^2(8) = 3.9$, p >0.05), indicating that it was able to distinguish between completers and dropouts. The model correctly classified 64.7% of cases, and explained between 10.2% (Cox & Snell R^2) and 13.7% (Nagelkerke R^2) of the variance in participation status. As shown in Table 5.6, two independent variables (season of participation, smoking status) made a statistically significant contribution to the model. Those who were referred in the autumn, spring or summer months, when compared to winter, had a two to three times higher odds of completing the scheme. Being a smoker at the time of the initial consultation corresponded with half the odds of completing.

Table 5.6. Logistic regression to predict adherence (n=802), and the % of participants per category who completed.

Predictor	**% completion**	**n**	**Adj OR**	**95% CI**	***p***
Age					0.908
75+	64.6	65	1.000 (ref)		
65-74	64.2	193	1.063	0.562-2.009	0.851
55-64	56.4	211	0.887	0.448-1.758	0.732
45-54	52.1	167	0.804	0.377-1.718	0.574
35-44	49.4	81	0.696	0.294-1.648	0.410
25-34	48.6	72	0.657	0.270-1.602	0.356
18-24	53.8	13	0.871	0.225-3.381	0.842
Quintile of deprivation					0.551
Q1 (most deprived)	53.2	314	1.000 (ref)		
Q2	56.5	230	1.002	0.691-1.454	0.992
Q3	59.7	124	1.006	0.630-1.606	0.980
Q4	57.5	106	1.002	0.610-1.646	0.995
Q5 (least deprived)	78.6	28	2.367	0.880-6.365	0.088
Gender					
Female	54.7	468	1.000 (ref)		
Male	59.3	334	1.197	0.860-1.666	0.287
Ethnicity					
Caucasian	56.8	766	1.000 (ref)		
BME	52.8	36	0.820	0.388-1.733	0.603
Employment status					0.125
Retired/Not working	57.9	435	1.000 (ref)		
Unemployed	50	114	1.267	0.751-2.135	0.375
Part-time employment	71	69	2.044	1.094-3.816	0.025
Full-time employment	52.2	184	1.042	0.657-1.653	0.860
Disability status					
Non-declared	56.8	757	1.000 (ref)		
Disability declared	53.3	45	0.889	0.443-1.784	0.741
Referrer organisation					
Other	58.6	442	1.000 (ref)		
Medical practice	54.2	360	0.690	0.464-1.025	0.066
Referrer profession					0.345
AHP	55.2	431	1.000 (ref)		
Nurse	63.6	206	1.345	0.902-2.005	0.146
Doctor	51.5	165	1.182	0.727-1.921	0.500

Predictor	**% completion**	**n**	**Adj OR**	**95% CI**	***p***
Primary referral condition					0.482
Multiple conditions	56	266	1.000 (ref)		
CVD or high risk of CVD	64.6	130	1.034	0.638-1.675	0.892
Diabetes/metabolic syndrome	69	29	1.623	0.669-3.938	0.285
Respiratory condition	60	25	1.033	0.416-2.564	0.944
Musculoskeletal condition	49.7	175	0.739	0.470-1.162	0.190
≥Overweight	61.3	106	1.109	0.658-1.868	0.699
Other condition	61.9	21	1.005	0.363-2.783	0.992
Mental health issue	42	50	0.569	0.284-1.141	0.112
Season of participation					0.000
Winter	42.3	137	1.000 (ref)		
Autumn	59.9	269	2.621	1.646-4.174	0.000
Spring	52.3	176	1.978	1.206-3.243	0.007
Summer	65	220	3.241	2.005-5.237	0.000
TTM stage of change					
Pre-contemplation, contemplation	56.4	691	1.000 (ref)		
Action phases	57.7	111	1.032	0.621-1.717	0.902
MET mins PA/week	56.6	802	1.000	0.999-1.000	0.155
Sitting time mins/day	56.6	802	1.000	0.999-1.001	0.636
BMI	56.6	802	0.988	0.965-1.012	0.324
Systolic BP	56.6	802	0.998	0.985-1.011	0.749
Diastolic BP	56.6	802	0.997	0.976-1.019	0.795
EQ-5D-3L	56.6	802	1.506	0.890-2.548	0.128
EQ-5D VAS	56.6	802	1.003	0.994-1.012	0.543
WEMWBS	56.6	802	1.006	0.987-1.025	0.549
Smoking status					
Non-smoker	60.1	662	1.000 (ref)		
Smoker	40	140	0.499	0.329-0.756	0.001
Alcohol consumption frequency					0.995
Weekly	58.7	230	1.000 (ref)		
Monthly	57.7	104	0.980	0.587-1.636	0.939
< Monthly	54.7	254	1.037	0.690-1.557	0.862
Never	56.1	214	1.039	0.675-1.600	0.862

n, number of participants; Adj OR, adjusted odds ratio; CI, confidence interval; ref, reference category.

5.4.9. Summary of findings – study 1

- Referrals were predominantly received for participants who were: under the age of 65 years (67.7%, n=1345); female (60.2%, n=1197), Caucasian (82.3%, n=1635), and from the most deprived quintile (39.9%, n=793).
- Just under half of referrals were received from a medical practice (44.8%, n=890), with the remainder received from secondary care organisations or community services (55.2%, n=1097).
- The most common reasons for referral were for having: multiple (≥3) health conditions (33.7%, n=669), musculoskeletal related conditions (23.5%, n=467), existing CVD or significant risk factors (15.1%, n=300), or for being overweight or obese (12.5%, n=248).
- Using definitions provided in section 5.2, the uptake rate was 81%, and the adherence rate was 56.9%.
- Of the 29.3% (n=471) of participants who entered the scheme, but dropped out early, 52% left before week 6 (n=245), 33.8% left before week 12 (n=159), and 14.2% left before week 24 (n=67).
- The most common reasons for non-uptake and dropout were that the ERS could not establish contact with the participant, or the participant perceived they could not be active due to their health status.
- There was an association between uptake and higher (Q4) and lower (Q2) SES, and being referred in the summer months when compared to winter.
- Adherence was predicted by entering the scheme in the autumn (OR=2.62), spring (OR=1.97) or summer months (OR=3.24) (compared to winter), or being a non-smoker at the time of the initial consultation (OR=2.00).

5.5. Study 1 - Discussion

5.5.1. Main finding of study 1

The aims of this study were to evaluate engagement levels in the 'Live Active' ERS, whilst addressing some of the gaps and limitation of previous research, by seeking to further understand the factors leading to engagement, or non-uptake and dropout. These aims are important as low uptake and adherence to ERSs is an established problem, with further research required to understand the reasons behind non-engagement, and how to facilitate better outcomes (Morgan et al., 2016; Pavey et al., 2012). Furthermore, equal exploration of uptake and adherence was important, since previous research has explored adherence more so than uptake, with less known about uptake (Pavey et al., 2012).

Compared to engagement levels reported in a systematic review by Pavey et al. (2012) (average uptake rate 66-81%, average adherence rate of 43%-49%), uptake (81%) and adherence (56.9%) to this ERS were higher than average, with the majority of dropout occurring in the first six weeks of participation (52%). Reasons for non-uptake and dropout have rarely been reported (Campbell et al., 2015; NICE, 2014a), with potential reasons inferred from the views of those who do engage (Morgan et al., 2016). The findings indicated the existence of common barriers to PA such as time (Morgan et al., 2016), as well as unavoidable causes such as relocation, and inappropriate referrals for those with contraindications to PA. These findings indicate that non-engagement does not always necessarily reflect an ineffective ERS.

Logistic regression was conducted to identify variables that could predict participation, in order to compare with previous research and build upon its limitations (Campbell et al., 2015; Pavey et al., 2012). Variables such as ethnicity and disability status were included in the model of adherence. This was a strength of this study that previous research was often unable to explore due to poor participant profiling (Gidlow et al., 2005; NICE, 2014a; Sowden & Raine, 2008), with the exception of few studies such as Mills et al. (2012) and Kelly et al. (2017). Variables such as employment status (Hanson et al., 2013), referral organisation (Kelly et al., 2017), and smoking status (Ward et al., 2010) had also rarely been explored. Relatively few variables significantly contributed to prediction models of uptake and adherence, with SES and season of referral being associated with uptake, and smoking status and season of entry to the scheme predicting adherence.

It is not possible to note with certainty the reasons for higher than expected engagement in this ERS. In addition to the characteristics explored in this study, it is possible that engagement was related to aspects not explored, but known to influence adherence. This includes factors such as the longer than average scheme duration (Rowley et al., 2018), the subsidised price of activities, the personality and approach of the exercise professionals (Morgan et al., 2016), or because the scheme was developed based upon knowledge of the local context (Tobi et al., 2017). Further qualitative research may clarify these reasons.

5.5.2. Referrals received

The reach of the ‘Live Active’ ERS was wide, engaging hard to reach groups such as socio-economically disadvantaged, disabled and BME groups. Based on mid-year 2015 estimates, referrals were received for 1.2% of Tameside’s adult (18+ years) population during this analysis period (ONS, 2017b). The number of referrals received was within the range reported by other ERS evaluations (e.g. Harrison et al., 2005a; Kelly et al., 2017; Lee et al., 2009). This may suggest that health care professionals are interested in promoting PA through methods of referral, as recommended by global (WHO, 2018) and national policy (PHE, 2014a).

One distinguishing feature though, was the proportion of referrals received by primary health care professionals in comparison to other ERSs reporting these figures. In this study, 44.8% of referrals were made from GPs and nurses in primary care practices, compared to 74%-86% reported in other studies (Hanson et al., 2013; Kelly et al., 2017; McGeechan et al., 2018). In this ERS, the majority of referrals were therefore made from secondary care and community services (55.2%), which is in contrast to previous research. This demonstrates the value of not only primary health care professionals, but also those working in secondary care and community services, to identify those in need of intervention. This is an important consideration for ERSs that accept referrals from primary care services only (e.g. Duda et al., 2014; Edwards et al., 2013; Harrison et al., 2005a), potentially limiting the reach of ERSs in these areas.

5.5.3. Uptake and adherence levels

Compared to previous RCT and observational studies (average uptake 66-81%, average adherence 43%-49% (Pavey et al., 2012)), this scheme had high uptake (81%) and adherence (56.9%) levels. Previous research has suggested that observational studies more closely represent the engagement levels of interventions, as those participating in a RCT may potentially be more committed to the intervention (Pavey et al., 2012). Therefore, engagement levels of observational studies are expected to be much lower than those of RCTs (Pavey et al., 2012). However, this was not the case with this study, which is a positive, if unanticipated finding.

The findings indicated that the majority of dropout (52%) occurred within the first six weeks of participation, which is comparable to recent research (Kelly et al., 2017; Moore et al., 2013). This indicates that the first six weeks of an ERS are critical for preventing dropout. The reasons for early dropout, reported in this study, were similar to those reported by participants who dropped out after the first six weeks. Therefore, understanding the reasons for early dropout and how it can be prevented may be better explored through qualitative research. Stathi et al. (2004) have suggested, for example, that overcoming barriers to PA (e.g. familiarity with a leisure centre environment) is critical in the first weeks of participation.

Direct comparison of engagement levels between studies is difficult due to the varying range of definitions of adherence that are used (Pavey et al., 2012). This study, like others, considered adherence to be related to attendance at consultations. This is in contrast to studies that have understood adherence to be based on attendance at leisure centre-based exercise sessions (for e.g. see Table 5.1), which fails to account for those participating in activities outside of the leisure centre. Whilst participation in PA could be considered to be a better reflection of adherence, the logistical reality of tracking the physical activities of each participant may be beyond the scope of an ERS's capabilities. Tobi et al. (2012) proposed that adherence could be described in terms of completers and dropouts, but should ideally also be related to the frequency of attendance at consultations. They suggested that adherence could be determined by attendance at both a mid-way point and final consultation, to overcome the limitation of participants who do not engage in an ERS but still attend the final consultation. Whilst acknowledging that this could be a viable consensus definition of adherence, application of this definition was not possible in this study. This was due to: i) the ERS completing telephone consultations, including data collection, with participants who did not attend an appointment, and; ii) inconsistencies in the marking of whether participants had attended a consultation, or had not attended and had received a follow-up telephone consultation. Therefore, it was not possible to distinguish with accuracy which participants had not attended a consultation, and which consultations had been completed via telephone following non-attendance.

Previous studies have measured uptake and adherence over varying time frames such as 15 months (Hanson et al., 2013), three years (Gidlow et al., 2007) and five years (Kelly et al., 2017). The results from this study are representative of uptake and adherence across two years, and could not be directly compared to a scheme measuring these outcomes over a significantly different time frame. The rationale for this is based on the gradual decline in uptake and adherence that was observed during quarterly analyses of secondary data during the first two years of the scheme's operation (appendix 9). Therefore, if this study had evaluated engagement levels across the first year of the scheme's operation, reported uptake and adherence levels would have been higher. However, there are no identified published studies that have reported engagement levels over a series of time points, such as on a quarterly basis, and it is therefore not known if other schemes experience such variation in engagement levels.

5.5.4. Reasons for non-uptake and dropout

The reporting of reasons for non-uptake and dropout from a large sample is novel, and is usually inferred from the views of participants who uptake and adhere to a scheme (Morgan et al., 2016). The findings highlight that non-engagement occurs for reasons other than commonly reported barriers such as time (Morgan et al., 2016). Reasons can be classified as those relating to personal choice, such as having no interest in participation, or health-related reasons such as perceiving that they could not be active due to their health status (e.g. musculoskeletal pain). They also related to health inequalities, such as having no available finances to pay for activities.

A large proportion of participants were classified as non-uptake or dropout because they could not be contacted, or did not attend a consultation multiple times. This could be related to lack of readiness to change or other unknown factors such as self-esteem, rather than communication difficulties, particularly as participants were contacted using multiple methods over a three-month period before classification as non-uptake or dropout. In noting that little research had focused on referral processes in comparison to the effectiveness of schemes, Johnston et al. (2005) highlighted the inflation of non-engagement figures due to inappropriate referrals for participants with contraindications to PA, despite dissemination of exclusion criteria to referring professionals. This was also observed in this study. Those referred with no motivation to change could also be considered an inappropriate referral, due to this being a pre-requisite to participation (NICE, 2014a). However, it could be argued that these participants have been identified as requiring an intervention, and that an ERS should therefore attempt to engage these individuals with relevant behaviour change techniques. Therefore, the findings regarding reasons for non-engagement support the conclusions of Johnston et al. (2005), that inappropriate referral practices remain a problem, contributing to >8% of non-uptake in this study. It would therefore be beneficial for schemes to communicate the inclusion and exclusion criteria to referrers, and to provide feedback when inappropriate referrals have been received, to reduce the incidence of inappropriate referrals.

The findings also support the suggestion by Harrison et al. (2005a) that not all dropouts are negative events, with some participants in this study reporting to have become independently active, not requiring further support (7%). These participants may have been motivated to become active after only a brief exposure to the ERS (Pavey et al., 2012), and therefore may have been suitable for a less contact intensive model of exercise referral. This however is hypothetical, with exploration of PA behaviour after dropout previously identified as an area for future research (Sørensen, Skovgaard, & Puggaard, 2006; Tobi et al., 2012).

High levels of non-uptake and dropout reduce the cost-effectiveness of ERSs (Pavey et al., 2012), leading to arguments that ERSs should be decommissioned in light of frequently reported poor engagement levels and uncertain impact (Pavey et al., 2011). However, with little exploration of the reasons why participants non-uptake or dropout, this argument assumes that non-engagement is due to the ineffectiveness of ERSs at enagaging and retaining participants. The findings of this study refute this argument and demonstrate the value of exploring the reasons for non-engagement when interpreting engagement levels.

Exploring the reasons for non-engagement has also provided the ERS with an opportunity for learning and improvements in practice, to prevent non-uptake and dropout. The reasons for non-engagement were communicated to the ERS in a quarterly report, intended as formative feedback that could lead to improvements throughout. However, communication of this information did not appear to reverse the progressive decline in engagement during the study period (appendix 9).

It is not known if the reasons for non-engagement in this study are comparable to other ERSs due to the lack of reporting. It is therefore recommended that ERSs record and regularly review the reasons for non-engagement, and that evaluations report these reasons. This would identify areas for improvements in practice, which could ultimately improve engagement and therefore the likelihood of positive PA behaviour change. Knowledge could be greatly enhanced by further qualitative exploration to understand these reasons, which has previously been suggested as an area for future research, but not yet fulfilled (Hanson et al., 2013; Isaacs et al., 2007), potentially due to the difficulty in engaging with people who did not engage in an ERS. Such research could also potentially answer the question of why the majority of participants dropout within the first six weeks of participation, and provide recommendations on how schemes can adapt to reduce early dropout (Hanson et al., 2013; Isaacs et al., 2007).

5.5.5. Predictors of uptake and adherence

By including a wider range of variables in statistical models to predict uptake and adherence, this study has sought to overcome limitations of previous literature, where poor participant profiling has limited exploration of predictors (Gidlow et al., 2005). Whilst demographic variables such as ethnicity, employment status and disability status did not significantly contribute to prediction models in this study, inclusion of these predictors provides an understanding of engagement levels in these different groups, which may also be related to the geographical and strategic context of the 'Live Active' ERS. The ERS did, for example, target referrals from BME groups.

There was an association between uptake and SES (Q2, Q4). Therefore, these findings only partly support previous research identifying lower likelihood of uptake amongst referrals from more deprived areas (Gidlow et al., 2007; Hanson et al., 2013; Moore et al., 2013). They are also surprising since there are strong SE patterns in PA in the England (Farrell et al., 2014), and therefore likelihood of uptake would be expected to increase linearly with SES. This finding is promising, as it may suggest that health inequalities have not negatively impacted access to this ERS, which is important given the SE context of Tameside. However, it is also relevant to note that the majority of referrals made to the ERS were for individuals from areas of SE deprivation, with two thirds of referrals made for participants from the two most depived quintiles of Tameside.

There was also an association between uptake and seasonal of referral (summer). No other identified published studies have examined the association between season and engagement with an ERS. Season of referral to the scheme and season of entry to the scheme were included in multivariate models based on the evidence that seasonal variations in weather may be a barrier to PA in various populations (Tucker & Gilliland, 2007), and there may be seasonal variations in the status of various health conditions (Dopico et al., 2015). Poor weather conditions, in the winter for example, are a known barrier to PA participation in various populations (Tucker & Gilliland, 2007), and anecdotal observations from ‘Live Active’ staff suggest this may also have prevented travel to the leisure centres for consultations (e.g. fear of falling on ice). Seasonal variation can also be observed in health conditions and other diseases (Dopico et al., 2015) such as viral infection (Fisman, 2012), COPD exacerbation (Donaldson & Wedzicha, 2014), and perceived osteoarthritis related pain (Timmermans et al., 2014), which may prevent participation in PA. The impact of season on engagement seen in this study may therefore support these hypotheses, with further research, potentially qualitative, required to understand exactly why season affects engagement. These findings have, however, identified potential areas for improvement in practice in this ERS, such as the provision of more indoor activities during the winter season, additional staff training on directing coping strategies for seasonal barriers to PA, and exploration of multi-sector partnerships to promote self-management, accessible environments, and accessible transport during these periods, to promote continued engagement in PA.

The results do not add to the consensus regarding older age with increased likelihood of uptake (Campbell et al., 2015; Pavey et al., 2012), as age was not associated with uptake. Previous research has suggested that older age adults (i.e. retired adults) have more free time available to dedicate to activities such as PA (Tobi et al., 2012). However, this effect was not observed in this population, as level of uptake was similar across the age groups. Furthermore, the results are in contrast with the predominance of research finding that females are more likely than males to uptake (Dugdill et al., 2005; Hanson et al., 2013; Lord & Green, 1995; Moore et al., 2013; Sowden et al., 2008). Female participants represented a larger majority of referrals to this scheme (60.2%), but no association between gender and uptake was observed. Two other studies have also found no association with gender (Harrison et al., 2005a; Gidlow et al., 2007). Previous research exploring the association between primary referral condition and uptake have produced equivocal results (Pavey et al., 2012), and no association between referral condition and uptake was found in this study. However, this may be due to the variation in reporting of medical conditions, and varying inclusion criteria of different schemes (Pavey et al., 2012). Furthermore, in this study, whilst the primary reason for referral could be identified from the referral information, the majority of participants also had additional comorbid conditions, and these conditions may have contributed to the ability or willingness to uptake more so than the primary referral condition.

Adherence was predicted by entering the scheme in autumn, spring or summer, compared to winter, or by being a non-smoker at the time of the initial consultation. As previously noted, no identified published studies have examined the impact of season on engagement, and the results may support previously mentioned hypotheses that seasonal variation in habitual PA and health status may impact engagement. Smoking status has been reported to a lesser extent in the literature compared to other variables, but the findings are in agreement with previous research identifying higher likelihood of dropout amongst smokers (Kelly et al., 2017; Ward et al., 2010). It has been suggested that negative health behaviours such as smoking and physical inactivity cluster (Kaczynski et al., 2008; Saint Onge & Krueger, 2017; Spring et al., 2012), and that the physiological impacts of smoking (i.e. reduced lung function) may impair physical capabilities (Kaczynski et al., 2008), which may have contributed to dropout from this ERS. Further research to understand this association is important, as previous research suggests that the combined effects of smoking and inactivity on mortality are greater than either behaviour alone (Borrell, 2014). Qualitative research, for example, may help researchers and practitioners to understand if smokers require additional support to become physically active, or if it would be beneficial to address smoking cessation before referral to an ERS.

Similar to the literature regarding uptake, results of this second multivariate analysis do not add to the consensus regarding older age with adherence (Campbell et al., 2015; Pavey et al., 2012), as age was not a significant predictor. They also do not support previously reported predictor variables related to gender (Gidlow et al., 2007; Hanson et al., 2013; Moore et al., 2013), primary referral condition (Crone et al., 2008; Dugdill et al., 2005; Tobi et al., 2012), referrer profession (Dugdill et al., 2005; Hanson et al., 2013; James et al., 2008), and baseline health status (Lee et al., 2009), as none of these were significant variables. Where primary referral condition has previously been identified as a predictor variable, it has been suggested that this may be due to the varying and unpredictable nature of condition related symptoms, and perceived restriction of PA by these symptoms, in turn leading to dropout (Mills et al., 2012; Tobi et al., 2012). In this study, 15.1% of participants who dropped out reported that they felt that they could not be active due to their current health status (e.g. musculoskeletal related pain), which would support this hypothesis. Despite this, primary referral condition was not a significant variable, but as previously mentioned this may be due to the majority of participants having multiple co-morbidities.

It is important to note that differences between the findings of this study and others may relate to the specific contexts that ERSs operate in, with different strategic aims, referral criteria, and scheme deliveries (Arsenijevic & Groot, 2017; Tobi et al., 2017). There may be, for example, considerable differences between schemes in the distribution of SES, age range, and eligible medical conditions. It is also important to recognise the substantial heterogeneity in scheme delivery, in which ERSs may offer leisure facilities of different standards, at different prices ranges, with a range of different PA options and social features, run by exercise professionals with different attributes, all of which could impact engagement (Morgan et al., 2016).

5.5.6. Strengths and limitations of study 1

The primary strengths of this study are the relatively large sample size, the more extensive participant profiling, and the reporting of new information such as the reasons for non-engagement. As this was an evaluation of engagement in a real-life ERS, this has provided stakeholders with an opportunity for learning, and direction to improve practices which could potentially maximise engagement in the future. The data can also be explored further in the future, to address additional questions from stakeholders.

As previously discussed, it was not possible to apply the proposed definition of adherence by Tobi et al. (2012) with accuracy. Therefore, it is possible that participants who completed this ERS did not fully engage with the scheme but still attended the week 24 consultation, not representing adherence.

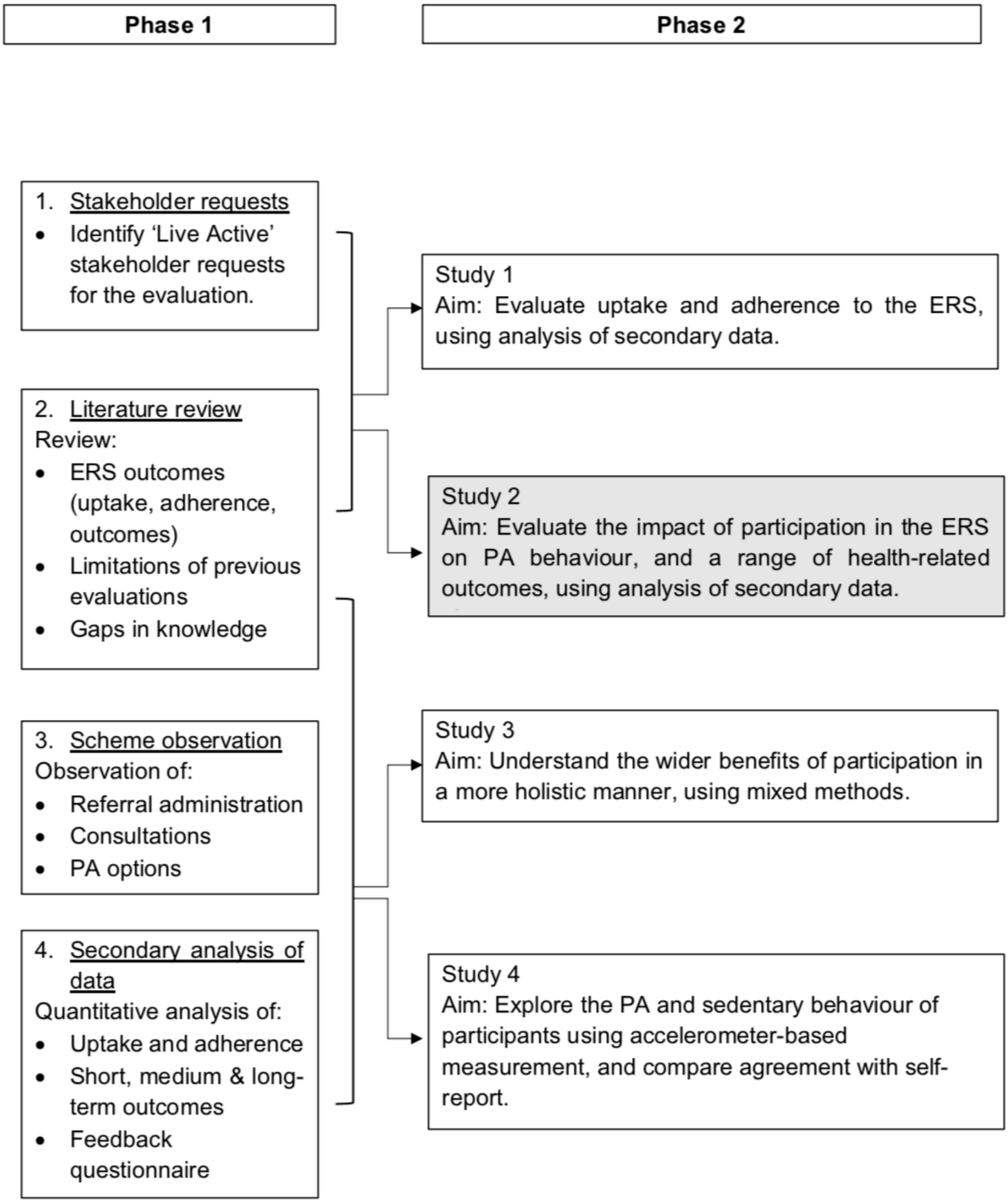
Missing data is a common issue with secondary data (McKnight & McKnight, 2011), as previously discussed in section 4.4.1. In this study, the limitations of previous research regarding poor participant profiling could not be overcome in the prediction model of uptake, due to limited data available from the referral form. Furthermore, there were high levels of missing data for BP and WEMWBS in the prediction model of adherence, although this did not affect the outcome as removal of the variables did not improve the fit of the model, or impact variables making a statistically significant contribution. Prediction models could also only predict 83% of uptake, and 65% of adherence, suggesting that factors other than those routinely collected may have influenced these outcomes.

Lastly, as discussed above, ERSs are complex interventions operating in different contexts, and the findings of this study may relate to its context within a specific geographical area, with particular strategic aims, referral pathways, methods of delivery and inclusion criteria, and may not be comparable to other ERSs (Arsenijevic & Groot, 2017; Tobi et al., 2017). This is why it is important for evaluations to describe the context of ERSs, and referral processes, to understand if the findings of one ERS are transferable to another (Moore et al., 2015).

5.5.7. Conclusions of study 1

Engagement levels with this ERS were higher than previously reported. Those who did not uptake a referral to this ERS, or dropped out early did so for a variety of reasons. Some cases of non-engagement were due to inappropriate referrals, indicating that non-engagement does not always necessarily reflect an ineffective ERS. Uptake was associated with SES (Q2, Q4) and season of referral (summer). Adherence was predicted by entering the scheme in autumn, spring or summer, compared to winter, or being a non-smoker at the time of the initial consultation.

These findings have been communicated to the ERS to inform improvements in practice. However, further qualitative research is required to build upon these quantitative findings, to better understand the reasons for higher than expected engagement, why these relationships with uptake and adherence occured, and to gain richer understanding of the reasons for non-uptake and dropout. This may identify potential adaptations to an ERS which could facilitate uptake and adherence for certain groups such as smokers, or those who experience seasonal variation in health status.



Overview of the research framework (study 2)

Chapter 6 Study 2: Evaluation of the routine outcomes of the 'Live Active' exercise referral scheme

6.1. Study 2 – Introduction, aims and objectives

A non-specific aim of ERSs to positively benefit health (NICE, 2014a), and a lack of clear guidelines as to how to evaluate the effectiveness of an ERS at meeting this aim, has led to uncertainty over what outcome measures to collect. This uncertainty has contributed to the wide variation in the measures collected in schemes across the country to date, limiting the ability for comparison of results across schemes (Campbell et al., 2015; Pavey et al., 2011). This chapter begins by looking at the variety of quantitative outcomes that have been reported in ERS evaluations. It considers the evidence for the impact of ERSs on PA level, physical health outcomes, mental wellbeing and HRQoL. The introduction concludes with a discussion of the limitations of previous evaluations, namely the lack of research evaluating the longer-term impact of ERS participation, and whether participation is more beneficial for particular sub-groups.

These limitations informed the objectives of this research study, which were to evaluate the medium to long-term impact of ERS participation on PA, and health-related outcomes, at the level of population and primary referral condition sub-group. These objectives contribute new knowledge, by providing information on the long-term impact of participation, and by acknowledging the heterogeneity of this population with a sub-group analysis, as has been recommended by previous research (Anokye et al., 2011; Pavey et al., 2011; Rowley et al., 2018).

Study 2 aim: Evaluate the impact of participation in the ERS on PA behaviour, and a range of health-related outcomes, using analysis of secondary data.

Study 2 objectives:

2.1. Identify changes in PA level, BMI, BP, HRQoL (EQ-5D-3L, EQ-5D VAS), mental wellbeing (WEMWBS), and smoking and alcohol consumption between weeks 1, 12, and 24 (medium-term), and weeks 1, 24 and 52 (long-term).

2.2. Investigate whether differences in PA level and health-related outcomes exist between primary referral condition groups.

6.2. Literature review of the quantitative outcomes of ERS participation

6.2.1. Physical activity

When previous research has attempted to establish the clinical-effectiveness (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007) and cost-effectiveness of ERSs (Anokye et al., 2011; Trueman & Anokye, 2013), the primary outcome of importance has been PA. It has been the most frequently reported outcome, predominantly measured using the 7-Day PA Recall (7-Day PAR) questionnaire (Sallis et al., 1985) (see for e.g. Harrison, Roberts, & Elton, 2005b; Murphy et al., 2012). No identified studies have measured total weekly PA in ERS participants using device-based measures such as accelerometers (Campbell et al., 2015), with the exception of Gallegos-Carrillo et al. (2017) in Mexico, which cannot be generalised to a UK ERS setting, and Giné-Garriga et al. (2017), which is in-progress and published as a study protocol. These studies will be further discussed in chapter 8, which looks specifically at the measurement of PA.

Several methods have been used to report self-reported PA in ERSs. The most common of these have been average minutes of PA per week, and the proportion of participants achieving ≥90-150 minutes of PA per week, which has been the most frequent reporting method (Campbell et al., 2015; Pavey et al., 2011). Looking at these in turn, within the systematic review of Campbell et al. (2015), pooled data demonstrated a mean increase in PA of 55.10 minutes per week in the ERS group, compared to a comparison group at 6–12 months follow-up. In respect of the proportion of participants achieving ≥90-150 minutes of PA per week following ERS participation, systematic reviews have found no strong evidence for a long-term improvement (Campbell et al., 2015; Pavey et al., 2011). Over the previous decade, there appears to have been a gradual reduction in the proportion of participants achieving this outcome at 6–12 months follow-up. To illustrate, in the systematic review of Williams et al. (2007), participants in the ERS group were 20% more likely to achieve ≥90-150 minutes of PA per week compared to a comparison group, falling to 16% (Pavey et al., 2011), and then 12% (Campbell et al., 2015) in later systematic reviews. However, only three of the included studies in the latest systematic reviews (Campbell et al., 2015; Pavey et al., 2011) measured long-term change in PA (≥12 months). Therefore, the impact of ERS participation on long-term behaviour change remains uncertain, with further research required that measures PA over 12 months or longer (Campbell et al., 2015; Pavey et al., 2011).

When interpreting the literature regarding the impact of ERSs on PA levels, it is important to take in to account the limitations of previous studies that have reported this outcome, and the limitations of systematic reviews that have attempted to provide a consensus on the impact of ERSs. It is difficult, for example, to make comparisons between studies, or generalisations about the impact of ERSs on PA, due to limitations in study design, analysis and reporting. ERSs are also, as previously discussed, highly heterogenous interventions. There is variation between

ERSs, for example, in the questionnaires used to assess PA level, and in how PA level is reported (total minutes of PA vs proportion meeting PA guidelines) (Campbell et al., 2015; Pavey et al., 2011). PA in ERSs has predominantly been measured using the 7-Day PAR questionnaire (Campbell et al., 2015; Pavey et al., 2011). However, more recent studies have measured PA using the IPAQ short form (e.g. Gallegos-Carrillo et al., 2017; Martín-Borràs et al., 2018; Webb et al., 2016). Differences exist between these questionnaires, for example the main outcome of the 7-Day PAR is minutes of PA (Sallis et al., 1985), whereas the main outcome of the IPAQ is MET minutes of PA (IPAQ, 2005). Furthermore, the 7-Day PAR collects data on time spent in light, moderate, vigorous and very vigorous intensity PA, whereas the IPAQ short form surveys time spent walking or in moderate or vigorous intensity PA. In addition, different MET values are assigned to these intensities, for example the IPAQ assigns 8 METs to vigorous intensity PA, whereas the 7-Day PAR assigns 6-10 METs. Differences in the measurement of PA can therefore complicate the comparison of studies using different questionnaires.

The most common method of reporting PA level in ERS evaluations has been the proportion of participants achieving ≥90-150 minutes of PA per week. However, the rationale to categorise according to this threshold has not been clearly described, but is linked to national PA guidelines such as the UK DoH's (2011) PA guidelines for adults, despite these guidelines recommending 150 minutes or more of moderate intensity PA per week, rather than 90 minutes or more. These are also generalised guidelines for healthy adults, intended to be interpreted with further consideration when dealing with those with physical and mental health needs. Applying this threshold in evaluations potentially means that a significant number of newly active participants, reporting less than 90 minutes per week, are labelled as inactive

(Campbell et al., 2015), despite recognition that activity below this threshold is beneficial, with significant improvements in mortality risk when moving from physical inactivity to a low volume of weekly PA (Arem et al., 2015; Hupin et al., 2015; Wen et al., 2011). Findings based on this threshold would therefore not fully reflect the impact of an ERS intervention on changes in PA level. This is particularly important for economic evaluations of ERSs, which determine success by meeting the ≥90-150 minutes threshold, therefore not capturing economic benefits gained from PA below this level (NICE, 2014a). In this regard, there is a need for future evaluations to reconsider this threshold, and whether it should differ between referral conditions, in order to capture those being active at beneficial levels for their respective conditions, but disregarded as 'inactive' by current methods. It would also be beneficial for studies to come to a consensus on how PA should be reported, as minutes of PA per week, the proportion meeting the ≥90-150 minutes threshold, or both.

Furthermore, all studies included in systematic reviews to-date, and the wider literature with the exception of Gallegos-Carrillo et al. (2017) (published as a trial protocol), have measured PA through self-report. These measures are vulnerable to bias, such as poor recall and social desirability (Haskell, 2012; Helmerhorst et al., 2012; Strath et al., 2013), and the outcomes should therefore be interpreted with caution, as it is known that participants may under or over-report PA levels (Cerin et al., 2016; Craig et al., 2003; Steene-Johannessen et al., 2016). The current consensus on the effectiveness of ERSs is therefore based on self-reported data (Campbell et al., 2015; Pavey et al., 2011), which may under or over-estimate the impact of the intervention.

Lastly, systematic reviews that attempt to provide a consensus on the impact of ERSs, and inform policy, impose strict inclusion criteria largely limited to evidence from RCTs (Gidlow et al., 2008). In the case of ERSs, this has led to generalisations based on eight studies, representing a small fraction of the 600 ERSs in operation in the UK (Campbell et al., 2015; Pavey et al., 2011). Furthermore, in all of the included RCTs, participants in the control group were given information related to PA, such as brief advice, that may have led to an increase in PA in both groups (Campbell et al., 2015). Control group contamination may also have occurred by volunteers in the control group who were motivated to become active when they agreed to take part in the trial, which would reduce the apparent effectiveness of ERSs when both groups are compared (Campbell et al., 2015; Gidlow et al., 2008). Therefore, in future it is important for researchers to consider alternative forms of evidence, such as high quality observational and qualitative studies, when evaluating real world practice (Gidlow et al., 2008), as previously discussed in section 2.5.

6.2.2. Physical health outcomes

Physical health outcomes have also been a primary indicator of success of ERSs, although with inconsistent evidence for an improvement in outcomes such as BMI and BP (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). Body composition has been assessed using a variety of methods such as body weight (Isaacs et al., 2007; Stewart, Dolan, Carver, & Swinton, 2017), BMI (Isaacs et al., 2007; Webb et al., 2016), body fat percentage (Isaacs et al., 2007; Ward et al., 2010), waist circumference (McGeechan et al., 2018; Webb et al., 2016) and waist-to-hip ratio (Isaacs et al., 2007; Stewart et al., 2017). Systematic reviews have reported no significant difference in body composition, compared to comparison groups, at short and long-term follow-ups, with any improvement in the ERS group mirrored by an improvement in the comparison group (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). In subsequent studies since the latest systematic review, Stewart et al. (2017) found significant improvements in body weight and waist-to-hip ratio in a 12-week ERS, whilst McGeechan et al. (2018) also found improvements in BMI. Parretti et al. (2017) recently attempted to review the impact of ERSs on the health of participants with obesity, but were unable to do so as no health-related outcomes had been stratified according to BMI class. The authors suggested that such analyses and reporting would be useful, as participants of different BMI classes may reduce BMI to different extents, which would not be evident when reporting a population average.

BP has also been an outcome of interest, since it is a risk factor for CVD (Goff et al., 2013; Piepoli et al., 2016) that can be improved by PA (Börjesson, Onerup, Lundqvist, & Dahlöf, 2016; Cornelissen & Smart, 2013; Diaz & Shimbo, 2013). Again, systematic reviews have reported no significant differences in resting BP, compared to comparison groups, at short and long-term follow-ups (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). A subsequent study published since the last systematic review also found no significant short-term (12-weeks) changes in resting BP (Stewart et al., 2017).

Few studies have examined parameters beyond body composition or resting BP, such as markers of physical fitness, lung function or biochemistry. The impact of ERS participation on these outcomes has been identified as an area for future research (Pavey et al., 2011). However, this may be difficult for researchers to address, if evaluations such as the present study are based on secondary data. In this respect, the researcher cannot always influence which outcomes are collected (McKnight & McKnight, 2011; Vartanian, 2011), and it may not be possible to conduct a further study to explore alternative outcomes based on the recommendations of the literature.

Mixed results have been found for changes in physical fitness. Stewart et al. (2017) found significant short-term (12 weeks) improvements in peak oxygen uptake during a shuttle walk test, and performance in a ‘sit to stand’ test. Sørensen et al. (2008) found no long-term (10 months) improvements in maximal oxygen uptake. In contrast, Isaacs et al. (2007) found long-term improvements in various markers such as time to exhaustion, heart rate response to exercise, and leg extensor power, but the majority of these markers were not significantly different to comparison groups who received advice only, or attended a walking group.

Considering respiratory function, Stewart et al. (2017) found significant short-term improvements in lung function (ratio of forced expiratory volume over one and six seconds) over the first six weeks of participation. In contrast, Isaacs et al. (2007) found no long-term (12 months) differences in markers of lung function, such as forced expiratory volume in one second, forced vital capacity, and peak expiratory flow, compared to groups who received advice only or attended a walking group.

Various blood markers have been measured in a small number of studies. Sørensen et al. (2008) measured a marker of blood glucose, glycated haemoglobin (HbA1c), but found that the reduction in HbA1c over 10 months in the ERS intervention group was mirrored by a reduction in the comparison group accessing PA counselling. Webb et al. (2016) demonstrated short-term (8 week follow-up) improvements in total cholesterol and low-density lipoprotein cholesterol, but not triglycerides, high-density lipoprotein cholesterol, glucose or insulin markers. Isaacs et al. (2007) also found similar small, but long-term (12 months) improvements, in total cholesterol and low-density lipoprotein cholesterol, but not high-density lipoprotein cholesterol. However, these changes over 12 months were also mirrored in comparison groups who received advice only or attended a walking group. Further physical health outcomes, including blood markers, are currently being explored in an ERS evaluation, which will add further knowledge of the impact of ERSs on physical health parameters (Giné-Garriga et al., 2017). These include markers of frailty and inflammation such as interleukin-6, as well as muscle biopsies to examine sarcopenia-associated markers of muscle quality.

6.2.3. Psychological health and health-related quality of life

Psychological health related outcomes have been measured to a lesser extent compared to physical health outcomes. They have been assessed using a variety of questionnaires such as the Hospital Anxiety and Depression Scale (HADS) (e.g. Duda et al., 2014; Isaacs et al., 2007; Murphy et al., 2012), the Profile of Mood States (POMS) (e.g. Stewart et al., 2017), the Physical Self-Perception Profile (e.g. Taylor & Fox, 2005), and the anxiety/depression component of the EQ-5D (e.g. Gusi, Reyes, Gonzalez-Guerrero, Herrera, & Garcia, 2008). Whilst some studies have reported significant improvements in psychological health outcomes (Gusi et al., 2008; Lord & Green, 1995; Murphy et al., 2012; Stewart et al., 2017), systematic reviews have concluded no overall significant impact of ERS interventions on psychological health (Campbell et al., 2015; Pavey et al., 2011), with the exception of a weak short-term reduction in depression level (Pavey et al., 2011). Mixed results have also been demonstrated for changes in HRQoL, measured predominantly using the EQ-5D or the 12-item short form survey (SF-12). Several studies have found improvements in HRQoL (Gusi et al., 2008; Pardo et al., 2014; Stewart et al., 2017; Ward et al., 2010), whilst others have found no significant differences (Isaacs et al., 2007; Murphy et al., 2012; Sørensen et al., 2011). An ERS evaluation that is currently in progress by Giné-Garriga et al. (2017), and published as a trial protocol, will add further knowledge concerning overall quality of life by including measures not previously used in ERS evaluations, such as change in social network, loneliness, executive functioning and fear of falling.

6.2.4. Limitations of analyses

There are many complexities to ERS evaluations, which may impact outcomes such as those described. Some of the most significant limitations of evaluations are the heterogeneity in the quality and nature of schemes offered, the diverse range of individuals that are referred to them, and the variation in outcomes collected, which may contribute to the non-significant and inconsistent outcomes reported to-date (NICE, 2014a; Pavey et al., 2011).

The inclusion criteria for referral, based on health conditions present, may differ between ERSs (Campbell et al., 2015; Pavey et al., 2011). Therefore, depending upon this criteria, schemes may accept referrals from a range of participants including the young and old, able bodied and disabled, and from those who are physically inactive and at risk of ill health to those with multiple LTCs. If one considers how diverse the ERS population can be, finding non-significant outcomes at the population level can be reasonably anticipated, with improvements potentially masked by population heterogeneity.

Emerging evidence indicates that referral condition and adherence level may be important characteristics that influence outcomes. When considering PA, one large scale evaluation found significant increases in PA for those referred for CVD risk factors, which were masked at the population level by unchanged PA levels in those referred for mental health conditions (Murphy et al., 2012). In contrast, participants referred for mental health conditions benefitted from greater improvements in psychological health from social engagement (Murphy et al., 2012). Secondly, as could be expected, high adherence to a scheme is positively associated with greater increases in PA (Campbell et al., 2015), which increases the likelihood of positive physiological changes (Williams et al., 2007). This is demonstrated by research finding differences in outcomes between low and high adherers, with high adherers obtaining a greater magnitude of benefit in outcomes such as BMI (Taylor et al., 1998).

Whilst these differences have been acknowledged by systematic reviews, such sub-group analysis, particularly at the level of medical condition is rare, leaving reviewers unable to draw any consensus (Campbell et al., 2015; Pavey et al., 2011). Recently, Rowley et al. (2018) reviewed the evidence for the impact of ERSs on participants with CVD, mental health or musculoskeletal conditions. However, due to lack of sub-group analyses, the authors found limited evidence of impact for these sub-groups.

Considering the above, and that decisions to commission ERSs can be made published evaluations, it is important for future evaluations to acknowledge this heterogeneity and to treat the population as a diverse group, and to identify effectiveness between groups (NICE, 2014a; Pavey et al., 2011; Rowley et al., 2018). This should first include careful consideration of which data are relevant to collect for a population, where the nature of outcome goals may vary between groups. Methods of analysis, and presentation of results should also then be considered, to identify which groups benefit from participation and in what way (Anokye et al., 2011; Pavey et al., 2011), rather than presenting aggregated outcomes for the whole population that offer little opportunity for learning. In doing so, areas for improvement in practice for particular groups may be identified, or referrals targetted for groups in which the benefits of participation are shown to be greatest (Anokye et al., 2011).

A further limitation has been the tendency for previous research to examine the impact of participation over a short to medium duration, such as eight weeks (Webb et al., 2016), three months (Stewart et al., 2017), or six months (McGeechan et al., 2018). Few studies have evaluated the long-term impact of participation in an ERS, such as 12 months or more, and whether any short-term changes in outcomes are sustained (Isaacs et al., 2007; Martín-Borràs et al., 2018; Murphy et al., 2012). This limits the understanding of how effective ERSs are in the long-term, and is a fundamental issue for cost-effectiveness analyses (Anokye et al., 2011). Understanding the long-term impact of participation is therefore an area for future research (Campbell et al., 2015; Pavey et al., 2011).

6.2.5. Study 2 -summary

Systematic reviews have found no strong evidence for a long-term benefit of participation in an ERS on change in PA level and several health-related outcomes (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007), and have drawn the conclusion that schemes have only marginal additional benefits compared to other interventions such as brief advice (NICE, 2014a). Pavey et al. (2011) therefore argue that there is a case for disinvestment in ERSs. However, evaluations have been limited by the complexity of ERS interventions, such as the heterogeneous nature of persons referred, meaning that significant changes in health outcomes may be masked when the population is analysed as a whole (Campbell et al., 2015; Murphy et al., 2012; NICE, 2014a; Pavey et al., 2011).

6.3. Study 2 - Methods

This section provides a description of the methods used to address the aims and objectives detailed in section 6.1.

This study was a quantitative outcome evaluation using secondary data from the

‘Live Active’ ERS scheme. Outcome evaluations are concerned with how effective an intervention has been at producing change in a set of outcomes (Hoggarth & Comfort, 2010; MRC, 2015).

6.3.1. Participants

Individuals who were referred to the ‘Live Active’ scheme between July 2015 – July 2017, and had attended either a week 24 or week 52 consultation were included in this study. New referrals awaiting an initial consultation, and individuals who were participating but had not yet attended a week 24 consultation were excluded.

6.3.2. Referral pathway

The referral pathway was previously detailed in section 3.2.5. and is summarised in this section. Referred participants were invited by the ‘Live Active’ scheme to attend an initial consultation with an exercise professional. During this appointment an overview of the scheme was provided, and the reason for referral and the individual benefits of increasing PA were discussed with the participant. A range of short-term process (12 weeks), medium (24 weeks) and long-term (52 weeks) outcome goals were agreed between the exercise professional and participant, and a programme of weekly physical activities based upon the participant’s needs and preferences was provided. Participants were risk assessed using the Physical Activity Readiness Questionnaire (PAR-Q), and a baseline evaluation was completed (further detailed in section 6.2.3. below). Participants then attended a follow up consultation at weeks 6, 12, and 24 to: evaluate progression in PA behaviour; repeat the baseline evaluation (weeks 12 and 24 only); make appropriate changes to the PA programme if required, and; to offer support and motivation. At week 24, participants were discharged from the scheme, and no additional consultations were offered, with the exception of a follow-up consultation at week 52 to repeat the baseline evaluation.

6.3.3. Data collection

Data routinely collected by the ‘Live Active’ scheme at four time points (weeks 1, 12, 24 and 52) were made available for secondary analysis. The content of the routine data were selected by the developers of the scheme, and was based upon the requirements of the funding organisation and the recommendations of the SEF (Cavill et al., 2012). The findings of this literature review were therefore unable to influence which outcomes were collected. The outcomes collected included readiness to change (TTM Stage of Change), current PA and sedentary behaviour level (IPAQ-SF), HRQoL (EQ-5D-3L and EQ-5D VAS), mental wellbeing (WEMWBS), BMI, BP, alcohol consumption, smoking status, and achievement of goals. These measures were administered face-to-face in consultations, or via telephone consultation by ‘Live Active’ exercise professionals. These outcomes were critically discussed in section 3.5, and are summarised below.

6.3.3.1. Readiness to change

TTM stage of change (Prochaska & DiClemente, 1982) was assessed to identify readiness to become physically active. In order to categorise stage of change, the following statements were presented to individuals:

1. I am not currently very physically active and I don't intend to become more active in the next six months/I'm too busy right now (Pre-contemplation)
2. I am not currently very physically active, but I am thinking about increasing the amount of activity I take in the next six months (Contemplation)
3. The amount of activity I take varies: sometimes I am physically active, other times not (Preparation)
4. I am currently physically active on most days, but have only just begun to be so within the last six months (Action)

5. I am currently physically active on most days, and have been so for longer than six months (Maintenance)
6. A year ago I was physically active on most days, but in the last few months I have been less active (Relapse)

These outcomes were coded categorically based on stage of change (i.e. 1-6).

6.3.3.2. Physical activity and sedentary behaviour

PA and sedentary behaviour were assessed using the IPAQ-SF questionnaire (IPAQ, 2005), which provides both categorical and continuous data (appendix 4). Sedentary behaviour was assessed using the single item sitting time per day (mins) question (appendix 4), and reported as average sitting time per day in minutes.

For the PA section of the IPAQ-SF, participants were asked to recall over the last 7-days, on how many days and for how long they did 10 minutes or more of i) walking, ii) moderate intensity PA, and iii) vigorous intensity PA. According to the IPAQ guidelines (IPAQ, 2005), continuous outputs for MET minutes of PA per week, and kilocalorie expenditure per week were calculated, and reported (see Table 6.1. below for calculations), using the following MET equivalents of activities: walking = 3.3 METs; moderate intensity PA = 4.0 METs, and; vigorous intensity PA = 8.0 METs. The continuous outcomes for MET minutes of PA per week and kilocalorie expenditure per week were presented as a median and interquartile range, as recommended by the developers of the IPAQ (IPAQ, 2005).

Table 6.1. Calculations for IPAQ continuous outputs.

Continuous output	Calculation	Units
Total walking per week	3.3 x minutes x days	MET mins/week
Total moderate intensity PA per week	4.0 x minutes x days	MET mins/week
Total vigorous intensity PA per week	8.0 x minutes x days	MET mins/week
Total PA per week	Walking (3.3 x minutes x days) + Moderate (4.0 x minutes x days) + Vigorous (8.0 x minutes x days)	MET mins/week
Total kilocalorie expenditure per week	MET-mins/week x (weight in kilograms/60 kilograms)	Kilocalories/week

In a further component of the IPAQ-SF scoring, participants were classified into PA categories based on the frequency and intensity of PA reported. Using a scoring algorithm (IPAQ, 2005), participants were assigned to one of three categories: low, moderate or high PA. High referred to those completing at least one hour of moderate intensity PA, or half an hour of vigorous intensity PA, on most days. Moderate referred to those completing at least half an hour of moderate intensity PA on most days. Low referred to those who were active below these thresholds (IPAQ, 2005). These results were presented as categorical data.

One of the most frequent methods of reporting PA has been the proportion of participants achieving ≥90-150 minutes of PA per week (Campbell et al., 2015; Pavey et al., 2011). Therefore, to allow for comparison of PA levels with previous research, participant’s total PA per week (mins) was categorised according to whether they were: completely inactive (0 mins of moderate intensity PA per week); active between 1-89 mins of moderate intensity PA per week, or; active ≥90 mins of moderate intensity PA per week. This was repeated, including PA of any intensity.

6.3.3.3. Body mass index

BMI was calculated using the formula body mass/height^2 (Keys et al., 1972; WHO, 2000), and recorded as continuous data to the nearest 0.01 kg/m^2. Furthermore, the corresponding BMI class was recorded based on the categories: underweight (<18.5 kg/m^2), normal weight (18.5-24.99 kg/m^2), overweight (25-29.99 kg/m^2), or obese (≥30 kg/m^2) (WHO, 2000).

A standardised protocol for the measurement of height was used, whereby height was recorded by the exercise professional to the nearest centimetre using a wall mounted measure. Individuals were asked to stand upright, with their back straight against the wall, and the head in the Frankfort plane. Body weight was measured using scales, and reported to the nearest 0.01 kilogram. Calibration of scales could not be verified. Where height or body weight could not be measured, approximate self-report measures were recorded, although this is not generally an accepted method of measurement as it is liable to under- or over-estimation (Connor Gorber et al., 2007; Maukonen et al., 2018). It was not possible to determine which participants had self-reported height or body weight.

6.3.3.4. Blood pressure

BP measurement was taken by the exercise professional using an automatic BP monitor, with a cuff size appropriate for the size of the upper arm. Calibration of the BP monitor could not be verified. Measurement was taken rested, and in the seated position, at least 5 minutes after assuming this position. The arm was supported, with the forearm at the level of the heart. The measurements were recorded as systolic BP over diastolic BP in mmHg. BP measurement was also categorised as either normal BP, or hypertension (≥140 mmHg and/or ≥90 mmHg) (Williams et al., 2018).

6.3.3.5 Health-related quality of life

HRQoL was assessed using the EQ-5D-3L health profile and EQ-5D VAS (EuroQol Group, 1990) (appendix 5). The responses to the health profile were recorded as a five-digit code (e.g. 11213), which was transformed in to a single summary score, ranging from <0 (worse than being dead) to 1 (full health), using a Microsoft Excel formula (EuroQol Group, n.d.) that assigns weights to each health dimension (Dolan, 1997). The response to the EQ-5D VAS was a single value between 0 (worst imaginable health) and 100 (best imaginable health), and was therefore recorded as continuous data.

6.3.3.6. Mental wellbeing

Mental wellbeing was assessed using the 14-item WEMWBS questionnaire (Tennant et al., 2007) (appendix 6). The score was calculated by summing the response to each item answered on a 1 to 5 Likert scale. The minimum possible score was 14 and the maximum was 70.

6.3.3.7. Smoking status

Smoking status was assessed with a single item measure: ‘Do you currently smoke?’, the response options were ‘yes, or, no’. Use of electronic cigarettes was classified as not smoking. The responses were categorically coded.

6.3.3.8. Alcohol consumption

Alcohol consumption was assessed using two multiple response measures. The first question was ‘How often do you have a drink containing alcohol?’. The response options were: never, monthly or less, 2-4 times per month, 2-3 times per week, 4+ per week. The second question was ‘How many units of alcohol do you drink on a typical day when you are drinking?’. The response options were: 1 or 2, 3 or 4, 5 or

6, 7-9, 10+. The responses to each question were categorically coded.

6.3.3.9 Participant’s goals

During the initial consultation, participants set a range of short-term process, medium-term outcome (1-4 goals), and long-term outcome (1-4 goals) goals. At respective time points (week 12, 24, 52), participants were asked if they had reached these goals, recorded with a yes or no answer. For this study, participant’s medium and long-term outcome goals were categorised by the researcher according to theme, and explored using frequency counts.

6.3.4. Sub-group analysis

A sub-group analysis of medium-term outcomes was conducted according to an individual’s primary referral condition, identified from the referral form and noted on the ERS database. The groups, which aligned with the scheme’s referral pathways were: 1) CVD, or at high risk of CVD; 2) metabolic syndrome or diabetes; 3) any mental health condition; 4) any respiratory condition; 5) any musculoskeletal condition; 6) overweight or obese; 7) multiple (≥3) conditions, or; 8) any other health condition or risk factor. Long-term outcomes (52 weeks) were not analysed due to small sample sizes in the sub-groups.

It is possible that differences in outcomes may also exist between other groups such as age and SES. However, the focus of this study was to build upon emerging findings that outcomes may differ between referral condition groups (Murphy et al., 2012; Rowley et al., 2018). Furthermore, the demographics (e.g. age, SES) of the ‘Live Active’ population were fairly homogenous in comparison to the diverse eligibility criteria regarding health conditions (see earlier section 5.4.2).

6.3.5. Statistical analyses

Data were downloaded from the ERS database and analysed using IBM SPSS version 23 (SPSS, Inc., Chicago, IL, USA). Medium-term outcomes were analysed on a variable-by-variable basis for those with complete data for each respective variable at weeks 1, 12 and 24. Long-term outcomes were analysed on the same basis for those with complete data for each respective variable at weeks 1, 24 and 52. Analysis of missing data (Little's MCAR test) indicated that missing data from the medium-term outcomes were not missing completely at random ($p \leq 0.05$), whilst missing data from the long-term outcomes were missing completely at random ($p > 0.05$). There are no dichotomous choices for treating missing data (Garson, 2015). However, if data are not missing completely at random, multiple imputation of missing values in usually recommended (Donders, van der Heijden, Stijnen, & Moons, 2006; Garson, 2015; Rubin, 1987). However, multiple imputation was not recommended in this case, as the data were not normally distributed (not resolved by transformation), predictive variables had missing values, and more than 50% of participants did not have complete data for BP and WEMWBS values (Table 6.2) (Garson, 2015; Sterne et al., 2009). Therefore, participants were included in the analysis, on a variable-by-variable basis, if they had complete data for the variable in question.

The distributions of continuous outcomes were tested for normality using the Shapiro-Wilk test, and visual inspection of Q-Q plots and histograms. If the distribution was non-normal ($p \leq 0.05$), the median and IQR were reported. A Friedman test (χ^2) was carried out for each continuous variable with non-normal distribution, to identify any significant differences between time points. Where differences were identified ($p \leq 0.05$), post-hoc Wilcoxon signed-rank tests (with

Bonferroni adjustment, $p \leq 0.017$) were used to identify where the differences existed. Categorical data were also analysed using a Friedman test (χ^2), with post-hoc Wilcoxon signed-rank tests (with Bonferroni adjustment, $p \leq 0.017$) when significant differences were indicated ($p \leq 0.05$).

In order to conduct a sub-group analysis of medium-term outcomes, participants were grouped according to their primary referral condition. The analysis was conducted using continuous outcomes, on a variable-by-variable basis for those with complete data for each variable at weeks 1, 12 and 24. Continuous data were not found to be normally distributed. Therefore, the Kruskal Wallis test was used to explore differences in outcomes between groups ($p \leq 0.05$), and Friedman and Wilcoxon signed-rank tests ($p \leq 0.05$) to analyse differences in outcomes between weeks 1, 12 and 24 for each primary referral condition group.

6.4. Study 2 - Results

This section reports the results of data analyses, aiming to address the aims and objectives detailed in section 6.1. The medium-term outcomes for the whole population followed by the sub-group analysis will be presented first, followed by the long-term outcomes.

6.4.1. Participant flow

Between July 2015 and July 2017, 2104 referrals were made to the scheme. Of these, 117 were new referrals awaiting an initial consultation, 378 did not uptake their invitation to participate, 471 had dropped out before week 24, and 515 were still participating (between week 1 and 24), and have therefore been excluded from this analysis. This left 623 participants who had completed the scheme, of which 393 had reached the week 52 follow-up point. Of these 393 participants, 273 had attended a week 52 follow-up consultation in-person or via telephone consultation. A further 120 participants did not attend this consultation (20 declined, 76 unable to contact, 22 time elapsed beyond week 52, 2 died) (Figure 6.1, below).

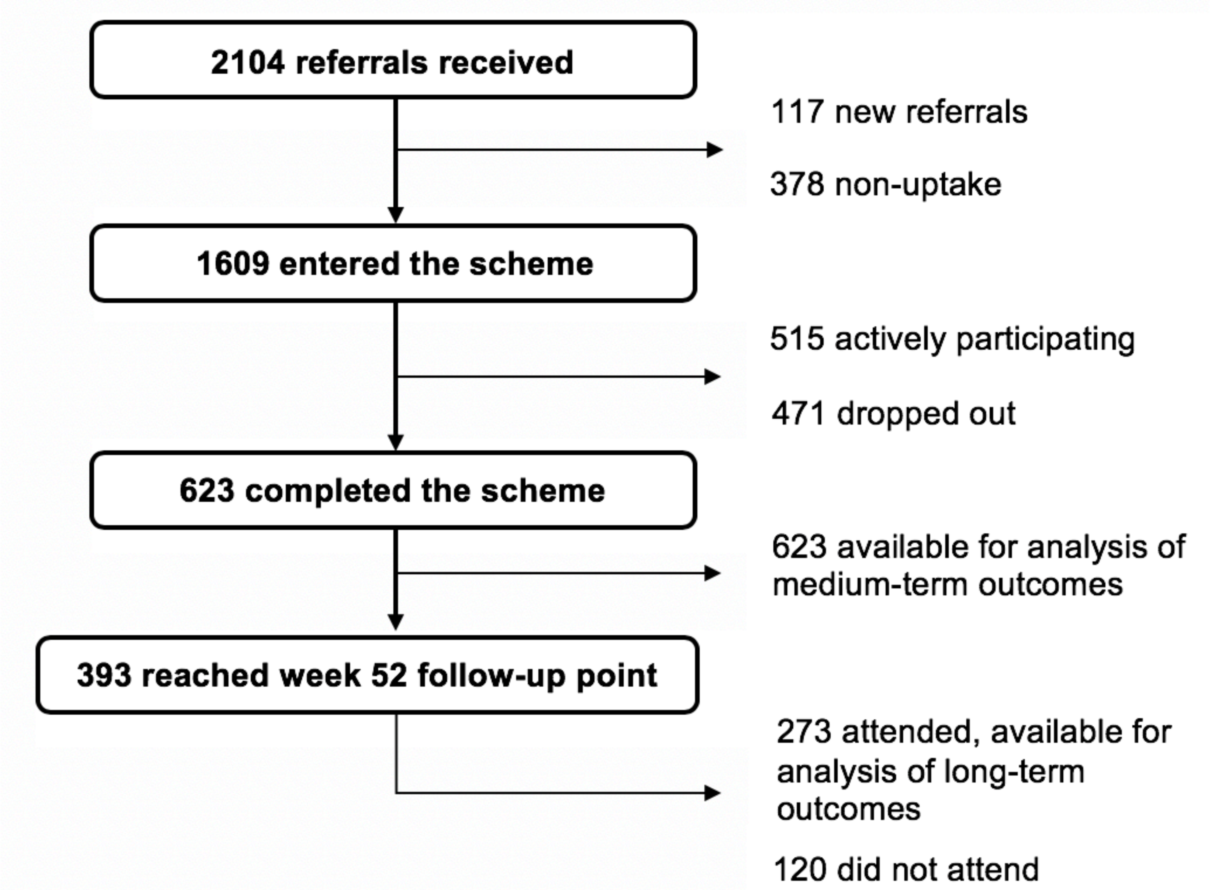


Figure 6.1. Participant flow through the ERS (July 2015 – July 2017).

6.4.2. Medium-term outcomes – differences between weeks 1, 12 and 24

Medium-term outcomes were analysed for 623 participants who had completed the scheme, and had data available at weeks 1, 12 and 24, on a variable by variable basis. Availability of data at all three time points differed between outcomes, as shown below in Table 6.2. Most outcomes had low levels of missing data (<10%), with the exception of BP (57.6% missing) and WEMWBS (51.7% missing).

Table 6.2. Number of participants with or without complete data, at all three time points (weeks 1, 12 and 24), for each outcome variable.

Outcome	Available data		Missing data	
	n	%	n	%
TTM Stage of change	602	96.6	21	3.4
Sitting time	598	96	25	4
MET mins of PA	602	96.6	21	3.4
Kilocalorie expenditure	514	82.5	109	17.5
IPAQ classification	602	96.6	21	3.4
Body weight	514	82.5	109	17.5
BMI	509	81.7	114	18.3
Systolic BP	264	42.4	359	57.6
Diastolic BP	264	42.4	359	57.6
EQ-5D-3L	589	94.5	34	5.5
EQ-5D VAS	590	94.7	33	5.3
WEMWBS	301	48.3	322	51.7
Alcohol consumption	586	94.1	37	5.9
Smoking status	587	94.2	36	5.8

The distributions of continuous variables were found to be non-normal (Shapiro-Wilk $p \leq 0.05$). Therefore, the median (IQR) for each variable at weeks 1, 12, and 24 is presented below in Table 6.3, and the difference between each time point, with statistically significant differences indicated. The table shows significant improvements in all outcomes between weeks 1 and 12, and weeks 1 and 24, with the exception of diastolic BP, which did not significantly change between weeks 1 and 12. In depth results for each variable are reported in the sections below.

Table 6.3. Median (IQR) outcomes for each variable, at each time point.

Outcome	W1	W12	W24	W1-W12	W1-W24	W12-W24
Sitting time (mins per day)	480 (420-720)	480 (300-600)	480 (360-540)	0**	0**	0*
MET mins (per week)	0 (0-231)	540 (240-960)	594 (240-960)	+540**	+594**	+54
Kilocalorie expenditure (per week)	0 (0-345)	819 (376-1484)	918 (347-1516)	+819**	+918**	+99
Body weight (kg)	90.5 (76.8-107.2)	89.75 (76-105)	88.5 (76.2-104)	-0.75**	-2**	-1.25**
BMI (kg/m^2)	32.06 (27.78-37.09)	31.67 (31.67-36.81)	31.25 (27.48-36.72)	-0.39**	-0.81**	-0.42**
Systolic BP (mmHg)	136 (122-146)	132 (121-141)	130 (120-140)	-4*	-6**	-2**
Diastolic BP (mmHg)	82 (73-88)	80.5 (73-86)	80 (73-85)	-1.5	-2**	-0.5
EQ-5D-3L	.727 (.587-.850)	.796 (.689-1.000)	.796 (.689-1.000	+.069**	+.069**	0
EQ-5D VAS	50 (40-70)	60 (50-75)	65 (50-80)	+10**	+15**	+5**
WEMWBS	48 (42-55)	52 (46-58)	52 (47-58)	+4**	+4**	0

*Week (W) 1-W12, W1-W24, and W12-W24 detail the median change in each outcome between the two respective time points. *p ≤0.017, **p ≤0.01 (Wilcoxon signed rank test).*

6.4.2.1. Readiness to change

Frequency counts for TTM stage of change at each time point are presented below in Table 6.4. At week 1, the majority of participants (84.9%) were in the contemplation phase. By week 12, the majority were in the preparation (39.2%) or action phases (41.9%), remaining in these phases at week 24. A small proportion remained in the contemplation phase, indicating that they were physically inactive. Differences in stage of change between the time points was significant ($\chi^2(2)=676.91$, *p* ≤0.001) between week 1 and 12 (z=-18.736, *p* ≤0.001), and week 1 and 24 (z=-18.606, *p* ≤0.001).

Table 6.4. Frequency counts for stage of change at each time point (n=602).

Stage of change	**Week 1**		**Week 12**		**Week 24**	
	n	**%**	**n**	**%**	**n**	**%**
Pre-contemplation	12	2	1	0.2	4	0.7
Contemplation	511	84.9	103	17.1	105	17.4
Preparation	64	10.6	236	39.2	205	34.1
Action	8	1.3	252	41.9	268	44.5
Maintenance	6	1	10	1.7	20	3.3
Relapse	1	0.2	0	0	0	0

Percentages may not always add up to 100 due to rounding errors.

6.4.2.2. Physical activity and sedentary behaviour

PA and sedentary behaviour were evaluated using outcomes from the IPAQ-SF. Percentage distribution of overall IPAQ classification (i.e. low, moderate, high) is presented below in Figure 6.2. Difference in overall classification between time points was significant ($\chi^2(2)=177.133$, $p \leq 0.001$) between weeks 1 and 12 ($z=-10.550$, $p \leq 0.001$), and weeks 1 and 24 ($z=-11.741$, $p \leq 0.001$). The majority of change occurred for participants transitioning between the low to moderate category, indicating an increase in weekly PA.

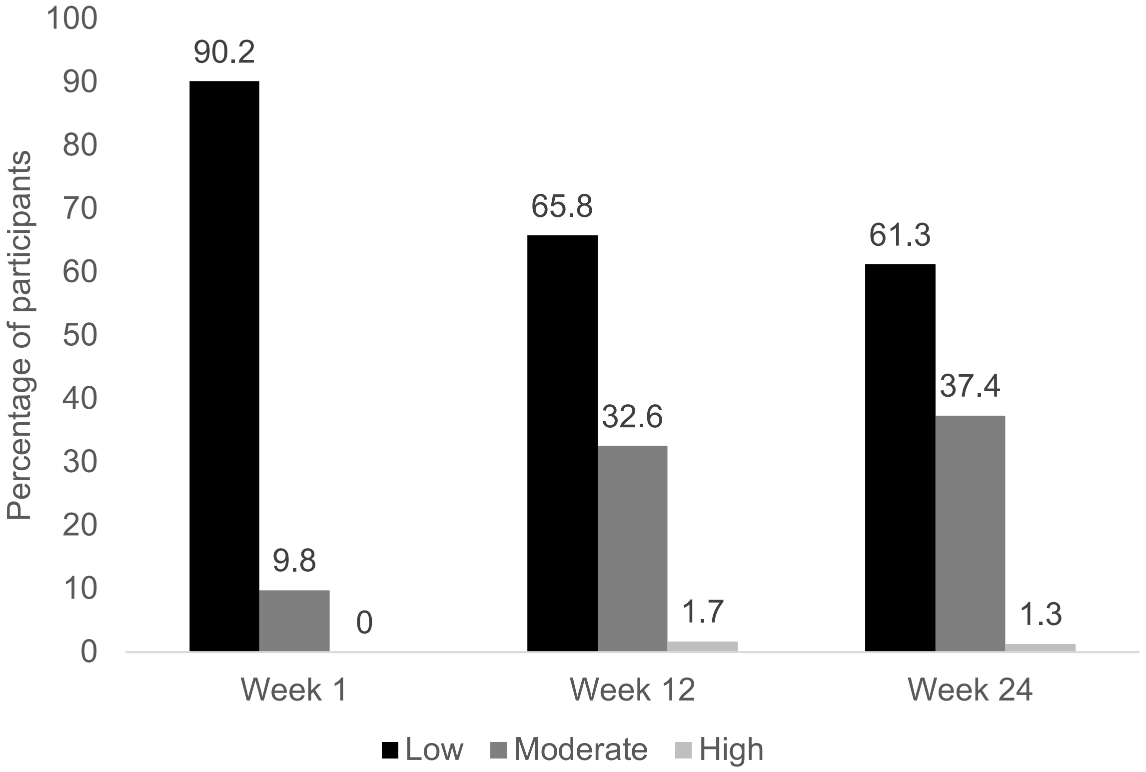


Figure 6.2. Percentage distribution of IPAQ classification at each time point (n=602). Percentages may not always add up to 100 due to rounding errors.

Friedman and Wilcoxon signed ranks tests showed significant improvements in sitting time per day ($\chi^2(2)=166.641$, $p \leq 0.001$), MET minutes of PA per week ($\chi^2(2)=446.121$, $p \leq 0.001$), and kilocalorie expenditure per week ($\chi^2(2)=373.644$, $p \leq 0.001$) between weeks 1 and 12, and weeks 1 and 24 (all $p \leq 0.001$). A breakdown of these outcomes, at each time point, was presented earlier in Table 6.3. Between weeks 1 and 24 there was a median increase in PA of 594 MET minutes per week, and therefore an increase in energy expenditure of 918 kilocalories per week. Whilst the median value for sitting time per day did not change, there was a change in the IQR. Rounded figures for sitting time may be explained by participants rounding sitting time per day to the nearest hour, supported by frequency counts and ERS staff observations.

Considering the frequency and duration of all PA reported, total minutes of PA per week (at any intensity) increased from 0 minutes at week 1, to 150 minutes at week 12, and 160 minutes at week 24 ($\chi^2(2)=433.037$, *p* ≤0.001). The increases between weeks 1 and 12, and weeks 1 and 24 were both significant (both *p* ≤0.001). There were also significant increases in MVPA and low intensity PA (see Table 6.5. below), which were significant between weeks 1 and 12, and weeks 1 and 24 (all *p* ≤0.001).

Table 6.5. Median minutes of PA, per week, at weeks 1, 12 and 24 (n=602).

	Week 1	Week 12	Week 24
Total mins of PA	0	150	160
Total mins of MVPA	0	60	60
Total mins of low intensity PA	0	40	50

For comparison with previous research, the percentage of participants who were physically active, at moderate-vigorous intensity, for ≥90 minutes per week was calculated at each time point, using total minutes of MVPA reported per week. At week 1, 5.6% (n=34) were active ≥90 minutes per week, rising to 44.4% (n=267) at week 12, and 43.2% (n=260) at week 24 ($\chi^2(2)=458.421$, *p* ≤0.001). However, a proportion of participants were active below this threshold (19.4% at week 24), and were therefore not completely inactive, as can be seen below in Table 6.6.

Table 6.6. Number of participants who were inactive, or active above or below the 90-minute MVPA (per week) threshold at each time point (n=602).

PA threshold	Week 1		Week 12		Week 24	
	n	%	n	%	n	%
0 mins	516	85.7	193	32.1	225	37.4
1-89 mins	52	8.6	142	23.6	117	19.4
≥90 mins	34	5.6	267	44.4	260	43.2

The differences in classification were significant between weeks 1 and 12 (z=-16.363, p ≤0.001), weeks 1 and 24 (z=-15.651, p ≤0.001), and weeks 12 and 24 (z=-1.756, p ≤0.001), Percentages may not always add up to 100 due to rounding errors.

The proportion of participants physically active for ≥90 minutes per week increased when PA of any intensity was included in the analysis. As can be observed below in Table 6.7, by taking in to account both total low intensity PA and MVPA, the proportion physically inactive was lower, and the proportion physically active for ≥90 minutes per week was higher. The differences in the proportion of participants physically active for ≥90 minutes per week, at any intensity, was significant between the three time points ($\chi^2(2)=474.964$, $p \leq 0.001$).

Table 6.7. Number of participants who were inactive, or active above or below the 90-minute (per week) threshold at each time point, for PA of any intensity (n=602).

PA threshold	Week 1		Week 12		Week 24	
	n	%	n	%	n	%
0 mins	378	62.8	91	15.1	104	17.3
1-89 mins	101	16.8	97	16.1	77	12.8
≥90 mins	123	20.4	414	68.8	421	69.9

The differences in classification were significant between weeks 1 and 12 (z=-16.409, p ≤0.001), and weeks 1 and 24 (z=-15.996, p ≤0.001). Percentages may not always add up to 100 due to rounding errors.

6.4.2.3. Body composition

Body composition was evaluated using body weight and BMI. There were significant differences in body weight ($\chi^2(2)=57.874$, $p \leq 0.001$) and BMI ($\chi^2(2)=54.638$, $p \leq 0.001$) between the three time points. Body weight was significantly reduced by 0.75kg by week 12 (z=-5.220, $p \leq 0.001$), and by 2kg by week 24 (z=-6.925, $p \leq 0.001$). BMI was significantly reduced by 0.39 kg/m^2 by week 12 (z=-5.104, $p \leq 0.001$), and by 0.81 kg/m^2 by week 24 (z=-6.677, $p \leq 0.001$). Despite reductions in BMI, when categorised according to BMI class, there were minimal, non-significant changes in BMI class between the three time points ($\chi^2(2)=4.022$, $p >0.05$).

When participants were stratified according to BMI class at week 1 (Table 6.8, below), there were no significant changes in BMI between weeks 1 and 24 for participants classified as normal weight (-0.22 kg/m^2, z=-1.209, $p >0.017$). There were, however, significant improvements in BMI for participants classified as overweight (-0.20 kg/m^2, z=-2.642, $p \leq 0.01$) or obese (-0.43 kg/m^2, z=-6.194, p

≤0.001). Changes in the underweight category were not analysed due to the sample size (n=1).

Table 6.8. Median BMI at weeks 1, 12 and 24 when stratified according to BMI class at week 1.

	n	Week 1	Week 12	Week 24
Underweight	1	18.05	17.31	17.83
Normal weight	55	23.59	23.37	23.37
Overweight	149	27.83	27.66	27.63
Obese	304	36.06	35.77	35.63

6.4.2.4. Blood pressure

Compliance with BP measurement was the lowest of any outcome, with data available for 42.4% (n=264) of participants. Between weeks 1 and 12 there was a significant improvement in systolic BP (z=-2.479, *p* ≤0.017) but not diastolic BP (z=-1.643, *p* >0.017). By week 24, compared to week 1, there were significant improvements in both systolic BP (z=-4.795, *p* ≤0.001) and diastolic BP (z=-3.479, *p* ≤0.001). Between weeks 1 and 24 systolic BP was reduced by 6mmHg from 136 (122-146) mmHg to 130 (120-140) mmHg, and diastolic BP was reduced by 2 mmHg from 82 (73-88) mmHg to 80 (73-85) mmHg.

BP was categorised according to whether the measurement could be classified as hypertension or not (systolic BP ≥140mmHg and/or diastolic BP ≥90 mmHg), to understand the impact of participation on changes in BP on classification. Significant changes were found in the proportion of participants with a BP measurement that could be classified as hypertension ($\chi^2(2)=27.75$, *p* ≤0.001). At week 1, 43.9% (n=116) of participants had a BP measurement classified as hypertension. This reduced to 33% (n=87) at week 12, and 27.7% (n=73) at week 24. The differences in classification between weeks 1 and 12 (z=-3.491, *p* ≤0.001), and weeks 1 and 24 (z=-4.965, *p* ≤0.001) were significant.

6.4.2.5. HRQoL

HRQoL outcomes included the EQ-5D-3L and the EQ-5D VAS. Significant improvements were found in the EQ-5D-3L ($\chi^2(2)=55.561$, $p \leq 0.001$) and EQ-5D VAS ($\chi^2(2)=191.141$, $p \leq 0.001$). Between weeks 1 and 12, the EQ-5D-3L summary score increased from .727 (.587-.850) to .796 (.689-1.000) (z=-5.735, $p \leq 0.001$), and was maintained at week 24 (z=-5.735, $p \leq 0.001$). The percentage of participants reporting problems with each health profile state is presented below in Figure 6.3. as recommended (Reenen & Oppe, 2015). The most noticeable change appeared to be in anxiety/depression; there was a 7.5% increase in participants reporting no problems with this health state between weeks 1 and 24. The EQ-5D VAS score increased from 50 (40-70) to 60 (50-75) between weeks 1 and 12 (z=-9.686, $p \leq 0.001$), and from 50 (40-70) to 65 (50-80) between weeks 1 and 24 (z=-11.107, $p \leq 0.001$).

6.4.2.6. Mental wellbeing

Compliance with collection of the WEMWBS questionnaire was low, with data available for 48.3% (n=301) of participants. Significant improvements in mental wellbeing were found between weeks 1 and 12 (z=-8.027, $p \leq 0.001$), and weeks 1 and 24 (z=-7.351, $p \leq 0.001$). The WEMWBS score increased from 48 (42-55) to 52 (46-58) between weeks 1 and 12, and was maintained at week 24.

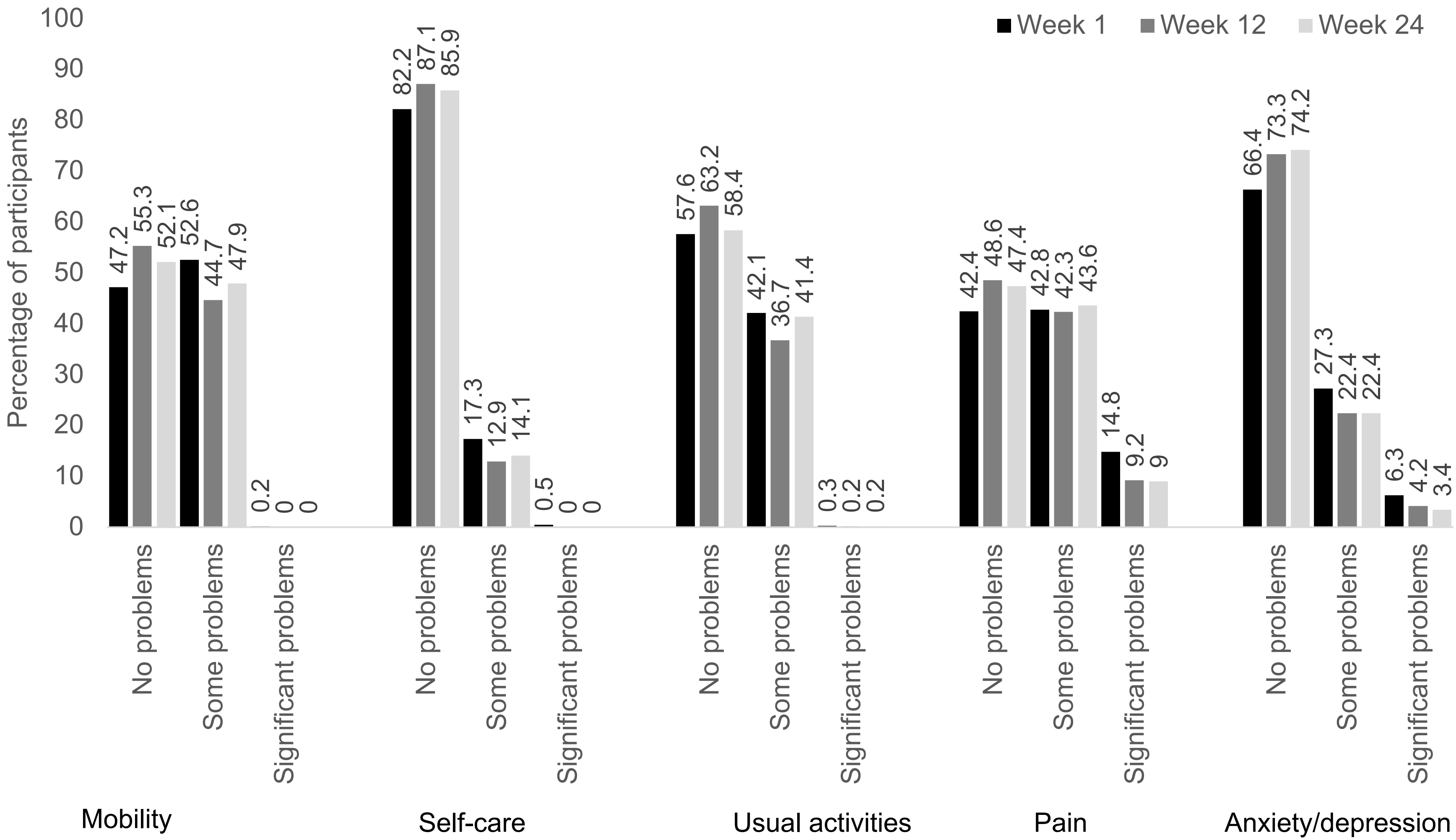


Figure 6.3. Percentage of participants reporting no/some/significant problems with each EQ-5D-3L health profile state at weeks 1, 12 and 24.

6.4.2.7. Smoking status and alcohol consumption

There were no significant differences in smoking status ($\chi^2(2)=4.348$, $p > 0.05$) or frequency of alcohol consumption ($\chi^2(2)=1.99$, $p > 0.05$) between the three time points. There were, however, significant differences in the number of alcohol units consumed per drinking episode ($\chi^2(2)=29.335$, $p \leq 0.001$), between week 1 and 12 ($z=-3.257$, $p \leq 0.001$), week 1 and 24 ($z=-5.119$, $p \leq 0.001$), and week 12 and 24 ($z=-2.627$, $p \leq 0.01$), with a shift in the level of consumption, from a higher number (≥5 units) to a lower number of units (≤4 units) consumed per drinking episode.

6.4.2.8. Medium term goals

Data on the achievement of medium-term goals were available for 613 out of 623 participants. The most frequently reported medium-term goals are presented below in Table 6.9. Of 613 participants, 447 (72.9%) reported that they had met their medium-term goals. Due to the method of reporting achievement of goals by the ERS, it was not possible to determine the percentage of participants that had achieved each individual goal. Rather, 72.9% of participants had achieved a cluster of goals, predominantly comprised of those reported in Table 6.9.

Table 6.9. Most frequently reported medium-term goals (n=623).

Goal	%	n
Increase or maintain PA level	51.2	319
Reduce BMI	21.3	133
Improve physical fitness	21	131
Improve physical strength	14.9	93
Improve mobility	11.4	71
Improve aspects of physical health (e.g. BP)	7.2	45
Improve flexibility	6.4	40
Reduce pain levels	4.8	30
Improve mental health	3.4	21
Improve limitation free walking distance (i.e. claudication)	3.4	21

% and n refers to the percentage and number of particpants reporting each goal.

6.4.3. Sub-group analysis

Sub-group analysis was conducted for participants with data available at weeks 1, 12 and 24, on a variable by variable basis. Groups were analysed by their primary referral condition. The median (IQR) for each outcome, by sub-group, is summarised below in Tables 6.10, 6.11. and 6.12.

Table 6.10. Median (IQR) outcomes for sitting time per day, MET mins per week and kilocalorie expenditure per week, at weeks 1 and 24, for each sub-group.

	Sitting time per day (mins)			MET mins per week			Kilocalorie expenditure per week		
	n	W1	W24	n	W1	WK24	n	W1	WK24
CVD/high risk of CVD	110	480 (420-720)	480 (300-480)*	111	0 (0-372)	720 (300-1116)*	96	0 (0-645)	1117 (480-1578)*
Diabetes/metabolic syndrome	27	480 (360-540)	420 (360-480)	27	0 (0-292)	742 (396-1108)*	26	0 (0-310)	1157 (607-1498)*
Mental health issue	31	540 (480-720)	480 (360-480)*	31	0 (0-66)	660 (180-990)*	25	0 (0-89)	1051 (310-1676)*
Respiratory condition	21	480 (360-600)	420 (360-480)	21	0 (0-132)	495 (240-772)*	18	0 (0-252)	777 (320-892)*
Musculoskeletal condition	107	480 (360-720)	420 (300-600)*	107	0 (0-198)	495 (0-1032)*	91	0 (0-287)	788 (209-1515)*
≥Overweight	94	540 (360-720)	420 (300-480)*	94	0 (0-264)	693 (297-1032)*	72	0 (0-419)	1066 (531-1796)*
Multiple conditions	193	600 (480-720)	480 (360-600)*	196	0 (0-132)	522 (198-852)*	171	0 (0-188)	833 (316-1473)*
Other condition	15	480 (360-600)	480 (240-480)	15	0 (0-240)	636 (300-876)*	15	0 (0-434)	964 (395-1426)*

*† significant difference between the groups at respective time point (Kruskal Wallis test p ≤0.05). * significant difference in outcome, for sub-group, between weeks 1 and 24 (Wilcoxon signed rank test p ≤0.05). n, number of participants; w,week.*

Table 6.11. Median (IQR) outcomes for BMI, systolic BP and diastolic BP, at weeks 1 and 24, for each sub-group.

	BMI			Systolic BP			Diastolic BP		
	n	W1†	W24†	n	W1	W24†	n	W1	W24
CVD/high risk of CVD	95	30.02 (27.46-34.54)	30.17 (27.07-35.08)*	55	139 (121-150)	130 (121-145)	55	81 (71-89)	78 (71-86)
Diabetes/metabolic syndrome	24	31.53 (28.74-34.38)	31.64 (29.15-34.15)	17	138 (135-145)	130 (125-136)*	17	81 (70-85)	78 (70-85)
Mental health issue	25	28.96 (26.13-33.67)	27.92 (26.32-31.96)	17	128 (118-137)	130 (123-137)	17	76 (69-85)	81 (75-86)
Respiratory condition	18	26.47 (23.65-27.89)	25.59 (22.23-28.76)	9	131 (121-140)	126 (122-135)	9	79 (72-82)	80 (72-81)
Musculoskeletal condition	91	28.73 (24.66-32.06)	28.03 (24.56-31.76)*	57	130 (121-141)	125 (115-134)*	57	81 (75-86)	76 (73-82)*
≥Overweight	71	36.07 (32.78-40.57)	35.03 (31.47-40.17)	27	135 (120-140)	133 (120-140)	27	84 (78-91)	80 (75-86)
Multiple conditions	170	34.94 (29.23-39.79)	33.65 (29.48-38.14)*	74	137 (123-146)	134 (122-142)	74	84 (75-89)	80 (75-87)
Other condition	15	30.52 (27.78-37.64)	32.21 (27.78-38.01)	8	142 (123-153)	132 (120-140)	8	88 (81-91)	85 (79-87)

*† significant difference between the groups at respective time point (Kruskal Wallis test p ≤0.05). * significant difference in outcome, for sub-group, between weeks 1 and 24 (Wilcoxon signed rank test p ≤0.05). n, number of participants; w,week.*

Table 6.12. Median (IQR) outcomes for EQ-5D-3L, EQ-5D VAS and WEMWBS, at weeks 1 and 24, for each sub-group.

	EQ-5D-3L			EQ-5D VAS			WEMWBS		
	n	W1†	W24†	n	W1†	W24†	n	W1†	W24†
CVD/high risk of CVD	109	.796 (.623-1.000)	.796 (.691-1.000)	109	60 (50-75)	70 (50-80)*	65	52 (46-58)	54 (50-61)*
Diabetes/metabolic syndrome	27	.848 (.796-1.000)	1.000 (.691-1.000)	27	70 (50-75)	80 (60-85)*	16	49 (45-55)	57 (53-61)*
Mental health issue	30	.734 (.378-.848)	.804 (.620-1.000)*	30	50 (40-60)	60 (50-75)*	13	38 (27-41)	50 (42-55)*
Respiratory condition	21	.814 (.691-.883)	.814 (.743-1.000)	21	66 (50-75)	60 (50-70)	14	48 (45-51)	50 (46-51)
Musculoskeletal condition	105	.691 (.516-.796)	.691 (.620-.848)*	106	50 (45-65)	60 (50-80)*	57	50 (41-57)	50 (45-57)*
≥Overweight	88	.848 (.656-1.000)	.848 (.709-1.000)*	88	50 (40-65)	65 (53-75)*	41	47 (44-54)	51 (45-55)*
Multiple conditions	194	.691 (.561-.848)	.727 (.656-848)*	194	50 (30-65)	60 (50-75)*	86	48 (42-54)	51 (45-58)*
Other condition	15	.710 (.639-1.000)	.814 (.710-1.000)	15	60 (40-70)	70 (50-75)	9	39 (39-47)	49 (49-55)

*† significant difference between the groups at respective time point (Kruskal Wallis test p ≤0.05). * significant difference in outcome, for sub-group, between weeks 1 and 24 (Wilcoxon signed rank test p ≤0.05). n, number of participants; w,week.*

Kruskal Wallis tests showed no significant differences in PA between primary referral condition groups at week 1 or week 24. At week 1, all groups were physically inactive, and by week 24 all groups had significantly increased PA levels. The largest increase in PA was seen in those referred for diabetes or metabolic syndrome (+742 MET mins, *p* ≤0.001), with the smallest increases seen in those referred for respiratory (+495 MET mins, *p* ≤0.01) or musculoskeletal conditions (+495 MET mins, *p* ≤0.001).

BMI significantly differed between the groups at both time points (both *p* ≤0.001). Those referred for being overweight or obese had the highest BMI at both time points, but did not significantly reduce BMI (both *p* >0.05). Only three groups experienced significant reductions, those referred for: CVD or being at high risk of CVD (*p* ≤0.05); musculoskeletal conditions (*p* ≤0.05), or; multiple conditions (*p* ≤0.001).

There were few significant changes in BP between the groups. Only those referred for musculoskeletal conditions reduced both systolic BP (*p* ≤0.01) and diastolic BP (*p* ≤0.01). The only remaining significant improvement in BP was for those referred for diabetes or metabolic syndrome, who reduced systolic BP only (*p* ≤0.01). There were no other significant changes in BP.

There were significant differences between the groups in HRQoL and mental wellbeing at both time points. However, those referred for respiratory or other conditions did not significantly improve in any of these measures. The largest significant improvement in HRQoL was seen in those referred for mental health conditions (EQ-5D-3L, +.070, *p* ≤0.05), as was the largest improvement in mental wellbeing (WEMWBS +12, *p* ≤0.01).

In summary, sub-group analyses demonstrated participation, in terms of these outcomes, to be least effective for those referred for respiratory related, or other conditions. These groups significantly increased their PA levels, but did not significantly improve in any other outcomes measured. Participation in this ERS, in terms of these outcomes, appeared to be most effective for those referred for musculoskeletal conditions, who had the lowest median increase in PA, but were the only group to experience significant changes in all outcomes.

6.4.4. Long-term outcomes – differences between weeks 1, 24 and 52

Long-term outcomes were analysed for 273 participants who attended a week 52 consultation, and had data available at weeks 1, 24 and 52, on a variable by variable basis. The demographics, source of referral, and primary referral condition did not differ significantly between those who did or did not attend the week 52 follow-up (Chi-square test $p > 0.05$) with the exception of age; those who attended were significantly older ($t(385) = 2.306$, $p \leq 0.05$).

Availability of data at all three time points differed between outcomes, as shown below in Table 6.13. Most outcomes had low levels of missing data (<20%), with the exception of BP (70.3% missing) and WEMWBS (61.9% missing).

Table 6.13. Number of participants with or without complete data, at all three time points (weeks 1, 24 and 52), for each outcome variable.

Outcome	Available data		Missing data	
	n	%	n	%
TTM Stage of change	271	99.3	2	0.7
Sitting time	269	98.5	4	1.5
MET mins	271	99.3	2	0.7
Kilocalorie expenditure	223	81.7	50	18.3
IPAQ classification	271	99.3	2	0.7
Body weight	224	82.1	49	17.9
BMI	222	81.3	51	18.7
Systolic BP	81	29.7	192	70.3
Diastolic BP	81	29.7	192	70.3
EQ-5D-3L	266	97.4	7	2.6
EQ-5D VAS	266	97.4	7	2.6
WEMWBS	104	38.1	169	61.9
Alcohol consumption	268	98.2	5	1.8
Smoking status	268	98.2	5	1.8

The distributions of continuous variables were found to be non-normal (Shapiro-Wilk *p* ≤0.05). Therefore, the median (IQR) for each variable at week 1, 24, and 52 is presented below in Table 6.14, and the difference between each time point, with statistically significant differences indicated. The table shows significant improvements in all outcomes between weeks 1 and 24, and weeks 1 and 52, with the exception of diastolic BP, which did not significantly differ at any time point. In depth results for each variable are reported in the sections below.

Table 6.14. Median (IQR) outcomes for each variable, at each time point.

Outcome	W1	W24	W52	W1-W24	W1-W52	W24-W52
Sitting time (mins per day)	480 (360-660)	420 (300-480)	360 (300-480)	-60**	-120**	-60
MET mins (per week)	0 (0-198)	660 (360-960)	636 (248-960)	+660**	+636**	-24
Kilocalorie expenditure (per week)	0 (0-289)	1051 (582-1583)	941 (461-1456)	+1051**	+941**	-110
Body weight (kg)	91.3 (76.6-107.2)	89 (76.8-104.8)	88 (76.1-104)	-2.3**	-3.3**	-1**
BMI (kg/m^2)	32.7 (27.82-37.64)	31.68 (27.68-36.91)	31.42 (27.51-35.99)	-1.02**	-1.28**	-0.26**
Systolic BP (mmHg)	138 (127-148)	131 (120-140)	129 (120-138)	-7**	-9**	-2
Diastolic BP (mmHg)	80 (75-89)	80 (73-85)	81 (76-85)	0	+1	+1
EQ-5D-3L	.727 (.587-.850)	.796 (.691-1.000)	.796 (.691-1.000)	+.069**	+.069**	0
EQ-5D VAS	50 (40-70)	70 (50-80)	70 (50-80)	+20**	+20**	0**
WEMWBS	50 (42-58)	53 (48-59)	55 (50-60)	+3**	+5**	+2**

*W1-W24, W1-W52, and W24-W52 detail the median change in each outcome between the two respective time points. *≤.017, **<.01 (Wilcoxon signed rank test).*

6.4.4.1. Readiness to change

Frequency counts for TTM stage of change at each time point are presented in Table 6.15. At week 1, the majority of participants (86%) were in the pre-contemplation phase. By week 24, the majority were in the preparation (37.6%) or action phases (46.5%), where they either remained or moved to the maintenance phase (30.5%) by week 52. A small proportion remained in the contemplation phase at both weeks, indicating that they were physically inactive. Differences in stage of change between the time points was significant ($\chi^2(2)=314.179$, $p \leq 0.001$) between weeks 1 and 24 ($z=-13.019$, $p \leq 0.001$), weeks 1 and 52 ($z=-12.643$, $p \leq 0.001$), and weeks 24 and 52 ($z=-4.742$, $p \leq 0.001$).

Table 6.15. Frequency counts for stage of change at each time point (n=271).

	Week 1		Week 24		Week 52	
Stage of change	**n**	**%**	**n**	**%**	**n**	**%**
Pre-contemplation	4	1.5	0	0	1	0.4
Contemplation	233	86	31	11.4	45	16.6
Preparation	27	10	102	37.6	75	27.7
Action	3	1.1	126	46.5	42	15.5
Maintenance	3	1.1	12	4.4	107	39.5
Relapse	1	0.4	0	0	1	0.4

Percentages may not always add up to 100 due to rounding errors.

6.4.4.2. Physical activity and sedentary behaviour

Percentage distribution of overall IPAQ classification is presented below in Figure 6.4. There was a significant change in overall IPAQ classification ($\chi^2(2)=115.209$, $p \leq 0.001$), between weeks 1 and 24 ($z=-8.324$, $p \leq 0.001$), and weeks 1 and 52 ($z=-9.466$, $p \leq 0.001$), with the majority of change occurring for participants transitioning from the low to moderate category, indicating an increase in weekly PA.

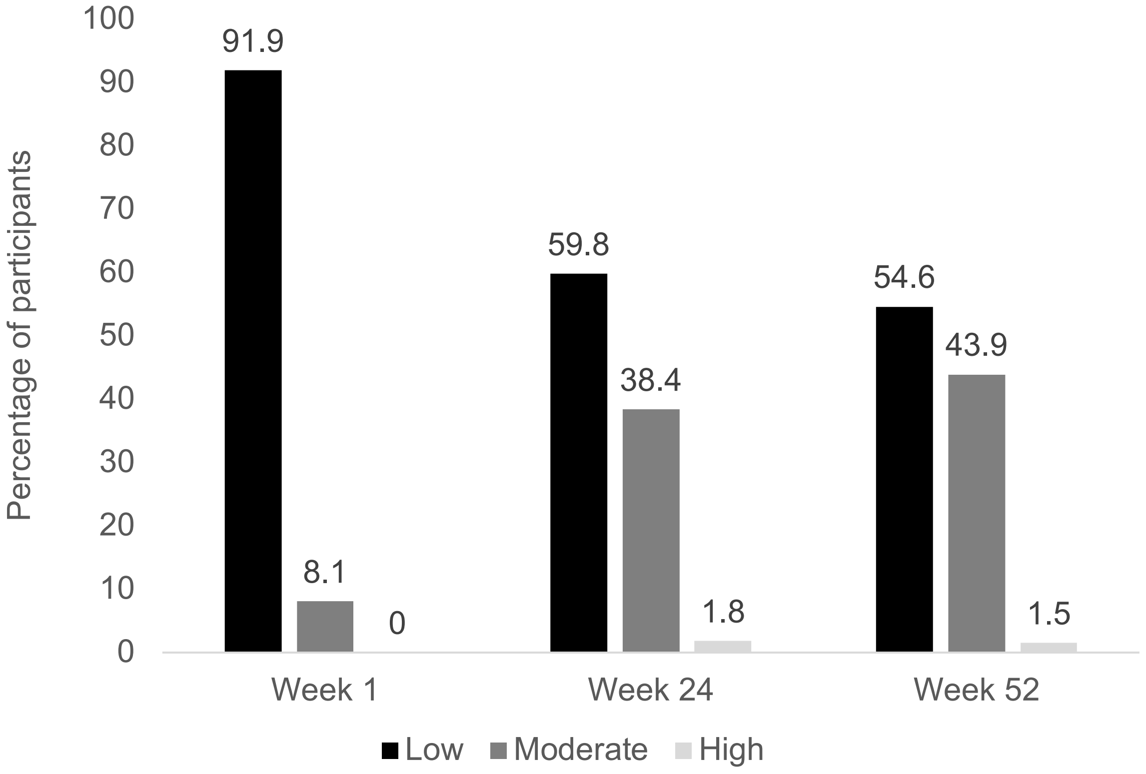


Figure 6.4. Percentage distribution of overall IPAQ classification at each time point (n=271). Percentages may not always add up to 100 due to rounding errors.

Friedman and Wilcoxon signed ranks tests showed significant improvements in sitting time per day, total MET minutes of PA per day, and kilocalorie expenditure per day between weeks 1 and 24 (all *p* ≤.001), and weeks 1 and 52 (all *p* ≤0.001). Sitting time per day reduced by 60 minutes between weeks 1 and 24 (z=-7.667, *p* ≤0.001), and by 120 minutes between weeks 1 and 52 (z=-7.836, *p* ≤0.001). Rounded figures for sitting time may be explained by participants rounding sitting time per day to the nearest hour, supported by frequency counts and ERS staff observations.

MET minutes of PA per week increased by 660 MET minutes between weeks 1 and 24 (z=-11.772, *p* ≤0.001), and by 636 MET minutes between weeks 1 and 52 (z=-11.923, *p* ≤0.001). As a result of increased PA, kilocalorie expenditure per week increased by 1051 kilocalories between weeks 1 and 24 (z=-10.816, *p* ≤0.001), and by 941 kilocalories between weeks 1 and 52 (z=-11.025, *p* ≤0.001). A breakdown of these outcomes at weeks 1, 24 and 52 were presented earlier in Table 6.14.

Considering the frequency and duration of PA reported, total minutes of PA per week (of any intensity) increased from 0 minutes at week 1, to 180 minutes at week 24, remaining at 180 minutes at week 52 ($\chi^2(2)=217.683$, $p \leq 0.001$). The increases between weeks 1 and 24, and weeks 1 and 52 were both significant (both $p \leq 0.001$). There were also significant increases in MVPA, and low intensity PA (see table 6.16. below), which were significant between weeks 1 and 24, and weeks 1 and 52 (all $p \leq 0.001$). The decrease in MVPA was also significant between weeks 24 and 52 ($z=-2.892$, $p \leq 0.001$).

Table 6.16. Median minutes of PA, per week, at weeks 1, 24 and 52 (n=271).

	Week 1	Week 24	Week 52
Total mins of PA	0	180	180
Total mins of MVPA	0	70	60
Total mins of low intensity PA	0	60	70

For comparison with previous research, the percentage of participants who were physically active, at moderate-vigorous intensity, for ≥90 minutes per week was calculated at each time point, using total minutes of MVPA reported per week. At week 1, 5.5% (n=15) were active ≥90 minutes per week, rising to 48% (n=130) at week 24, declining to 40.2% (n=109) at week 52 ($\chi^2(2)=212.498$, $p \leq 0.001$). However, a proportion of participants were active below this threshold (17.3% at week 52), and were therefore not completely inactive, as can be seen below in Table 6.17.

Table 6.17. Number of participants who were inactive, or active above or below the 90-minute MVPA (per week) threshold at each time point (n=271).

PA threshold	Week 1		Week 24		Week 52	
	n	%	n	%	n	%
0 mins	236	87.1	85	31.4	115	42.4
1-89 mins	20	7.4	56	20.7	47	17.3
≥90 mins	15	5.5	130	48	109	40.2

The differences in classification were significant between weeks 1 and 24 (z=-11.272, p ≤0.001), weeks 1 and 52 (z=-10.048, p ≤0.001), and weeks 24 and 52 (z=-3.627, p ≤0.001), Percentages may not always add up to 100 due to rounding errors.

The proportion physically active for ≥90 minutes per week increased when PA of any intensity was included in the analysis. As can be observed below in Table 6.18, by taking in to account both total low intensity PA and MVPA, the proportion inactive was lower, and the proportion physically active for ≥90 minutes per week was higher. The differences in the proportion of participants physically active for ≥90 minutes per week, at any intensity, was significant between the three time points ($\chi^2(2)=231.586$, $p \leq 0.001$).

Table 6.18. Number of participants who were inactive, or active above or below the 90-minute threshold (per week) at each time point, for PA of any intensity (n=271).

PA threshold	Week 1		Week 24		Week 52	
	n	%	n	%	n	%
0 mins	167	61.6	40	14.8	46	17
1-89 mins	57	21	25	9.2	28	10.3
≥90 mins	47	17.3	206	76	197	72.7

The differences in classification were significant between weeks 1 and 24 (z=-11.378, p ≤0.001), and weeks 1 and 52 (z=-11.213, p ≤0.001). Percentages may not always add up to 100 due to rounding errors.

Taking the conservative assumption that all those lost to follow-up made no improvement in PA level, 11 referrals needed to be made for one participant to become long-term active (≥90 minutes per week at any intensity), or 19 referrals for one participant to become long-term active at moderate-vigorous intensity.

6.4.4.3. Body composition

Analyses showed significant improvements in body weight ($\chi^2(2)=68.425$, $p \leq 0.001$) and BMI ($\chi^2(2)=67.499$, $p \leq 0.001$) between all three time points. Compared to week 1, the median reduction in body weight was 2.3kg at week 24 (z=-5.090, $p \leq 0.001$) and 3.3kg at week 52 (z=-7.139, $p \leq 0.001$). The difference in body weight between weeks 24 and 52 was also statistically significant (z=-4.678, $p \leq 0.001$). Changes in body composition were also observed through improvements in BMI. Compared to week 1, the median reduction in BMI was 1.02 kg/m^2 at week 24 (z=-4.936, $p \leq 0.001$), and 1.28 kg/m^2 at week 52 (z=-7.064, $p \leq 0.001$). The difference in BMI between weeks 24 and 52 was also statistically significant (z=-4.727, $p \leq 0.001$).

When categorised according to BMI class, at all three time points, the majority of participants were either overweight or obese. There were significant differences in BMI class between the three time points ($\chi^2(2)=9.023$, $p \leq 0.05$), with post hoc analysis finding significant differences between week 24 and 52 only (z=-2.683, $p \leq 0.01$). Frequency counts demonstrated a decline in the proportion classified as overweight obese, and an increase in the proportion classified as normal weight (Table 6.19).

Table 6.19. Number of participants in each BMI class, at each time point (n=222).

BMI Class	**Week 1**		**Week 24**		**Week 52**	
	n	**%**	**n**	**%**	**n**	**%**
Underweight	0	0	0	0	0	0
Normal weight	26	11.7	29	13	34	15.2
Overweight	64	28.7	58	26	60	26.9
Obese	133	59.6	136	61	129	57.8

Percentages may not always add up to 100 due to rounding errors.

When participants were stratified according to BMI class at week 1 (Table 6.20), there were significant improvements in BMI between weeks 1 and 52 for participants classified as normal weight (-0.52 kg/m^2, z=-2.829, *p* ≤0.01) and overweight (-0.59 kg/m^2, z=-2.475, *p* ≤0.017). The biggest improvement was in those classified as obese at baseline (-1.42 kg/m^2, z=-6.313, *p* ≤0.001).

Table 6.20. Median BMI (kg/m^2) at weeks 1, 24 and 52 when stratified according to BMI class at week 1.

	n	Week 1	Week 24	Week 52
Normal weight	26	23.68	23.22	23.16
Overweight	64	28.25	27.85	27.66
Obese	133	36.30	35.68	34.88

6.4.4.4. Blood pressure

Compliance with BP measurement was the lowest of any outcome, with data only available for 29.7% (n=81) participants. Conversations with exercise professionals collecting the data suggests that this was due to many of the week 52 consultations being conducted via telephone. There were no significant differences in diastolic BP between the three time points ($\chi^2(2)=5.353$, *p* >0.05). There were however significant improvements in systolic BP ($\chi^2(2)=32.589$, *p* ≤0.001) between weeks 1 and 24 (z=-4.010, *p* ≤0.001), and weeks 1 and 52 (z=-4.150, *p* ≤0.001). Between week 1 and 52, systolic BP was reduced by 9mmHg from 138 (127-148) mmHg to 129 (120-138) mmHg.

There were also significant changes in the proportion of participants with a BP measurement that could be classified as hypertension ($\chi^2(2)=15.316$, *p* ≤0.001). At week 1, 49.4% (n=40) of participants had a BP measurement that could be classified as hypertension (systolic BP ≥140mmHg and/or diastolic BP ≥90 mmHg). This reduced to 32.1% (n=26) at week 24, and 25.9% (n=21) at week 52. The reductions between weeks 1 and 24 (z=-2.646, *p* ≤0.01), and weeks 1 and 52 (z=-3.413, *p* ≤0.001) were significant.

6.4.4.5. HRQoL

Significant improvements were found in the EQ-5D-3L ($\chi^2(2)=41.913$, $p \leq 0.001$) and EQ-5D-VAS ($\chi^2(2)=115.480$, $p \leq 0.001$). Between weeks 1 and 24, the EQ-5D-3L summary score increased from .727 (.587-.850) to .796 (.691-1.000) (z=-4.329, $p \leq 0.001$), and was maintained at week 52 (z=-5.992, $p \leq 0.001$). The percentage of participants reporting a problem with each health profile state is presented below in Figure 6.5, as recommended (Reenen & Oppe, 2015). The most noticeable changes appeared to be in anxiety/depression, pain, and mobility; there was a 15.4%, 10.9% and 10.6% increase in participants reporting no problems with these health states respectively between weeks 1 and 52. The EQ-5D-VAS score increased from 50 (40-70) to 70 (50-80) (z=-7.279, $p \leq 0.001$) between weeks 1 and 24, and was maintained at week 52 (z=-8.812, $p \leq 0.001$).

6.4.4.6. Mental wellbeing

Compliance with collection of the WEMWBS questionnaire was low, with data available for 38.1% (n=104) of participants. Conversations with exercise professionals collecting the data suggests that this was due to time constraints encountered when conducting telephone consultations at week 52. There were significant improvements in mental wellbeing between the three time points ($\chi^2(2)=48.873$, $p \leq 0.001$). The WEMWBS score increased from 50 (42-58) to 53 (48-59) between weeks 1 and 24 (z=-4.395, $p \leq 0.001$), and from 50 (42-58) to 55 (50-60) between weeks 1 and 52 (z=-5.895, $p \leq 0.001$).

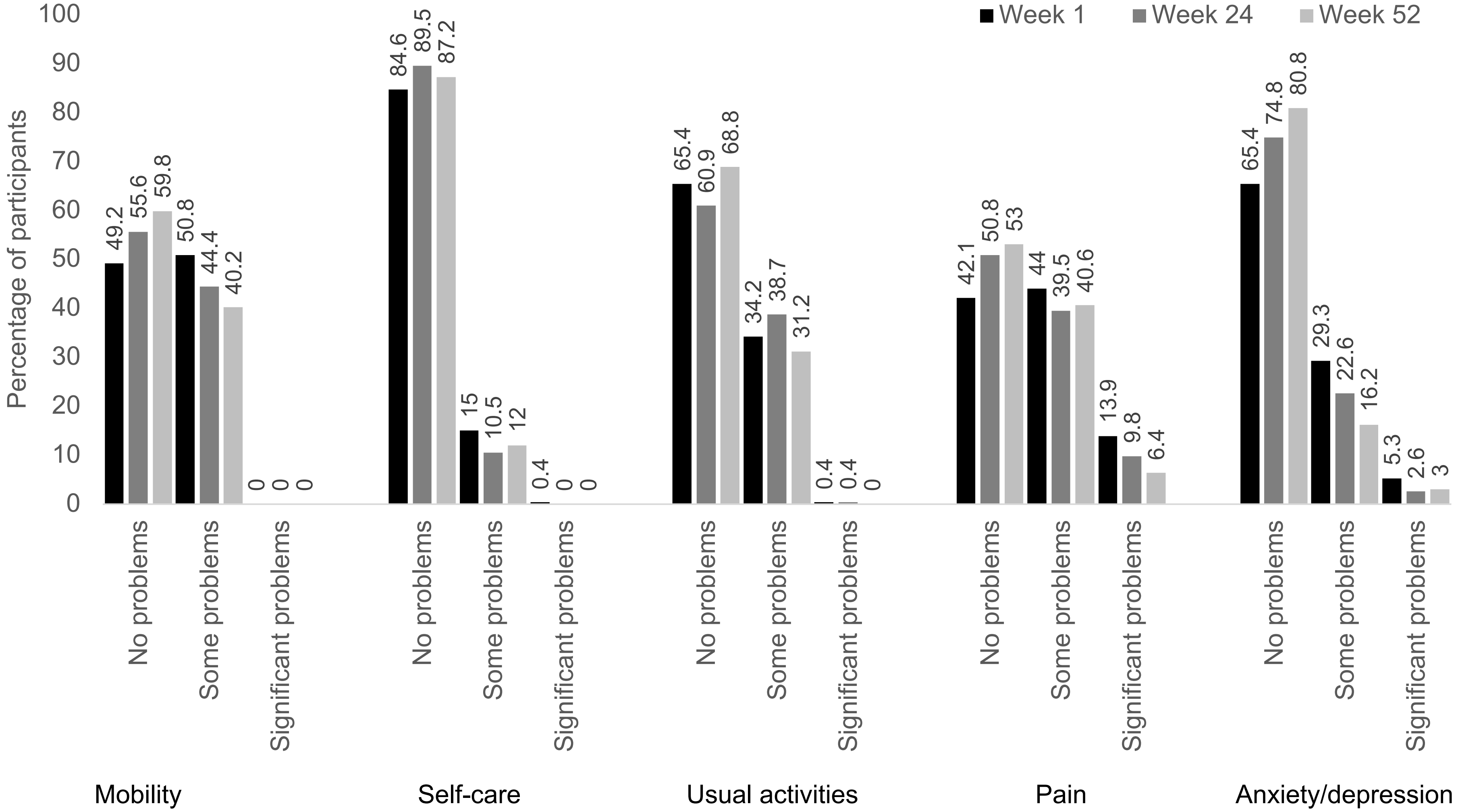


Figure 6.5. Percentage of participants reporting no/some/significant problems with each EQ-5D-3L health profile state at weeks 1, 24 and 52.

6.4.4.7. Number of referrals required

In total, 254 participants showed a long-term improvement of any degree in at least one health indicator (excluding smoking status and alcohol consumption). Taking the conservative assumption that all those lost to follow-up made no improvement, eight referrals needed to be made to have a positive long-term impact on one participant.

6.4.4.8. Smoking status and alcohol consumption

There were no significant differences in smoking status ($\chi^2(2)=2.154$, $p > 0.05$) or frequency of alcohol consumption ($\chi^2(2)=2.122$, $p > 0.05$) between the three time points. There were, however, significant differences in the number of alcohol units consumed per drinking episode ($\chi^2(2)=28.427$, $p \leq 0.001$), with a shift in the level of consumption, from a higher number (≥5 units) to a lower number (≤4 units) of units consumed per drinking episode. The changes between weeks 1 and 24 ($z=-4.145$, $p \leq 0.001$), and weeks 1 and 52 ($z=-4.486$, $p \leq 0.001$) were significant.

6.4.4.9. Long-term goals

Data on the achievement of long-term goals were available for 269 out of 273 participants. The most frequently reported long-term goals are reported below in Table 6.21. Of 269 participants, 192 (71.4%) reported that they had met their long-term goal. Due to the method of reporting achievement of goals by the ERS, it was not possible to determine the percentage of participants that had achieved each individual goal. Rather, 71.4% of participants had achieved a cluster of goals, predominantly comprised of those reported in Table 6.21.

Table 6.21. Most frequently reported long-term goals (n=273).

Goal	%	n
Reduce BMI	49.8	136
Increase or maintain PA level	26.7	73
Improve aspects of physical health (e.g. BP)	16.1	44
Improve physical fitness	11	30
Improve mobility	9.5	26
Improve physical strength	8.1	22
Reduce pain levels	6.2	17
Improve limitation free walking distance (i.e. claudication)	3.7	10

% and n refers to the percentage and number of particpants reporting each goal.

6.4.5. Summary of findings – study 2

- There were significant increases in PA in the medium (+594 MET mins) and long term (+636 MET mins). When considering long-term changes in PA, there was a 55.4% significant increase in the proportion of participants who were active above the 90-minutes per week threshold. For one participant to become long-term active (≥90 minutes per week), 11 referrals needed to be made.
- There were significant improvements, in the medium and long term, in BMI, systolic BP, HRQoL and mental wellbeing (WEMWBS). To have a positive long-term health impact on one participant, eight referrals needed to be made.
- There were no changes in smoking status, or alcohol consumption frequency. However, there was a significant decrease in the number of alcohol units consumed per drinking episode.
- Over two thirds of participants met their medium-term goal (72.9%), and 71.4% met their long-term goal. The most frequent long-term goals (of those who attended week 52) were to: reduce or maintain BMI (49.8%); increase or maintain PA levels (26.7%), and; improve aspects of physical health (16.1%).
- Sub-group analyses demonstrated participation to be least effective for those referred for respiratory related, or other conditions, and most effective for

those referred for musculoskeletal conditions.

6.5. Study 2 - Discussion

6.5.1. Main findings of study 2

The objectives of this study were to evaluate the medium to long-term impact of ERS participation on PA level and health-related outcomes, including a sub-group analysis based on primary referral condition. These objectives contribute new knowledge, by reporting the long-term impact of participation, and by acknowledging the heterogeneity of this population with a sub-group analysis, as was recommended by previous research (NICE, 2014a; Pavey et al., 2011; Rowley et al., 2018).

The most important findings of this study were that participation led to significant medium to long-term improvements in self-reported PA level and the majority of health-related outcomes. For every 11 participants referred, one participant went on to report long-term PA behaviour change (≥90 mins per week), and for every eight participants referred, one participant went on to show long-term improvements in at least one health indicator. Sub-group analyses indicated that participation was most effective for those referred for musculoskeletal conditions.

These findings are important because previous systematic reviews that are based on a small number of RCTs, but inform policy, have found no consistent evidence for an improvement in PA and health outcomes (Campbell et al., 2015; Pavey et al., 2011). Currently, there is an understanding that ERSs have a small impact on short-term PA changes, with the long-term impact largely unknown (Campbell et al., 2015). The results of this study indicate that this scheme was effective at initiating and sustaining long-term PA behaviour change, and improved health outcomes.

6.5.2. Change in physical activity

The results of this study are in contrast with the conclusions of systematic reviews, and indicate that this ERS was successful at improving long-term PA behaviour change (Campbell et al., 2015; Pavey et al., 2011). The significant increase in long-term PA (+636 MET mins per week, or 180 mins per week), was much higher than expected (average +55 mins per week, Campbell et al., 2015). There was also a significant increase in the proportion of participants achieving ≥90 minutes of moderate intensity PA per week (from 5.5% at week 1 to 40.2% at week 52). This was within the range previously reported (25-68% at week 52, Campbell et al., 2015).

However, applying the ≥90-150 minutes threshold in evaluations disregards individuals who became active for less than 90 minutes per week. This is an important limitation, as significant improvements in mortality risk are achieved when moving from being physically inactive to achieving a low volume of weekly PA (Arem et al., 2015; Hupin et al., 2015; Wen et al., 2011). In this study, at week 52, whilst 40.2% were active at moderate-vigorous intensity for ≥90-150 minutes per week, a further 17.3% were active below this threshold. A further 42.4% were completely inactive at this intensity. When considering total PA, and therefore the contribution of low intensity PA, 72.7% were active for ≥90-150 minutes per week, 10.3% were active below this threshold, and only 17% were completely inactive. These distinctions are important to report, as in previous studies, those active for less than 90 minutes per week would have been described as inactive, despite the benefits associated with this volume of weekly PA. This is also an important limitation of cost-effectiveness analyses, which determine success by meeting the ≥90-150 minute threshold (Anokye et al., 2011; Trueman & Anokye, 2013), therefore failing to capturing economic benefits gained from PA below this level (NICE, 2014a).

Since the latest systematic reviews (Campbell et al., 2015; Pavey et al., 2011), several published studies have reported significant increases in self-reported PA in the short (8-12 weeks) (Stewart et al., 2017; Webb et al., 2016), medium (24 weeks)

(McGeechan et al., 2018) and long term (>12 months) (Martín-Borràs et al., 2018). Furthermore, the studies of McGeechan et al. (2018) and Martín-Borràs et al. (2018) demonstrated that increases in PA achieved during the scheme were sustained beyond its completion, similar to the results of this scheme. Two of these studies measured PA using the IPAQ, and the long-term change in total PA per week in this study (+636 MET mins) was higher than that reported by Webb et al. (2016) after 8 weeks (+296 MET mins), but lower than that reported by Martín-Borràs et al. (2018) (+1373 MET mins) after 15 months. Whether or not the increases in PA observed in this study are truly long-term changes (i.e. after several years) should be revisited in future research.

The self-reported increases in PA observed in this study, and recent RCT and observational studies (e.g. Martín-Borràs et al., 2018; McGeechan et al., 2018), supports the value of ERSs for increasing levels of PA in indivuals with health conditions. In these cases, policy actions which recommend partnerships between health care services and the sports and leisure sector, to provide routes of referral to PA interventions, have been successful (e.g. Department for Culture, Media & Sport, 2015; PHE, 2014a; WHO, 2018).

6.5.3. Change in physical health outcomes

When considering physical health outcomes, systematic reviews have reported no significant difference in body composition or resting BP, compared to comparison groups, at short and long-term follow-ups, with any improvement in the ERS group mirrored by an improvement in the comparison group (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). This may be due to control group contamination whereby, for example, individuals in the comparison group are motivated to become physically active, thereby achieving physical health benefits (Campbell et al., 2015).

In contrast, the results of this study indicate that the scheme was effective at improving physical health outcomes, finding significant medium-term improvements in BMI, systolic BP and diastolic BP, but only long-term improvements in BMI and systolic BP. As long-term changes in PA are required to sustain long-term changes in health outcomes (Martín-Borràs et al., 2018), this may suggest that self-reported levels of PA of individuals in this study may have been close to as reported, or maintained at a sufficiently high level to sustain improvements in physical health outcomes.

When interpreting the impact of participation on change in BMI, Stewart et al. (2017) noted that individuals referred to an ERS because they were overweight were more likely to receive PA prescriptions designed to improve body composition, and therefore greater declines in BMI were to be expected. Furthermore, McGeechan et al. (2018) identified differences in change in BMI when participants were grouped according to their primary referral condition, but could not determine the direction of the relationship. Parretti et al. (2017) also suggested that overweight and obese participants may improve BMI to a greater extent than normal weight participants, which was observed in this study when stratifying the results according to BMI class. Sub-group differences, in relation to factors such as existing obesity, or the objective of the PA prescription, may explain why some studies find no significant improvements in BMI when it is aggregated across all participants.

Furthermore, when interpreting the impact of participation on change in BP, it is important to note that BP can be variable and influenced by several external factors such as temperature, smoking and emotion (Mancia, 2012), as previously discussed in section 3.5.6. Participants may also begin BP treatments during participation, and as noted by Ward et al. (2010), improvements in BP from participation in an ERS may be limited in magnitude by medication prescribed to bring these outcomes to an optimal level. BP may also be reduced in greater magnitude in those with existing hypertension, for example, Cornelissen and Smart (2013) observed greater PA induced changes in BP in those with hypertension compared to those without. It is also known that different types of PA, such as aerobic, resistance or combined PA, can cause different magnitudes of improvement in BP (Cornelissen & Smart, 2013), which is an important consideration given that ERS participants are not given standardised PA prescriptions, and may participate in different types of activities. Therefore, the interpretation of the impact of ERS participation at the population level, on physical outcomes such as BP, is difficult as change in BP can be affected by various factors.

6.5.4. Change in psychological health and health-related quality of life

Whilst some studies have reported significant improvements in psychological health outcomes (Gusi et al., 2008; Lord & Green, 1995; Murphy et al., 2012; Stewart et al., 2017) and HRQoL (Gusi et al., 2008; Pardo et al., 2014; Stewart et al., 2017; Ward et al., 2010), systematic reviews have concluded no overall significant impact of ERS interventions on psychological health or HRQoL, with the exception of a weak short-term reduction in depression level (Campbell et al., 2015; Pavey et al., 2011).

This study found significant medium and long-term improvements in HRQoL and mental wellbeing, and therefore the findings support previous studies that have found similar outcomes (e.g. Gusi et al., 2008; Stewart et al., 2017). It would be beneficial for future research to determine which elements contribute to improved mental wellbeing, and whether this is due to increased PA, or other elements of ERSs such as social interaction and social support (Murphy et al., 2012).

6.5.5. Changes in smoking status and alcohol consumption

Few published evaluations have explored the impact of ERS participation on smoking and alcohol behaviours. A small number of studies have reported descriptive changes in smoking status (Isaacs et al., 2007; Lord & Green, 1995; Taylor et al., 1998; Ward et al., 2010), such as a 3.1% decline in the number of participants reporting to be smokers (Isaacs et al., 2007), or that one smoker became a non-smoker (Ward et al., 2010). None of these studies provided statistical significance of change, and therefore the findings of this study are presumed to have a similar, non-significant impact on smoking status. The ERS evaluated in this study had a protocol to signpost all smokers to a smoking cessation service, but anecdotal observations communicated by ERS staff suggested low uptake of this referral. Fewer studies have reported alcohol consumptions levels, with only Lord and Green (1995) reporting that nine participants modified their alcohol consumption. It is therefore difficult to compare these findings with others, or to understand if other ERSs have positively impacted lifestyle behaviours. Although there was a significant impact of participation on units of alcohol consumption (per drinking episode), there was no significant impact on frequency of alcohol consumption or smoking status. There is mixed evidence regarding the impact of PA interventions on smoking status (Ussher, Taylor, & Faulkner, 2014) and alcohol consumption (Hallgren, Vancampfort, Giesen, Lundin, & Stubbs, 2017), and the findings of this study are not surprising, given that these behaviours were not the focus of consultations, with little more than signposting to recommended guidelines or support services.

6.5.6. Differences according to primary referral condition

The findings of this study support arguments for recognition of population heterogeneity when analysing outcomes (NICE, 2014a; Pavey et al., 2011; Rowley et al., 2018). This study focused on differences in outcomes according to primary referral condition. It is possible that differences in outcomes may also exist between other groups such as age and SES. However, this study focused on primary referral condition to build upon emerging findings that outcomes may differ between these groups (Murphy et al., 2012; Rowley et al., 2018), and to fulfil recommendations for sub-group analyses based on health condition (Anokye et al., 2011; Pavey et al., 2011; Rowley et al., 2018). Furthermore, the eligibility criteria for this ERS, regarding health conditions was more diverse than previous studies (e.g. Murphy et al., 2012), whilst the demographics (e.g. age, SES) of the population were more homogenous. The majority of referrals to this ERS, for example, were for individuals from areas of low SES; 81% of which were from the three most deprived quintiles of Tameside (see earlier section 5.4.2).

Although the reasons remain unknown, in terms of the outcomes analysed in this study, the results indicated that participation was most effective for those referred for musculoskeletal conditions, who were the only group to experience significant changes in all outcomes, despite having the lowest median increase in self-reported PA. Rowley et al. (2018) recently noted that evidence for the impact of ERS participation for participants with musculoskeletal conditions was limited. The results therefore contribute new knowledge in this area. They are also important considering the current context of musculoskeletal conditions in the UK, which affect more than 17.8 million adults (Arthritis Research UK, 2018), and are the nation's leading cause of years lived with disability (James et al., 2018). Musculoskeletal conditions can affect all aspects of an individual's life (Arthritis Research UK, 2018). However, they also have a substantial indirect burden on society; in 2017, almost 30% of all work days lost to sickness absence were attributed to musculoskeletal conditions (ONS, 2018b). The economic burden of these conditions to the nation is therefore high, and estimated to be £103–129 billion per year (Arthritis UK Research, 2018). As a result, there is an imperative to reduce the impact of musculoskeletal conditions in the UK at this time, to both individuals and wider society (Arthritis UK Research, 2018), and the results of this study indicate that ERSs can contribute to achieving this aim. Consequently, there is scope for the targeted referral of individuals with musculoskeletal conditions to this ERS.

The sub-group analysis also highlighted some important findings, for example those referred specifically for being overweight did not significantly reduce BMI, and those referred for existing CVD or being at high risk of CVD did not significantly improve BP. However, those referred for an existing mental health condition had the largest median increase in mental wellbeing (+12 points) and HRQoL, which is promising given recent mandates to urgently address outcomes in this population (Department of Health & Social Care, 2018a; NHS, 2016). This analysis has provided information on how the different groups benefit from participation, and whether they improved outcomes relevant to their primary referral condition, or whether potential improvements in practice are required. However, it is important to note the limitations of this analysis. Whilst participants were grouped according to their primary referral condition, the majority had multiple comorbidities, and it is possible that one or more of these comorbidities had an impact on the outcomes rather than the primary referral condition itself. Despite this, considering that decisions to commission and continue ERSs are made from evaluations such as this one, it is important for future evaluations to acknowledge this heterogeneity, and to treat the population as a diverse group and identify effectiveness between groups, rather than solely reporting population outcomes (Anokye et al., 2011; NICE, 2014a; Pavey et al., 2011; Rowley et al., 2018). This can be strengthened by the inclusion of recommendations to conduct sub-group analyses in evaluation frameworks and ERS guidelines.

6.5.7. Strengths and limitations of study 2

The primary strengths of this study are that it is an evaluation of the outcomes of a real-life ERS, which provides a higher level of ecological validity (Gidlow et al., 2008).

The study also had a relatively large sample size, with a longitudinal follow-up.

There are, however, a number of limitations in study design and analysis. As this study did not have a control group, it cannot be said with certainty that the results of this study were due to participation in the ERS. The role of RCTs in the evaluation of public health interventions operating in real world contexts has received criticism (Rutter 2017, Pettman et al., 2012), as previously discussed in section 2.5. Recent evaluations, for example, did not include a comparison group because the ERS already existed before the evaluation started (McGeechan et al., 2018) or because of resource constraints (Stewart et al., 2017). Similarly, this study did not include a control or comparison group because the scheme was pre-existing.

One of the most important limitations of this study is the loss to follow-up of 30.5% of the sample at week 52. It is not possible to quantify the bias that this introduces, although it is noted that, other than being younger, those lost to follow-up were not systematically different at baseline. It is possible that those who were not followed up at week 52 were less likely to have maintained their participation in PA.

Another important limitation is the self-reporting of PA and sedentary behaviour, as questionnaires are vulnerable to bias and poor validity (Haskell, 2012; Helmerhorst et al., 2012; Strath et al., 2013), with the potential for under or over-reporting of behaviours (Cerin et al., 2016; Chastin et al., 2018; Craig et al., 2003; Steene-Johannessen et al., 2016). Despite this, few studies have assessed PA levels in ERS participants using device-based methods (Gallegos-Carrillo et al., 2017), perhaps due to the convenience and low cost of questionnaires (Strath et al., 2013). Therefore, the current consensus on the effectiveness of ERSs at increasing PA levels is based predominantly on self-reported data (Campbell et al., 2015; Pavey et al., 2011), which may under or over-estimate the impact of the intervention. Future evaluations would therefore benefit from the inclusion of device-based measurement of PA to support self-reported outcomes (Campbell et al., 2015).

Furthermore, although considered to be objective measures, BMI and BP are susceptible to variation, as previously discussed in section 3.5. Body weight, for example, was not measured at the same time of day, and therefore may have varied due to food and water intake/excretion (Madden & Smith, 2016). BP is also variable, influenced by behavioural and environmental factors such as temperature, emotion, recent PA, and anxiety towards having BP measured (Mancia, 2012). It is also limited by technical and human errors, such as infrequent calibration of automatic BP monitors (O’Brien et al., 2003; Pickering et al., 2005). Calibration of the monitors in this study could not be verified. If variation in these measures occurred, then the results may not accurately reflect the extent to which the outcomes changed over time. However, this is likely to introduce random error to the results rather than systematic bias (Blackmore, Medina, Ravenel, Silvestri, & Applegate, 2011).

In the general field of PA and health outcomes research, it is now well known that different types of PA may affect outcomes differently. BP, for example, is significantly reduced by participation in aerobic or resistance exercise, but not combined exercise (Cornelissen & Smart, 2013). Maximal oxygen consumption, may also be increased to a greater extent by interval training compared to continuous aerobic exercise (Milanović, Sporiš, & Weston, 2015). In an intervention such as an ERS, where individuals receive individualised PA prescriptions, the population may be participating in a wide range of activity types, at different intensities, for different lengths of time, and at different frequencies. It is therefore important to acknowledge the contribution of non-standardised PA prescriptions to population outcomes such as BMI and BP (Stewart et al., 2017), where there is likely to be variation in magnitude of improvement according to the PA prescription. Rowley et al. (2018) have recommended that future research compares the impact of different PA prescriptions on outcomes, and whether particular PA prescriptions are beneficial for those with certain health conditions.

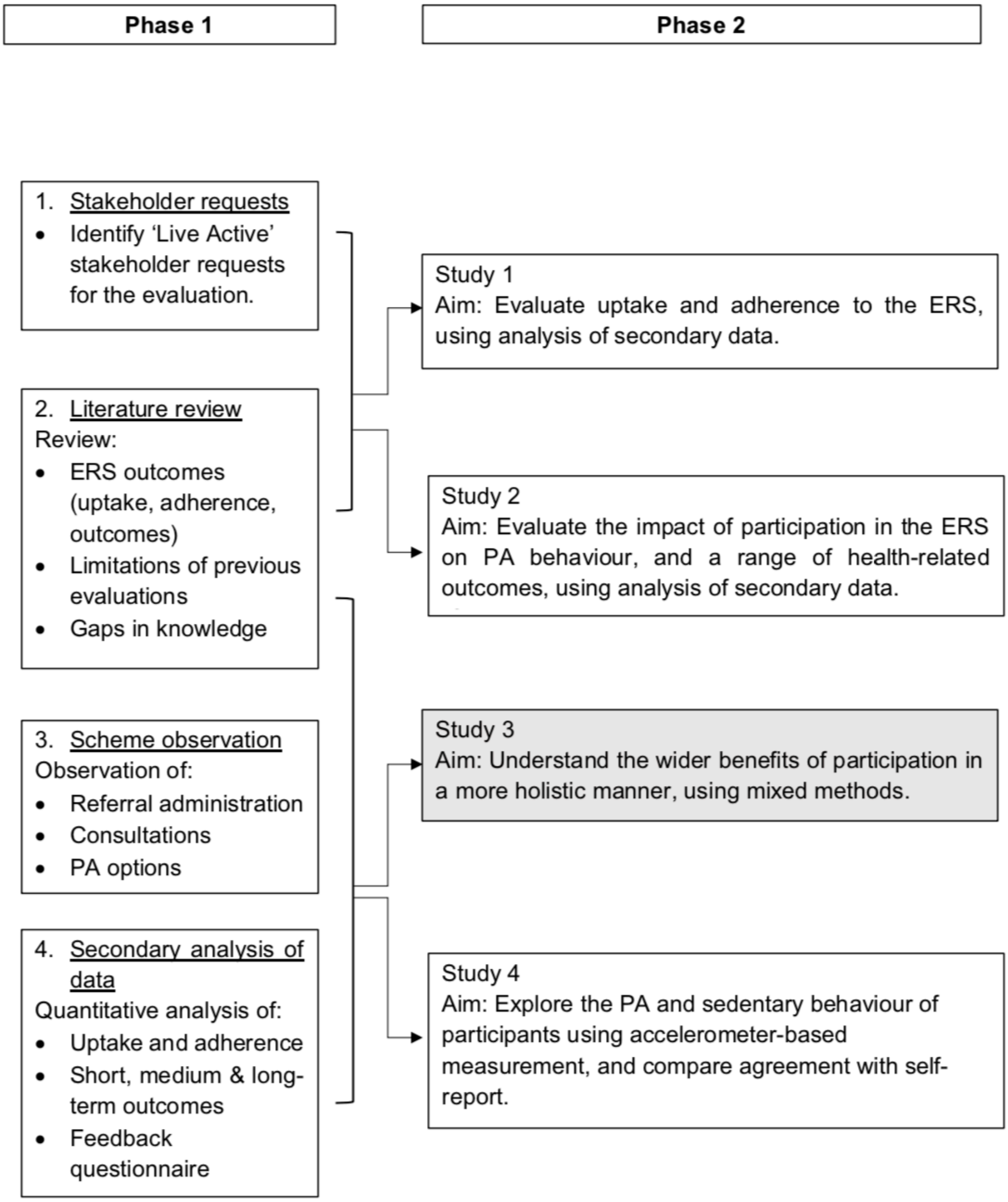
A further limitation of this study was missing data, in particular compliance with BP measurement and collection of the WEMWBS questionnaire. Where data were missing, the reasons were unknown to the researcher. It is known that several participants were not native English speakers, or had learning difficulties or disabilities that may have affected the ability to answer questionnaires such as the WEMWBS. It is also relevant to note the time associated with collection of the routine outcomes during a consultation, some of which were questions of a sensitive nature. It is therefore possible that time was a barrier to the full collection of data. In future, scheme developers may wish to consider both the outcomes required to demonstrate effectiveness, but also the feasibility and experience of the data collection process for both the exercise professional and participant, giving the priority of consultations to encouraging behaviour change.

In the discussion, the outcomes of this study have been compared to the findings of systematic reviews and subsequent studies. However, these systematic reviews, and the studies they included, are not without limitations, which may explain the difference in findings between this study and the conclusions of previous reviews. Reviews have noted, for example, the heterogeneous nature of ERS intervention delivery, and the outcomes collected, which can contribute to inconsistent findings (Campbell et al., 2015; Pavey et al., 2011). Studies included in these reviews had variable but often short scheme durations (10-12 weeks), with variable follow-up durations (2-12 months). The participants were mostly middle aged, referred by a GP, and had no medical diagnoses, except for CVD risk factors such as obesity and raised BP. Most schemes had standardised PA prescriptions, usually twice weekly attendance at a 30-60 minute PA class. The characteristics of the included studies are therefore in contrast with the characteristics of the ‘Live Active’ scheme, which was longer in duration (6-months), reported a long-term follow-up (12-months), accepted referrals from a variety of health care professionals for participants who had pre-existing medical diagnoses, and prescribed individually tailored PA programmes utilising a range of PA types. Difference in findings between this study and others may therefore be due to the characteristics of the scheme. Longer duration schemes (20+ weeks), for example, have been associated with superior outcomes, and tailored PA prescriptions are also expected to produce greater improvements in outcomes (Rowley et al., 2018).

Lastly, this study has evaluated the impact of participation in an ERS based on a small set of pre-determined quantitative outcomes, similar to the predominance of previous research (Campbell et al., 2015; Pavey et al., 2011). However, the aim of an ERS is to increase an individual’s PA levels in order to improve their health (NICE, 2014a), and in this context health may include a wide range of potential outcomes not limited to those collected. It would be beneficial for future research to explore the full spectrum of potential benefits to participation in an ERS, potentially using mixed methods research. This could, for example, include the collection of condition specific health markers using methods such as biochemistry (e.g. Webb et al., 2016). The wider physical, mental and social benefits of participation could also be explored through qualitative inquiry, with common benefits translated in to quantitative measures for inclusion in future research, to provide a more representative evaluation of impact (Mills et al., 2012).

6.5.8. Conclusions of study 2

Referral to the ‘Live Active’ ERS led to significant improvements in PA and a range of health-related outcomes, which were sustained in the long term beyond the scheme’s duration. This indicates that this scheme was effective at initiating and sustaining long-term PA behaviour change, and subsequently health outcomes. The sub-group analysis in this study has directed areas for improvement in practice, to improve outcomes across referral groups. It has also highlighted the value of referral to this ERS for groups in which there is a current imperative to improve health-related outcomes, such as those with musculoskeletal conditions (Arthritis Research UK, 2018) or mental health issues (Department of Health & Social Care, 2018a). However, consideration of the results of this study has highlighted the myriad of limitations that can be encountered when evaluating an ERS. Moving forwards, the evidence base requires more long-term evaluations, which examine effectiveness between groups, acknowledges the impact of non-standardised PA prescriptions, includes device-based measurement of PA to support self-reported outcomes, and explores the wider benefits of participation using qualitative methods.



Phase 1

1. Stakeholder requests
- Identify 'Live Active' stakeholder requests for the evaluation.

2. Literature review

Review:
- ERS outcomes (uptake, adherence, outcomes)
- Limitations of previous evaluations
- Gaps in knowledge

3. Scheme observation

Observation of:
- Referral administration
- Consultations
- PA options

4. Secondary analysis of data

Quantitative analysis of:
- Uptake and adherence
- Short, medium & long-term outcomes
- Feedback questionnaire

Phase 2

Study 1
Aim: Evaluate uptake and adherence to the ERS, using analysis of secondary data.

Study 2
Aim: Evaluate the impact of participation in the ERS on PA behaviour, and a range of health-related outcomes, using analysis of secondary data.

Study 3
Aim: Understand the wider benefits of participation in a more holistic manner, using mixed methods.

Study 4
Aim: Explore the PA and sedentary behaviour of participants using accelerometer-based measurement, and compare agreement with self-report.

Overview of the research framework (study 3)

Chapter 7 Study 3: Mixed methods evaluation of the perceived benefits of participation in the 'Live Active' exercise referral scheme

7.1. Study 3 – Introduction, aims and objectives

In a field dominated by positivist research methods that some researchers argue do not capture the full impact of an intervention, this study answers calls for more pluralistic evaluations of ERSs, which attempt to capture the full impact of the intervention. This chapter begins by critically discussing the literature on the benefits of participation uncovered by quantitative research, compared to that collected through qualitative studies, to highlight the broad range of benefits reported from participants. Particular consideration is given to the adjunct benefits of ERS participation, which are rarely captured in quantitative evaluations. The introduction concludes by considering the importance of pluralistic methods and recognition of adjunct benefits when conducting an evaluation, and introduces the findings of secondary data analyses that support the rationale for conducting this study.

Study 3 aim: Understand the wider benefits of participation in a more holistic manner, using mixed methods.

Study 3 objectives:

3.1. Evaluate the medium to long-term (6-12 months) benefits of participation in the 'Live Active' ERS, using semi-structured interviews and secondary analysis of the 'Live Active' feedback questionnaire.

3.2. Identify if these emerging benefits are outcomes measured by the 'Live Active' ERS, previous ERS evaluations, or ERS evaluation frameworks.

7.2. Literature review of the benefits of ERS participation

In 2005, Dugdill et al. noted that the field of ERS evaluation had been dominated by positivist research methods that predominantly captured the impact of participation on physiological markers of health such as body composition and BP. The authors argued that positivist methods such as RCTs could not embrace the complexity of real world interventions such as ERSs, nor could they capture the holistic worth of schemes on other important but often overlooked parameters such as social and psychological health. These sentiments have been echoed by others (Crone et al., 2005; Mills et al., 2012; Riddoch et al., 1998), with recognition that pluralistic methods may provide a more holistic evaluation of a scheme's impact (Crone et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998).

To-date, developers of schemes, researchers and other stakeholders have not been prompted by evaluation frameworks to consider the alternative benefits of participation. Previous evaluation guidelines have been ambiguous (DoH, 2001), with no strong evaluation framework available until more recently (The SEF – Cavill et al., 2012). Whilst this new framework (Cavill et al., 2012) discussed the importance of understanding the objectives of an intervention, in order determine the outcomes that should be measured, the suggested outcomes continued the focus on the collection and evaluation of PA and physical health-related outcomes, despite several calls for greater recognition of other potential adjunct benefits (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; Mills et al., 2012).

It has been suggested that interventions targeting behaviours such as PA are likely to have unintentional ripple effects on multiple outcomes beyond their specific target, spanning physical, mental, and social health (Smith & Petticrew, 2010; Wilson, 2015). Therefore, the effectiveness of an intervention and its return on investment may be easily underestimated if there is no attempt to capture these impacts (Wilson, 2015). Several evaluations leaning towards interpretivist approaches have now used qualitative or mixed research methods to understand participant's experiences of ERSs (Mills et al., 2012; Queen et al., 2015). However, more than a decade after calls for more pluralistic research (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998), the field of ERS evaluation is still dominated by positivist research methods that predominantly focus on physiological outcomes (see for e.g. Campbell et al., 2015). In the following sections, the evidence for the impact of ERS participation, explored through both quantitative and qualitative approaches, will be summarised.

7.2.1. Physical health outcomes

Historically, much of the focus of ERS evaluations has been placed on changes to PA and physical health. However, quantitative measures of physical health have rarely gone beyond that of BMI and BP. Outcomes beyond these have often been reported qualitatively (e.g. Fenton, Hill, Stocker, & House, 2015; Queen et al., 2015), and more recently quantitatively. Measures have been collected, albeit infrequently, for important physical health risk factors such as: physical fitness (Isaacs et al., 2007; Sørensen et al., 2008); lung function (Isaacs et al., 2007); HbA1c, glucose or insulin (Sørensen et al., 2008; Webb et al., 2016), and; cholesterol and triglycerides (Isaacs et al., 2007; Webb et al., 2016), as discussed previously in chapter 6. When collected, changes in these outcomes have been used to infer the clinical effectiveness of an ERS (Isaacs et al., 2007; Sørensen et al., 2008), such as impact on CVD risk lowering (Webb et al., 2016).

When qualitatively exploring the physical health outcomes of ERSs, the range of benefits that are reported, such as reducing BMI and improving physical fitness, and the importance that participants place upon these benefits, varies between respondents, and is perhaps dependent upon personal goals and circumstances (Wormald et al., 2006). Improved physical health status appears to be of importance to participants (Morgan et al., 2016), but anticipated outcomes go beyond the limited range of outcomes traditionally measured, such as body weight, BMI, and BP.

Weight loss, for example, is a commonly cited goal and benefit of participation (Campbell et al., 2015; Morgan et al., 2016), however there is also a strong focus on expected benefits to mobility, strength, flexibility, and physical fitness, often in order to improve activities of daily living (Morgan et al., 2016; Scott, Breckon, Copeland, & Hutchison, 2015; Wormald & Ingle, 2004; Wormald et al., 2006). Some participants recall how PA has benefitted them specifically in relation to their health conditions, such as improvements in joint stiffness associated with arthritis (Stathi et al., 2004; Wormald & Ingle, 2004). They also recall expectations that improvements in physical health outcomes would reduce the risk of further ill health (Morgan et al., 2016; Scott et al., 2015). This is a relevant expectation, given that ERSs are intended to improve the health status of those who have or are at risk of ill health (NICE, 2014a). However, the risk of future health events, or their risk factors, many of which are central to the outcome expectations of participants, have rarely been measured, as discussed above (Isaacs et al., 2007; Sørensen et al., 2008; Webb et al., 2016).

7.2.2. Mental health outcomes

Compared to physical health outcomes, mental health outcomes have been explored in published evaluations to a lesser extent. Systematic reviews have concluded no overall significant impact of ERS interventions on psychological health, with the exception of a weak short-term reduction in depression level (Campbell et al., 2015; Pavey et al., 2011).

Changes in mental health have also been explored frequently through qualitative research. Participants in several studies, for example, have reported improved mood (Fenton et al., 2015; Stathi et al., 2004), stress reduction (Stathi et al., 2004), improved self-confidence (Wormald & Ingle, 2004; Wormald et al., 2006) and better wellbeing (Stathi et al., 2004; Wormald & Ingle, 2004) from ERS participation. Despite improved mental health being frequently reported throughout qualitative research, it is still not regarded as an important indicator of success by stakeholders, and therefore has not been routinely included in ERS evaluations to date (Crone et al., 2005; Morgan et al., 2016).

7.2.3. Adjunct benefits

Qualitative research in relation to participants' experiences of an ERS highlights a range of adjunct benefits to participation that are under-addressed by previous evaluations and frameworks. Some of these benefits are important for the process of behaviour change and maintenance such as: new knowledge and awareness of the importance of PA for health, and how to incorporate PA in to daily life (Fenton et al., 2015; Wormald & Ingle, 2004; Wormald et al., 2006); overcoming preconceived ideas around PA (Wormald et al., 2006), and age-related beliefs (Fenton et al., 2015); developing the self-confidence to be physically active (Wormald & Ingle, 2004), and; having a new sense of responsibility to change behaviour (Fenton et al., 2015) and improve oneself (Stathi et al., 2004).

Qualitative research, and in fewer instances quantitative research, has shown positive changes to other lifestyle behaviours such as diet (Fenton et al., 2015; Isaacs et al., 2007; Wormald et al., 2006), alcohol consumption and smoking status (Lord & Green, 1995), although there was no change in smoking behaviour in the later study of Taylor et al. (1998). Furthermore, in some cases family and friends of the participants were also motivated to positively modify lifestyle behaviours such as diet and PA (Isaacs et al., 2007; Wormald et al., 2006). Other adjunct benefits include a reduction in medication (Queen et al., 2015) and healthcare utilisation (Isaacs et al., 2007; Queen et al., 2015), benefits that are potentially cost saving to healthcare systems and wider society.

A frequent finding is that ERSs provide a renewed sense of purpose and a break away from usual daily routines, especially for older age groups in which participants felt a loss of routine and purpose with retirement (Hardcastle & Taylor, 2001; Scott et al., 2015; Stathi et al., 2004; Wormald & Ingle, 2004; Wormald et al., 2006). There is also a strong theme around social relationships, with participants benefitting from peer support and a sense of community within schemes (Fenton et al., 2015), and establishing new social relationships that help to prevent or reduce social isolation (Fenton et al., 2015; Hardcastle & Taylor, 2001; Stathi et al., 2004). Whilst social benefits have not been routinely measured or acknowledged by previous ERS evaluations, previous literature has shown that they have significant benefits to health, and economical expenditure. In recent years, for example, increasing attention has focused on the contribution of social isolation to poor mental and physical health, leading to premature mortality (Holt-Lunstad, Smith, & Layton, 2010; Holt-Lunstad, Smith, Baker, Harris, & Stephenson, 2015; Mushtaq, Shoin, Shah, & Mushtaq, 2014; Valtorta & Hanratty, 2012). This decline in health contributes to higher rates of health care utilisation amongst socially isolated individuals, and therefore higher economical expenditure (Gerst-Emerson & Jayawardhana, 2015; Molloy, McGee, O'Neill, & Conroy, 2010; Valtorta & Hanratty, 2012).

7.2.4. Considerations for the evaluation of ERSs

The overall aim of ERSs, according to NICE (2014a, p. 7) is "to increase someone's physical activity levels on the basis that physical activity has a range of positive health benefits". In this context, no definition of health is provided, and in the wider literature its definition is a contentious issue. Health was defined in the Constitution of the WHO in 1948 (2006, p. 1) as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". This definition has received criticism surrounding the concept of complete wellbeing, and the definition's relevance given changes in global demographics and patterns of illness since its conception (Huber, 2010; Huber et al., 2011; Jadad & O'Grady, 2008; Misselbrook, 2014). The argument is well concluded by Jadad & O'Grady (2008, p. a2900), "any attempt to define health is futile; that health, like beauty, is in the eye of the beholder; and that a definition cannot capture its complexity". Despite growing uncertainty around how health can be defined, the original definition (WHO, 2006) is important in the context of an ERS evaluation, in considering that the concept of health is holistic and encompasses more than just characteristics of physical health.

Despite various definitions of health, the ERS evidence base continues to be predominated by positivist evaluations which reinforce the measurement of changes in PA and physical health-related markers such as BMI (Campbell et al., 2015; Pavey et al., 2011). If these variables are not improved then an ERS is considered to be ineffective, when in fact participants may have experienced other benefits not captured by the evaluation, which could represent effectiveness (Wilson, 2015). This has been demonstrated in section 7.2.3. through the discussion of the adjunct benefits reported in previous literature. Therefore, positivist evaluations of past and future evaluations may underestimate the effectiveness of an ERS and its return on investment if there is no attempt to capture and report these alternative benefits (Wilson, 2015).

It is apparent that when reviewing the qualitative experiences of participants, although published to a lesser extent than quantitative evaluations, a wide range of benefits to participation can be uncovered that collectively demonstrate a holistic benefit of ERSs (see for e.g. Mills et al., 2012; Queen et al., 2015; Wormald et al., 2006), and the ripple effect of behavioural interventions (Wilson, 2015). The variation in benefits gained, as described throughout section 7.2, may reflect the diverse populations that ERSs engage with, in which many of the participants may have a range of not just health, but complex emotional and social problems, and therefore have many potential ways of benefitting from participation (Wormald et al., 2006).

In order to make more representative evaluations of the impact of ERSs, several researchers have called for future evaluations to include appropriate methods to capture the holistic worth of schemes (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998). This includes exploring the experiences of participants through qualitative methods (Crone et al., 2005; Dugdill et al., 2005), and the inclusion of outcome measures that evaluate the impact on not just physical, but mental and social health (Dugdill et al., 2005). This has been followed by the recommendations of NICE (2014a) for future evaluations to gather further information regarding which measures and outcomes are best to establish effectiveness, recognising that some may be adjunct benefits to participation that are not routinely collected.

7.2.5. Emerging evidence from the 'Live Active' scheme

Secondary analysis of data routinely collected by the 'Live Active' ERS, between July 2015-2016, indicated that participants were frequently self-reporting alternative benefits to participation such as improved: sense of wellbeing; pain; stress; confidence; self-esteem; socialisation, and; independence (see appendix 7). These self-reported benefits, however, were reported via a 'Live Active' feedback questionnaire where tick-box options for a variety of potential outcomes, in addition to space for free text, may have encouraged the reporting of these particular benefits (Choi & Pak, 2005). However, the most frequently reported goals that participants wished to achieve also reflected that they expected to attain significantly different outcomes compared to the outcomes that this scheme, and others, have been designed to evaluate. This included typical goals such as increasing PA level and reducing BMI, but also included a range of different goals that are important for health such improving: physical fitness; mobility; risk factors such as blood glucose;

level of pain, and; mental health (see previous section 6.4.2.8 for examples).

7.2.6. Summary of gaps in knowledge

To date, systematic reviews have found no consistent evidence for an improvement in the primary outcomes of an ERS such as PA level and physical health (see previous section 6.2.) (Campbell et al., 2015; Pavey et al., 2011; Williams et al., 2007). However, for the most part, these predominantly positivistic evaluations of ERSs have not considered or sought to capture the holistic worth of ERS participation, therefore potentially not making an accurate and representative evaluation of ERSs in respect of their broader outcomes (Crone et al., 2005; Dugdill et al., 2005; Mills et al., 2012). Very few identied studies, for example, have collected quantitative data on adjunct benefits such as smoking behaviour (Lord & Green, 1995; Taylor et al., 1998), dietary improvements (Isaacs et al., 2007), or change in healthcare utilisation (Isaacs et al., 2007). Quantitative data on the adjunct benefits of participation collected in the 'Live Active' feedback questionnaire, from a large segment of the population, could therefore be considered novel.

However, for more than a decade, researchers have recommended that future evaluations take a pluralistic approach to evaluation (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998), with an emphasis on the qualitative exploration of how participants benefit from an ERS. To-date, few qualitative explorations of ERS impact have been published (e.g. Mills et al., 2012; Queen et al., 2015; Wormald et al., 2006), especially in comparison to the number of quantitative evaluations published (Campbell et al., 2015), with further studies therefore necessary to reach a consensus.

A mixed methods evaluation of an ERS can help to prevent underestimation of impact, by identifying the wider benefits of participation (Wilson, 2015). Knowledge of the wider benefits collected in this manner would help to further determine the impact and cost-effectiveness of schemes (Dugdill et al., 2005; NICE, 2014a; Riddoch et al., 1998). It could, for example, support evaluators to develop a new range of outcome indicators for inclusion in future research (Mills et al., 2012; NICE, 2014a), and facilitate the development of evaluation frameworks that guide evaluators to capture the impact of ERS participation in a more comprehensive manner. However, to-date, very few studies have explored the wider benefits of ERS participation (e.g. Mills et al., 2012; Queen et al., 2015; Wormald et al., 2006). Therefore, a concensus on the expected outcomes of ERS participation, which can be translated into quantitative outcomes, cannot yet be reached.

Secondary analysis of the feedback questionnaire routinely collected by the ‘Live Active’ ERS, conducted between July 2015-2016, indicated that participants were self-reporting a variety of adjunct benefits to participation (appendix 7). The aim of this study was therefore to build upon this knowledge, and to address gaps in the literature, by seeking to understand how participants perceive they have benefitted from participation in the ‘Live Active’ ERS, using mixed methods. In doing so, the results can be considered alongside the results of study 2, to make a balanced evaluation of the impact of this ERS, and to understand how future evaluations can judge the effectiveness of ERSs.

7.3. Study 3 Methods

This section provides a description of the methods used to address the aims and objectives detailed in section 7.1.

7.3.1. Research design

This was a mixed methods study, using a triangulation design to collect separate quantitative and qualitative data sets about the benefits of ERS participation. The quantitative data set was the existing ‘Live Active’ feedback questionnaire, completed with ‘Live Active’ staff, and the qualitative data set was semi-structured interviews with ERS participants, conducted by the researcher as part of this study. The quantitative data set was included as it could provide an indication of the frequency of several adjunct benefits from a large sample of participants. However, in order to extend the existing literature in this area, it was important to realise that a benefit is perceived and experienced by the individual and cannot be fully pre-conceived by evaluators and captured through pre-selected quantitative measures. This is particularly important when dealing with the diverse populations encountered within ERSs, where what constitutes a benefit may vary between persons, and the extent to which an outcome can be improved may be limited by the nature of an individual’s LTCs (Mills et al., 2012; Wormald et al., 2006). Therefore, it was imperative to include qualitative methods to understand participants’ experiences of the ERS.

A one-phase design was used (convergent parallel), whereby both quantitative and qualitative methods were collected and analysed separately during the same time frame, with the results then compared, contrasted and interpreted together thereafter (Creswell, 2003; Creswell & Plano Clark, 2011). Using quantitative or qualitative methods alone may not always capture data that provides a deep understanding of the research question. Using both methods can be complementary, and provide a more complete analysis (Creswell, 2003; Creswell & Plano Clark, 2011). Therefore, this design allowed for the collection of two different but complementary data sets, which provided more depth in the understanding of the adjunct benefits of participation in the ‘Live Active’ ERS (Creswell, 2003; Creswell & Plano Clark, 2011).

7.3.2. Recruitment plan: semi-structured interviews

The demographics of participants in the ‘Live Active’ scheme were wide ranging. The age range at the time of commencement of data collection was 18-90 years, with participants referred for a range of health conditions. Purposive sampling (proportionate, stratified purposeful) was therefore used to capture perspectives from a wide range of participants, and to ensure the demographic spread of the population was included, in terms of gender, age and medical condition.

Purposive sampling is a non-random strategy for ensuring that particular cases are represented within a research sample (Patton, 2002; Robinson, 2014). It is a common method used in qualitative research today, however there exists ambiguity in its definition and sampling methods (Gentles, Charles, Ploeg, & McKibbon, 2015; Gentles & Vilches, 2017). Therefore, researchers using purposive sampling must describe their reasoning and approach within the context of their research study (Gentles et al., 2015).

There are 40 different variants of purposive sampling, one of which is proportionate stratified purposeful sampling (Patton, 2015). This sampling method is used to ensure representation of groups that will have importance to the research (Henry, 1990). It assumes that based on theory or previous research, certain sub-groups may have different perspectives on a phenomenon, and therefore must be included in a sample (Mason, 2002; Trost, 1986). There are no requirements for the selection of sub-groups, other than a clear theoretical rationale for their inclusion (Robinson, 2014). When the sub-groups have been decided, a target number of participants is then decided for each group (Patton, 2002).

When using stratified purposeful sampling, to control for sampling error, the researcher must identify and justify sub-groups that are theoretically based (Robinson, 2014) and representative of the population (MacNealy, 1999). A weakness of this method however is that it may be necessary to limit the number of sub-groups depending upon the time and resources available to the researcher, as the more sub-groups that are included the more complicated and time consuming the recruitment becomes (Robinson, 2014).

In this study, the sub-groups were gender (female/male), age group (18-64 years/≥65 years of age) and referral reason (has a mental health condition/has only physical health conditions) (Table 7.1). This was to capture the demographic spread of the population, without introducing the complication of too many sub-groups (Robinson, 2014). It was also based upon previous research that has found differences in the experiences of participating in an ERS between these groups (e.g. Crone et al., 2008; Dugdill et al., 2005; Mills et al., 2012; Murphy et al., 2012; Tobi et al., 2012, 2017). Participants with mental health issues, for example, are more likely to dropout from an ERS (Tobi et al., 2017), and may value the experiences of participation rather than pre and post outcome changes (Crone et al., 2005). In comparison, participants with physical health conditions may increase their PA levels more (Murphy et al., 2012), but may also feel restricted by the symptoms of these conditions, which may influence how they perceive successful participation (Mills et al., 2012).

Table 7.1. Target recruitment for each sample sub-group.

Gender	Age	Referral reason	Target recruitment
Male	18-64	Mental health issues	3-4
Male	18-64	Physical health issues	3-4
Female	18-64	Mental health issues	3-4
Female	18-64	Physical health issues	3-4
Male	≥65	Mental health issues	3-4
Male	≥65	Physical health issues	3-4
Female	≥65	Mental health issues	3-4
Female	≥65	Physical health issues	3-4
			Total: 24-32

Eligible participants were those who had passed their sixth month of participation with the scheme, and had therefore completed the ERS. Eligible participants were identified, in discussion with the ERS exercise professionals, within the research setting in consultations or supervised activities. On any given day, for example, the scheme's calendar was reviewed with ERS staff to identify eligible participants attending consultations that day, or present in supervised activities. Identified participants were informed of the research study by the exercise professional, and if they expressed an interest in taking part, they were formally invited to participate by the researcher.

The target sample size was 25 (lower range 15, upper range 35). Guidelines for qualitative research, dependent upon methodology, have recommended anywhere between 5 and 50 interviews to be sufficient (Creswell, 1998; Morse, 1994; Russel Bernard, 2000). Bertaux (1981) recommended that for all qualitative research, 15 is a minimum accepted sample size. Previous research conducting qualitative investigation into the experiences of ERS participants have used sample sizes of between 9 and 36 participants (Eynon, O'Donnell, & Williams, 2016, n=9; Mills et al., 2012, n=17; Queen et al., 2015, n=12; Stathi et al., 2004, n=13; Wormald & Ingle, 2004, n=26; Wormald et al., 2006, n=16). However, not all of these studies pertained to the perceived benefits of participation. Some qualitatively investigated other areas such as adherence. The sample size of 25 was therefore based on recommendations, previous research, and the resources required to conduct, transcribe and analyse qualitative interviews. Participants were recruited in consideration of this target until data saturation was reached and no new major themes emerged.

7.3.3. Research protocol: semi-structured interview

The purpose of the interviews was to collect information on how participants perceived to have benefitted from participation in the 'Live Active' scheme to date. Semi-structured interviews were chosen since they can direct the interaction whilst still allowing participants the flexibility to express their opinions at a deeper level, which often could not be gained from a structured interview (Sparkes & Smith, 2014). This type of interview also allowed for probing to clarify answers, or additional unplanned questions to gain further information (Corbetta, 2003; Gray, 2004). This method aligned with the objectives of the research study, in giving participants the flexibility to express how they have benefitted from participation in their own terms, free from structured questioning that might guide a desired response for the researcher (Edwards & Holland, 2013).

Interviews took place in private rooms within the 'Live Active' leisure centres, and were recorded using a digital audio recording device. An interview guide was used (see appendix 10), containing open-ended questions based on existing literature, and the emerging findings of secondary data analysed to date. Additional probing questions were used during the interview for clarification or elaboration (Russel Bernard, 2000).

Semi-structured interviews are not without limitations. The success of the interview depends upon the extent to which the participant voices their true opinions, which can be prevented if there are barriers between participant and interviewer, preventing the full sharing of information (Denscombe, 2007; Sparkes & Smith, 2014). The researcher must also be proficient in using probing questions when necessary, whilst avoiding using leading questions that elicit answers in line with their preconceived ideas. There is also an added level of difficulty to analysing semi-structured interviews compared to structured interviews (Sparkes & Smith, 2014).

7.3.4. Research protocol: secondary analysis of data

Feedback about participants' experiences of the 'Live Active' scheme was routinely collected by the exercise professional between the 12th week and 24th week of participation. This was collected in the form of a questionnaire (see appendix 2), with categorical and open ended responses, and was stored on the ERS database. All available feedback questionnaire data were retrieved from the ERS database upon completion of the final semi-structured interview. Ethical approval was previously granted for the secondary analysis of these data (HSCR15/124).

7.3.5. Data analysis: interviews

Qualitative data from interviews were transcribed verbatim, with any identifying information anonymised (e.g. name, location). Transcriptions were managed via Microsoft Word (version 15.20), and data analysis managed via Microsoft Excel (version 15.20).

Framework analysis (FA) was used to identify, analyse and interpret themes within the interview data. FA was first developed by applied qualitative researchers working in social policy, and is a now a popular pragmatic approach to research in several fields including health sciences (Ritchie & Spencer, 1994; Ritchie, Spencer, & O'Connor, 2003; Ritchie, Lewis, McNaughton-Nicholls, & Ormston, 2013).

The advantage of FA is that it provides a systematic approach to data management and analysis, with visible stages that increase the transparency of the analysis process, so that others can see how interpretations and concepts were reached in relation to the data (Ritchie & Spencer, 1994; Ritchie et al., 2003, 2013). The defining feature of FA is the matrix output, a structure in which data are summarised and analysed by case and by code (Gale, Heath, Cameron, Rashid, & Redwood, 2013; Ritchie et al., 2003). It has been most often used for the analysis of semi-structured qualitative interviews (Gale et al., 2013).

FA shares many similarities with thematic analysis, such as the identification of recurring and significant themes, and relationships between different parts of the data (Gale et al., 2013). However, FA is distinct in emphasising the transparency of data analysis (Smith & Firth, 2011). There is also no requirement for solely inductive or deductive analysis, as FA can be adapted to include both when appropriate (Gale et al., 2013; Lacey & Luff, 2009; Ward, Furber, Tierney, & Swallow, 2013). Therefore, codes and themes in FA can include both a-priori as well as emergent concepts, where specific issues can be addressed based on previous literature or stakeholder requests (Lacey & Luff, 2009), whilst allowing for the generation of new, unrestricted themes based on unexpected or alternative themes (Gale et al., 2013). This combined approach is appropriate for research studies that have prior issues to explore, but also aim to allow other unexpected concepts to emerge (Gale et al., 2013; Parkinson, Eatough, Holmes, Stapley, & Midgley, 2016).

FA has five key stages: familiarisation; identifying a thematic framework; indexing; charting, and; mapping and interpretation (Ritchie et al., 2003, 2013). All of the data may be collected before analysis begins, or the researcher may choose to collect and analyse data concurrently (Ritchie et al., 2003, 2013).

The first stage, familiarisation, involves the researcher reading and re-reading transcripts, to gain in-depth familiarity with the data (Ritchie et al., 2003, 2013). In the second stage, identifying a thematic framework, the researcher notes themes in the data, informed by a-priori issues and emerging issues identified during familiarisation. A list of themes and sub-themes is created that are then organised in to a thematic framework, or index (Ritchie et al., 2003, 2013). The a-priori issues included in this analysis were the broad themes of physical health benefits, mental health benefits, social health benefits, and adjunct benefits.

In the third stage, indexing, the draft thematic framework is applied to the data. The researcher re-reads the transcripts, highlighting and annotating relevant passages with the theme they correspond to (Ritchie et al., 2003, 2013), also known as coding (Lacey & Luff, 2009). During this stage the thematic framework may be further developed and refined, with new themes emerging, or previous themes merged together (Ritchie et al., 2003).

In the fourth stage, charting, the data are organised in a thematic chart, to facilitate the next stage of data analysis (Ritchie et al., 2003, 2013). Data are organised according to the themes and sub-themes of the thematic framework, usually in the format of a table, or matrix, in Microsoft Excel or Word. In this study, Microsoft Excel was used. The table is composed of rows for each case or participant, and columns for each theme, usually containing easily discernible quotes and summaries (Furber, 2010; Gale et al., 2013). Data are then much more easily read across the whole dataset, and visualised as a whole (Furber, 2010; Gale et al., 2013; Lacey & Luff, 2009).

The final stage is mapping and interpretation, where the researcher seeks to understand the data (Ritchie et al., 2003, 2013). The chart is reviewed to understand the data set as a whole, making final alterations to themes and sub-themes (Furber, 2010). Once the thematic framework has been finalised, concepts, patterns and relationships can be identified (Parkinson et al., 2016). Comparing and contrasting data is an important stage of qualitative analysis, which is made easier by the ability to compare both across and within participants when using charting (Gale et al., 2013).

This series of stages allows the researcher to move back and forth between the data and the thematic framework, until a coherent account emerges (Smith & Firth, 2011). Another important advantage is that as the data is charted according to each individual participant, analysis of themes can be conducted across the whole data set, whilst retaining the context of individual participant's views, which may have otherwise been lost using alternative methods (Gale et al., 2013).

Although FA is a highly systematic method of qualitative data analysis, it is not a solution to common issues faced in qualitative research. Skills are still required to interpret data correctly, and make analytic choices, requiring reflexivity and rigour from the researcher (Gale et al., 2013). Therefore, before commencing data collection, a review of FA methodology was undertaken, including reviews of worked examples of FA, to become familiarised with the analysis process. The FA process was then overseen by the supervisory team.

7.3.6. Data analysis: secondary data

Categorical and qualitative responses from the feedback questionnaire were coded, and the data were managed using IBM SPSS statistics, version 23.0 (SPSS Inc., Illinois, USA), and analysed using descriptive statistics.

7.3.7. Reflexive statement

In section 4.4.4. I reflected on how my philosophical stance has shaped my approach to this evaluation. I considered it important to reflect further in this section, on how my values, assumptions and skills could influence the research strategy in this study. This reflexivity is intended to increase the rigour of this research process (Jootun, McGhee, & Marland, 2009).

I have previously noted that my personal values align most with pragmatism and interpretivism, and therefore characteristics of my research are selecting research methods appropriate to the problem, and recognising the contribution of human experience to knowledge where appropriate. As a real world intervention participated in by individuals, I have noted how it is my view that I cannot not make a full evaluation of this ERS without attempting to capture the experiences of these individuals. This has contributed to the rationale for qualitative inquiry in this study, in that I cannot understand the impact of participation from quantitative data alone (i.e.

studies 1 & 2). However, as a pragmatist, I could not disregard the usefulness of the feedback questionnaire in seeking to understand how participants benefit from this ERS, particularly given the larger sample size that may give a stronger generalisation.

When deciding upon the qualitative methodology for this study, I perceived that unstructured, open ended interviews would provide the most bias free evaluation, and allow participants to talk freely about their experiences without elicitation of desired responses. However, as a relative novice in qualitative research, I felt that I did not have the right skills to conduct quality interviews in this manner. Semi-structured interviews, with open ended questions, aligned more closely to my skill set, and I perceived the structure they provided would prevent deviation of the conversation to alternative topics. In the interviews I wanted to identify commonalities between participants' experiences and those reported in previous literature and the feedback questionnaire. However, the primary aim was to allow participants to openly express how they perceived to have benefited from participation, without leading questions. This is why all interviews opened with a question such as "could you talk to me about how you've benefitted from participation in this scheme". Probing questions were introduced thereafter only where it felt appropriate, so as not to elicit specific answers that might conform to previous observations.

Furthermore, I chose to analyse the interviews using framework analysis in regard of my limited experience of qualitative analysis. I felt that compared to other methods, this method provided a structured approach to the analysis of semi-structured interviews, which would facilitate a systematic and comprehensive analysis that would answer my research questions. The thematic framework, presented as a matrix in Microsoft Excel, could also allow easy collaboration with my supervisory team, due to the simplicity and transparency of the matrix.

7.4. Study 3 Data analysis and discussion

This section will analyse and discuss the results of data analyses, aiming to address the aims and objectives detailed in section 7.1. The findings of the feedback questionnaire will be presented, followed by the analysis of qualitative data.

7.4.1. Feedback questionnaire data

Feedback questionnaire data were available for 909 participants (female n=526, 57.9%, mean age 57 years). In total, 88.9% (n=808) reported that they had enjoyed participating in the scheme, 98.6% (n=896) found their exercise professional to be helpful, 59.7% (n=543) had experienced noticeable improvements in their fitness level, 87.1% (n=792) had perceived an improvement in quality of life from participation, and 100% (n=909) of those asked would recommend the scheme to another person. The most frequently reported benefits of participation are detailed in Table 7.2. These included having an improved sense of wellbeing (45.7%), improved confidence (35%), weight loss (29.4%), and making new friends (24.8%).

Table 7.2. Frequency counts for benefits of participation, surveyed using the 'Live Active' feedback questionnaire (total n=909).

Benefit	Experienced this benefit	
	n	%
Better sense of wellbeing	415	45.7
Improved confidence	318	35.0
Weight loss	267	29.4
Made new friends	225	24.8
Reduction in pain	193	21.2
Reduction in stress	166	18.3
Feel less isolated	147	16.2
Able to be more independent at home	139	15.3
Improved self-esteem	139	15.3
Improved sleep	101	11.1

Ten percent (n=98) of participants reported having reduced or stopped taking a medication as a result of participation in the scheme. These medications were: pain related such as painkillers (51%, n=50); CVD related such as BP medication (12.2%, n=12); diabetes related such as insulin (10.2%, n=10); respiratory disease related such as inhalers (10.2%, n=10); mental health related such as antidepressants (9.2%, n=9); gastrointestinal health related such as proton pump inhibitors (5.1%, n=5), or; other/undisclosed medications (2%, n=2).

7.4.2. Semi-structured interviews – participant characteristics

In total, 30 participants were interviewed. The characteristics of each participant are reported below in Table 7.3. A similar number of males (n=16, 53%) and females (n=14, 47%) were interviewed, mean age 61 years, mean duration of participation 9 months. All participants had at least one physical health condition, whilst 12 participants also had a mental health condition. The primary referral condition for each participant is listed, however all participants had at least one additional health condition. In relation to the target recruitment (Table 7.1), all sub-groups had at least 3 participants, with the exception of male or female, ≥65 years of age with a mental health condition (n=2, n=1 respectively).

Table 7.3. Characteristics of participants in the semi-structured interview sample.

Participant	Gender	Age	Mental health condition	Physical health condition	Primary referral condition	Duration of participation (months)
1	M	68	✓	✓	Mental health	6
2	F	62	✓	✓	CVD	6
3	F	41	✓	✓	Fibromyalgia	8
4	F	43		✓	MS	10
5	M	41		✓	Obesity	12
6	F	58		✓	Obesity	18
7	M	72		✓	Obesity	11
8	M	65		✓	Obesity	15
9	F	78		✓	COPD	18
10	M	74		✓	Diabetes	13
11	F	42	✓	✓	ME	12
12	F	57	✓	✓	Mental health	6
13	M	54	✓	✓	Mental health	6
14	M	69		✓	MSK	7
15	M	76		✓	Parkinson’s	6
16	M	44	✓	✓	MSK	7
17	M	67		✓	MS	6
18	M	57	✓	✓	Diabetes	12
19	M	63		✓	CVD	11
20	F	72		✓	MSK	12
21	M	65		✓	Diabetes	6
22	F	70		✓	MSK	6
23	M	44		✓	CVD	7
24	M	59	✓	✓	Fibromyalgia	7
25	F	60	✓	✓	MSK	6
26	F	68	✓	✓	MSK	9
27	M	75	✓	✓	CVD	9
28	F	67		✓	IGT	6
29	F	48		✓	MSK	14
30	F	80		✓	MSK	6

CVD, cardiovascular disease; COPD, chronic obstructive pulmonary disease; IGT, impaired glucose tolerance; ME, myalgic encephalomyelitis; MS, multiple sclerosis; MSK, musculoskeletal condition.

7.4.3. Semi-structured interviews – analysis and discussion

The purpose of the interviews was to understand how participants had benefitted from participation in the ‘Live Active’ ERS. In order to understand the context of perceived benefits, it was also important to understand what had led to each participant being referred to the scheme. Eight major themes arose from the interviews, with the sub-themes illustrated below in Table 7.4:

- The referral process
- PA related benefits
- Social health related benefits
- Physical health related benefits
- Mental health related benefits
- Adjunct benefits
- Limitations to PA and outcomes
- General feedback

Table 7.4. The major and sub-themes to emerge from the interviews.

Referral process	**Physical activity**	**Social health**	**Mental health**	
Reason for referral	Current PA level	Socialisation	Mental health	
Knowledge of scheme	PA level if not referred	Socialisation external to the scheme	PA as an escape	
Outcome expectations	PA as a benefit			
	PA related knowledge	Being around people in similar circumstances		
Physical health	**Adjunct**	**General feedback**		**Limitations**
Fitness	Confidence	The ERS staff	To physical activity	Free time
Weight loss	Sense of purpose	Beneficial		Inconvenient timing of activities
Mobility	Support	Enjoyment		
Strength	Enjoyment	Atmosphere		Health related symptoms
ADLs	Mind set	Cost (negative)		Finance
Energy	Dietary changes			Weather
Acute benefits	Encouraging PA in others			Motivation
Blood pressure	Medication withdrawn			Confidence
Pain	Independence			Support
Lung function	Quality of life			
Blood glucose	Cognitive health		To outcomes	Health related symptoms
Balance	Sleep quality			Medication
	Operation not required			Age
	PA as an alternative to medication			
	Return to work			
	Mastery of skill			
	Rids the body of toxins			
	Controlled environment			

ADLs, activities of daily living

In the following sections, each major theme and the corresponding sub-themes will be explored. They will also be discussed in relation to the outcomes of the feedback questionnaire and previous literature.

7.4.3.1. The referral process

The first theme was related to the context of the participant’s referral and their expectations of the scheme. The sub-themes were: reason for referral, knowledge of the scheme at the point of referral, and outcome expectations from participation.

Participants were referred for a variety of reasons, some of whom had complex physical and/or mental health histories. The majority were referred with relatively low risk physical health conditions such as: obesity; type 2 diabetes; hypertension, or; osteoarthritis. A minority had complex health histories such as: multiple complex morbidities; previous organ transplants; serious mental health illnesses, or; long-term memory loss.

> *“I needed an operation on one of me [sic] knees because I’d ruptured the cartilage … it turned out that they couldn’t do the operation at the time because my blood pressure was too high due to being overweight, and that I had class two diabetes” (P18, male, 57, diabetes)*
>
> *“Well I’ve had a lot of health issues, I’ve had two heart attacks and a bypass … diabetes, I’ve had lots of operations, I’ve got mobility problems with me [sic] legs and me [sic] back, I’ve got heart disease, so I’ve got a lot of on-going things, arthritis being another one, gout, all sorts of things, so yeah I’m a bit of a mixed bag medically” (P25, female, 60, MSK)*
>
> *“I had a nervous breakdown … I was found in the back garden, I was suicidal, I couldn’t take anymore” (P1, male, 68, mental health issue)*

This illustrates the diverse range of participants that ERSs may support, and therefore how the goals and outcome expectations could potentially differ between participants (Wormald et al., 2006). The broad range of health conditions that participants were referred for may reflect the referral criteria for this scheme, which is much broader than that of ERSs that have previously been evaluated. These schemes typically included a smaller number of referral condition pathways (Campbell et al., 2015; Pavey et al., 2011).

When asked about their expectations of the scheme at the time of referral, participants generally did not have any knowledge about the scheme, what support it would offer, or what would be required of their participation. There was a recurring notion of how participants had no knowledge of what the scheme was, but intended to attend an initial consultation to find out more information.

> *"I'd not heard of it before… she recommended it [nurse] so I thought I'd give it a try … I just thought I'd turn up and see what it was like really, you know?" (P13, male, 54, mental health issue)*

Several participants made vague references to the expectation that they would take part in physical activities such as "low impact circuit training" or "chair based classes", which had been described to them by their referrer. Two participants, referred from pulmonary rehabilitation and cardiac rehabilitation, expected the ERS to be delivered in a similar manner to these programmes.

> *"Just low impact circuit training was how it was described to me" (P7, male, 72, obesity)*

Two participants described how they did not know what to expect, but anticipated that they would not enjoy the scheme and would likely dropout. It was to their surprise that they found participation enjoyable. Females in particular described their pre-conceptions of the gym environment, such as the type of people who might be in attendance (e.g. thin and dressed in leotards). There was a sense that these pre-conceptions were a potential barrier to attendance, prompting anxieties about stepping in to the leisure centre for the first time, which P29 described as *"the fear factor"*.

> *"I came here not expecting to stay, and I thought well I'll give it a go, and I did, and I enjoyed it … the only time I thought about a gym was all these stick thin people in leotards, I was expecting that, I thought 'oh I am not gonna [sic] like this', but of course it's not like that at all" (P6, female, 58, obesity)*

> *"I think a lot of time when someone mentions the word gym, ugh, you think 'oh I can't do that', don't you? And it's a barrier right away and you think 'oh no', and you dismiss it, but I thought well I'll come along and I'll have a see how it goes" (P25, female, 60, MSK)*

The finding that at the point of referral, the majority of participants had little knowledge of what the scheme would entail, may suggest that discussion of the referral between the referrer and participant may have been insufficient. Qualitative evaluation of an ERS by Wormald et al. (2006) also found that participants did not know what to expect when they were referred to the scheme. Furthermore, pre-conceptions of the gym environment, particularly from females referred to an ERS, have also been previously reported (Wormald & Ingle, 2004), and are a known barrier to uptake and adherence (Morgan et al., 2016). Referral practices are an under-researched area of ERSs, but it is known that barriers such as time restrictions can prevent healthcare professionals such as GPs from making a referral (Din et al., 2015; Graham, Dugdill, & Cable, 2005). These barriers may therefore also impact the depth of discussion between the referrer and participant when a referral is made, on aspects such as what an ERS is, what support it could provide, and how participation could be beneficial for the individual.

A number of participants also discussed expectations of the ERS in terms of what outcomes they would achieve. These expectations varied between participants, potentially in relation to their present health conditions, but often related to increasing PA levels, or maintaining or improving aspects of physical and mental health such as wellbeing, mobility and body composition.

> *"that's why I'm coming here as well, to try and lose weight, increase me [sic] mobility, me [sic] wellbeing, the way I feel about me [sic] self, they're all sort of positive things moving forwards that I've come for" (P25, female, 60, MSK)*

> *"I know I'm not going to improve, I'm not going to be Mr Universe, but if I can keep my level mentally and physically, and my balance, as what it is now compared to what it could be … it's life changing MS, so I'm trying to keep my life as normal as possible" (P17, male, 67, MS)*

The findings of this sub-theme indicated that participants had varied outcome expectations from participation, which were largely related to their health conditions. This would support the approach of ERSs, to prescribe PA programmes according to the individual's circumstances, goals and preferences (NICE, 2014a). However, it also indicates that prior to attending the scheme, participants already had different outcome expectations and goals to those of which the scheme measured and deemed markers of successful participation.

The characteristics of participants in this sample highlights the diverse range of participants that ERSs may support, and begins to underline how the goals and outcome expectations of participants could differ significantly according to the context of referral (Wormald et al., 2006). Referral practices are an underexplored area of ERS operations. It is known that schemes can receive inappropriate referrals, such as those detailed in study 1, for persons with no motivation to change or improper explanation of the referral, for example. The observation that the majority of participants were not well informed about what the ERS would entail may have in-part contributed to the non-uptake observed in chapter study 1. Discussion of the referral between participant and referrer is therefore imperative, to identify eligibility and provide expectations, and these findings highlight why more research on referral practices is necessary (Din et al., 2015; Graham et al., 2005).

7.4.3.2. Physical activity related benefits

Quantitative evaluations of ERSs have rarely explored PA beyond the change in overall PA level. In contrast, PA was appreciated by participants in this study in a broader manner, with the sub-themes to emerge including: current PA level; PA level if not referred; PA as a benefit, and; PA related knowledge.

Almost all of the participants who were asked about their current physical activities reported that they were active on a regular basis, although the frequencies and types of activities varied between participants. In addition to structured activities such as PA classes, some participants discussed additional ways in which they tried to incorporate PA into their daily life, such as by walking instead of driving, or parking their car further away from a shop entrance.

> *"I do a couple of days here… I try and do the class and an hour in the gym on a Monday, and an hour, well at least an hour on a Wednesday, and Tuesdays and Thursdays I do at least an hour's swim" (P5, male, 41, obesity)*
>
> *"I try and do a bit of walking rather than use my car, if it's a short distance I try and motivate myself to walk rather than you know use my car" (P13, male, 54, mental health issue)*

Future intentions were also discussed, with fourteen participants making specific reference to their intention to sustain their current PA level, and eight participants planning to add additional activities to their existing PA programme such as an extra

PA class or gym session.

> *"this is something I'm gonna [sic] carry on for the rest of my life" (P19, male, 63, CVD)*

> *"I do the exercise class here on a Wednesday … I'm trying to fit swimming in and go swimming, that's why I got the new booklet … so I'm gonna [sic] try and find a slot and get [sic] swimming for about an hour" (P1, male, 68, mental health issue)*

At the time of the interview, three participants reported that they were inactive. For P12 (female, 57, mental health issue) this was temporary due to an acute infection, and she planned to return to her previous PA level. P29 (female, 48, MSK) had been regularly active, until a combination of illness and the expiration of subsidised membership. P18 (male, 57, diabetes) had remained inactive since referral, attributing this to time, motivation and the impact of depression, which are known correlates of physical inactivity (Bauman et al., 2012).

> *"I've been off for five weeks, but before that I was just doing the Monday class … I was gonna [sic] do more but when I became ill I just stopped … hopefully I'll come back on Monday" (P12, female, 57, mental health issue)*

> *"I lack motivation, I think a bit of depression sneaks in now and again" (P18, male, 57, diabetes)*

The finding that the majority of participants interviewed were active on a regular basis supports the quantitative findings of study 2, in that referral to the 'Live Active' scheme increases PA levels. Qualitative investigation has however complemented this quantitative data, with additional findings that participants plan to sustain, and in some cases further increase this PA. This is similar to previous qualitative research, which was able to report not only that participants were active, but that the process of becoming active had been gradual, and involved making progressions in PA (Wormald et al., 2006), to the point where PA is either sustained in the long term or further increased (Queen et al., 2015).

Furthermore, the experiences of P12 and P29, whose regular PA was interrupted for several weeks by illness, raises an important observation that would not ordinarily be captured by quantitative analyses. Had these participants attended a follow-up consultation during this period of illness and completed a PA questionnaire, they would likely have been categorised as inactive. Subsequent quantitative analyses of this data may have led researchers to incorrectly hypothesise that they were inactive due to the intervention being ineffective. These experiences add support for the unpredictable and non-linear journeys that can occur through an ERS, where factors such as illness can interrupt PA (James et al., 2009) and lead to an extended duration of participation in a scheme (Stewart et al., 2017).

All but one of those asked reported that if they had not been referred to the 'Live Active' scheme, they would most likely be inactive. Some attributed this to the support and structure that the scheme offered, and the motivational aspect of group based PA. In response to being asked if they would be active today if they had not been referred to the scheme:

> *"Probably not, no no, I think I was on a bit of a downwards spiral of you know becoming more and more depressed" (P13, male, 54, mental health issue)*
>
> *"Not as active as I am now, definitely not, and I suppose I'd have carried on eating you know the wrong sort of things, and then maybe I would have been diabetic, so it was a wake up call" (P28, female, 67, impaired glucose tolerance)*
>
> *"I don't think you can be as disciplined at home… I mean you get in to a class*

> *environment and you've always got someone cracking the whip you know saying 'right start off on that one, it's timed, see how long you can go for, right stop, change' so you've got, it's almost like the lion tamer you know? … I think that by coming here it gives you more discipline to your routine you know?" (P10, male, 74, diabetes)*

One male who had previously attended cardiac rehabilitation believed they may have otherwise maintained their PA levels, albeit with a struggle outside of the supported environment they had become accustomed to in cardiac rehabilitation:

> *"yeah but not as active I don't think… I'd have just been left to my own devices and I would have struggled because I was so used to being in that place [cardiac rehabilitation], anything else would have been a bit different" (P23, male, 44, CVD)*

The finding that the majority of participants would likely still be inactive if they had not been referred to the ERS supports the value of opportunistic identification and referral of inactive adults by health care professionals. The findings also suggest that participants find aspects such as the support given by ERS staff, and the motivational aspects of group based PA to be facilitators of PA. These two factors were previously identified by Morgan et al. (2016) as critical facilitators for adherence to ERSs, which developers should aim to incorporate. However, dependency upon these factors, and the impact on the likelihood of long-term independent PA has not been considered.

Six participants viewed becoming active as a benefit of participation. This included valuing the opportunity to become active (P27, male, 75, CVD), and to do activities of an intensity that they felt pushed them to make a concerted effort (P21, male, 65, diabetes). For P5, becoming active had been beyond their expectations.

> *"It's helped me to realise that I've got to exercise … without somewhere like*

> *this, what would I be doing? Or if I didn't know of anything like this?" (P27, male, 75, CVD)*

> *"It's something I look forwards to during the week because it's doing a bit more exercise that pushes your heart a bit more, cause [sic] I don't like running really, I wouldn't like do marathons and that, but I can see myself exercising and going a bit faster for you know, to catch something like a ball" (P21, male, 65, diabetes)*

> *"I'm doing the gym work and the swimming, which I wouldn't have dreamt… I mean I started swimming four months ago, this is the first time in over 20 years, and I've actually got back in to swimming now and I'm enjoying it" (P5, male, 41, obesity)*

No identified studies have reported participant's perceptions of becoming active as a benefit in this manner. These findings therefore contribute new knowledge of how some participants value the opportunities that an ERS referral provides.

New PA knowledge was a benefit for seven participants, which predominantly included perceptions of what constituted beneficial PA, but also how to be appropriately active with health conditions, and how to monitor and progress PA.

> *"I think at the beginning it was a bit of a learning curve, with my ME you've got to pace yourself, and when I love doing something I don't know when to stop, but I've got to stop before I get too tired, otherwise I won't be able to do it the next day" (P11, female, 42, ME)*

A recurrent theme was the perception that past activities, such as walking for leisure, were sufficient to maintain or improve health. Participants reported a realisation that, in comparison to activities such as gym programmes, these activities were not of a sufficient intensity to improve health, or did not provide comparable benefits.

> *"now there's a difference to when you've not used your muscles for years, you might think you are by just walking, but this is walking on a treadmill or doing two sets of bikes, lifting weights, it's different, different altogether, like it's proper exercise isn't it, you think walking is doing it" (P14, male, 69, MSK)*

New PA knowledge has previously been reported as a benefit of ERS participation (Fenton et al., 2015; Wormald & Ingle, 2004; Wormald et al., 2006). This is an important benefit, as previous research suggests that knowledge is a barrier to the maintenance of regular PA in individuals with health conditions (Dobson et al., 2016; Korkiakangas, Alahuhta, & Laitinen, 2009). It may therefore be relevant for future studies to evaluate change in PA knowledge in ERS participants, given that it may facilitate PA adherence beyond the scheme's duration.

The findings contained within this theme support the quantitative findings of study 2, that participation in the 'Live Active' ERS increases PA levels. However, this study has also provided the participant experiences behind this quantitative outcome, illustrating the types of activities that participants take part in, their intentions to sustain PA levels, and the development of PA knowledge to support the maintenance of PA. It has also highlighted that journeys through a scheme can be non-linear and interrupted by factors such as illness; information that would not have been available in the secondary analysis of data conducted in study 2. None of these findings appeared to be influenced by age, gender or health condition.

Given that PA is considered the primary outcome of ERS evaluations (Campbell et al., 2015; Pavey et al., 2011), relatively few qualitative studies have explored in-depth participants' experiences of becoming active through an ERS and their future intentions. Qualitative research has tended to focus on factors that facilitate adherence to PA (Morgan et al., 2016), or on broad evaluations of ERSs where PA experiences form one major theme (e.g. Queen et al., 2015; Wormald et al., 2006), such as in the present study. It may be useful for future research to focus on qualitative investigation of participants' PA experiences, in order to further understand, for example: why some participants become active and others do not, and the prevalence of interruptions to participation; to understand long-term intentions beyond the scheme, and; to understand if participants develop the relevant knowledge to facilitate long-term independent PA.

7.4.3.3. Social health related benefits

The most apparent benefit of participation in the 'Live Active' scheme was improved social health. The sub-themes of this category were socialisation, socialisation external to the scheme, and being around people in similar circumstances.

In total, 20 participants made references to socialisation as a benefit of participation, and this was often linked to the improvement of mental wellbeing and sense of purpose. This benefit was not exclusive to those attending PA classes, as those who were independently active in a gym setting also found socialisation amongst other gym users and staff members.

> *"the main thing I find actually is the social side of it … I do come for the exercise, but the meeting people is a part of the experience" (P7, male, 72, obesity, attends PA classes)*
>
> *"I didn't want to do any classes because of my depression, I wasn't in to like interacting as much, but what's been great is like you can go in to the gym and just… at first I wasn't speaking to people, but then like you get to know people, and I speak to people… but it's been dead gradual (P11, female, 42, ME, attends the gym)*

It was evident that a number of participants had been socially isolated and/or lonely prior to their referral, especially older, retired participants. This group recalled a growing sense of purposelessness in retirement, and a loss of friendships. It was interesting that these concerns were at the forefront of people's minds, openly recalling the social challenges of ageing, and reflecting upon how socialisation in the scheme was good for themselves, but also for other older aged participants who they perceived could be facing the same challenges.

> *"when I drove a bus every day, I was on the go all the time, but when you're retired, you feel as if society has had enough of me [sic], they don't want to know me" (P1, male, 68, mental health issue)*
>
> *"it's a good thing, a very good thing, especially for people of my age … the terrible thing with when you do get older is your friends are no longer there, you know? A lot of them at my age, it might sound morbid but they die you know?" (P27, male, 75, CVD)*
>
> *"I think that's good for everybody really, just to chat to people isn't it? … that's what I mean about this Monday class, I think some people perhaps don't see that many people in a week" (P26, female, 68, MSK)*

Social isolation was also present amongst younger age groups, particularly those who were unemployed or recovering from illness at home, and had reduced social contact or networks.

> *"with being poorly for so long… friendships had gone out the window, and socialising had completely gone and what not, so coming here to the class on a Monday was brilliant, I made loads of new friends, there was a socialising aspect to it… so it was good for the mind as well as the body" (P23, male, 44, CVD)*
>
> *"I have been out of work for a long time with my illnesses and that, I have done some voluntary work, I do voluntary work just to you know get out and*

> *meet people and that, and this is like an extension of that, it's just you know getting out and meeting people rather than being stuck at home" (P5, male, 41, obesity)*

It was clear that individuals who valued the social aspects of the scheme found this to be a facilitator for adherence. They noted how they looked forwards to seeing each other at the activities, and this motivated them to attend when they otherwise didn't want to.

> *"if you're not feeling up to coming, you want to come just to meet the other people" (P8, male, 65, obesity)*

Several qualitative studies have reported socialisation as a benefit of ERSs (Crone et al., 2005; Stathi et al., 2004; Vinson & Parker, 2012; Wormald et al., 2006). In this study, whilst socialisation was reported by 20 participants, it appeared to be a particularly important benefit for those who had previously felt socially isolated and/or lonely. Social isolation and loneliness are considered to be challenges associated with older age (Cotterell, Buffel, & Phillipson, 2018; Davidson & Rossall, 2015; Ong, Uchino, & Wethington, 2016). However, in this study social isolation was also present amongst younger age groups. In comparison to research on older age groups, investigation of social isolation in working age groups appears less frequent. However, it is known that the risk of social isolation is higher in working age adults who have LTCs (Meek et al., 2018), are unemployed, or are experiencing long-term sick leave (Vingård, Alexanderson, & Norlund, 2004), which may in part explain the experiences of these participants.

Furthermore, in this study socialisation was evident in both group and independent PA settings. Whilst the opportunity for socialisation may appear to be more likely in group PA settings, this is not guaranteed, and can depend upon the structure of the class and qualities of the class leader (Stathi et al., 2004). These findings therefore suggest that the PA classes in this scheme were effective at facilitating social interaction. The finding that socialisation contributed to adherence supports previous qualitative research, which has found that social inclusion can contribute to adherence to ERSs (Morgan et al., 2016).

The second sub-theme of this category was ‘socialisation external to the scheme’. In some instances, friendship groups had been formed during PA classes, and this had led to groups organising attendance at physical activities external to the scheme, such as bowling and line dancing.

> *“after a bit a few of us just got together… we go ten pin bowling and stuff like that, there’s about four of five of us go to line dancing now you know? Even that’s created by meeting them here” (P8, male, 65, obesity)*

Encouraging dependence upon supported PA is a potential concern for ERSs (Wormald et al., 2006). However, social networks can be a modifiable mediator of PA behaviour change (Bauman et al., 2012), and the formation of these groups, and the social support they offer, may therefore facilitate maintenance of PA beyond the scheme and reduce the likelihood of reliance on ERSs (Scott et al., 2015).

The final sub-theme to emerge from this category, albeit only from female participants attending group classes, was being active amongst people in similar circumstances, such as those with health conditions. This appeared to be related to feelings of comfort around others, as P4 for example felt as though they fitted in rather than stood out because they had a health condition:

> *“everybody is kind of like in the same boat, cause [sic] there’s something wrong with everybody, you know? Like it’s not that you come here and think ‘oh like I’ve got MS’, everybody’s got something” (P4, female, 43, MS).*

For P25, as well as comfort, it gave additional perspective about their circumstances and made them feel fortunate in comparison to others, their explanation synonymous with downwards comparisons (Wills, 1981) observed in social comparison theory (Festinger, 1954).

> *“you find that there’s other people here who are worse off than you are, I know it sounds… but that makes you feel better, thinking ‘oh well at least I’m you know not in a wheelchair’ … it’s like keeping yourself on a high if you get my meaning, making yourself feel better, more comfortable, and thinking well ‘I’m lucky’” (P25, female, 60, MSK)*

A review of predominantly qualitative research by Morgan et al. (2016) found engagement with likeminded individuals to be a facilitator of enjoyment and adherence to an ERS in several studies. Whilst this study did not link being active amongst people in similar circumstances with adherence, it does demonstrate that it is an important factor for some participants, particularly in terms of facilitating comfort in group environments. This may be related to participants’ physical self-perceptions, which may increase when they identify similarities between themselves and others in the PA class (Fox, 2000).

Improved social health was the most apparent major theme to emerge from the semi-structured interviews, and reporting of this benefit did not appear to differ between age, gender or health condition categories. It was reported across all of these categories, and in some instances was related to a person’s background, such as social isolation, rather than demographics. The results of the feedback questionnaire also identified frequent reporting of reduced isolation and making new friends, strengthening the evidence for improved social health as a result of participation in the scheme. This was surprising, as whilst several qualitative studies have reported social health related benefits of ERSs (Fenton et al., 2015; Hardcastle & Taylor, 2001; Stathi et al., 2004), this has not reported in any recent systematic reviews of

ERS impact (Campbell et al., 2015; Pavey et al., 2011), nor has it been accepted as an indicator of success by the wider community of researchers (Fenton et al., 2015; Hardcastle & Taylor, 2001; Stathi et al., 2004). Quantitative measures of social health from a large sample have rarely been collected in ERS evaluations, and it is therefore novel to find that around a quarter of the feedback questionnaire sample experienced improved social health. It was also surprising that as many as two thirds of the interview sample identified socialisation as a benefit. These findings regarding the impact of ERS participation on social health are promising, given the recent drive to reduce social isolation and loneliness in the UK, with the appointment of the first ‘minister for loneliness’ (Department for Digital, Culture, Media & Sport, 2018).

The reluctance of positivist evaluations, and evaluation frameworks to incorporate social health indicators is surprising as a growing body of research has linked improved social health to improved overall health and economical expenditure. Social isolation and loneliness, for example, are associated with negative health behaviours such as physical inactivity and smoking, and biological processes such as rising BP and systemic inflammation, all of which are risk factors for LTCs (Shankar, McMunn, Banks, & Steptoe, 2011; Smith, Stewart, Riddell, & Victor, 2018). They have both therefore been established as independent risk factors for mortality, comparable with well established risk factors such as smoking and obesity (Holt-Lunstad et al., 2010, 2015). Deteriorations in health associated with social isolation and/or loneliness also lead to higher rates of health care utilisation and therefore direct healthcare costs (Gerst-Emerson & Jayawardhana, 2015; Molloy et al., 2010; Valtorta & Hanratty, 2012). Given the benefits of positive social health, and the recent mandate to reduce social isolation and loneliness in the UK (Department for Digital, Culture, Media & Sport, 2018), the value of improved social health through participation in an ERS cannot be overlooked any further. These findings therefore support arguments for the inclusion of social health indicators in future evaluations and evaluation frameworks (Dugdill et al., 2005; Mills et al., 2012).

7.4.3.4. Physical health related benefits

As a consequence of participation in the scheme, participants discussed a variety of physical health related benefits such as improved: physical fitness; mobility; weight loss; physical strength; ability to carry out activities of daily living; energy; pain; BP, and; blood glucose (see Table 7.4). Some benefits, such as improved lung function, blood glucose, and balance were discussed in relation to specific health conditions.

Improved physical fitness was the most frequently reported benefit in this theme, reported by participants with a wide range of health conditions, supported by the feedback questionnaire in which 60% of participants reported noticeable improvements in fitness. Several participants illustrated an improvement in fitness with examples of progression in exercise intensity, or by reporting that they found previously difficult activities much easier.

> *"I do the star jumps, nobody else in the class does the star jumps, I think the first time I came I did about three and I thought hey, I've done 60 today … and instead of running around the room I was only walking around, so it's helped me tremendously you know" (P10, male, 74, diabetes)*

Fitness has been highlighted as a benefit in previous qualitative research (e.g. Queen et al., 2015; Wormald & Ingle, 2004; Wormald et al., 2006). Despite this, very few quantitative evaluations have identified the impact of participation on fitness (e.g. Isaacs et al., 2007; Stewart et al., 2017).

Weight loss was also a common theme, particularly amongst participants who were overweight or obese. However, three participants, all of whom were overweight, noted that this was in combination with dietary changes. One of these participants had made dietary changes independently, and two had received specialist dietary support. Furthermore, P21 (male, 65, diabetes) could not discount the potential contribution of medication to their weight loss.

> *"I'm overweight but it's coming down and it's staying off which is good … the exercise definitely helps and I found it easier, it wasn't as difficult because I was eating a lot of rubbish … I see Referrer 5… she's a health worker and she's helped sort of recommending things to me to eat… it's really made a difference you know" (P28, female, 67, impaired glucose tolerance)*
>
> *"I know I've lost weight, but I don't know if it's the tablets, or if it's the exercise, I don't know" (P21, male, 65, diabetes)*

P11 (female, 42, ME) was the only participant to discuss changes in body composition rather than weight loss. By participating in resistance exercise alone she had experienced an increase in muscle mass and therefore bodyweight and BMI, but a reduction in girth measurements.

> *"it has helped my weight loss but I've started to put on now … but me [sic] figure's totally changed … It's all sort of sucked in a bit if you know what I mean, and I've got definition in me [sic] muscles, because I had no muscle … so I've more shrunk in inches rather than weight now" (P11, female, 42, ME)*

Weight loss was not evident for all participants though, and six participants discussed their disappointment with poor weight loss. Two participants felt that medication was inhibiting weight loss, one had lost weight but subsequently regained it after a period of inactivity due to illness, whilst a further participant struggled to make positive dietary changes to support weight loss without affecting the symptoms of colitis.

> *"I was hoping for weight loss, but it's not happening because I'm on steroids" (P23, male, 44, CVD)*

> *"I've been off for five weeks … since I've been off, I've put the weight back on" (P12, female, 57, mental health issue)*

> *"I thought I'd like to get this weight down … like I say, with trying to change me [sic] diet it's causing me [sic] colitis to flare up" (P20, female, 72, MSK)*

Similar to the findings of Mills et al. (2012), despite being considered an important quantitative outcome by researchers, weight loss was not central to participant's views, but considered alongside several other outcomes of equal importance. The mixed experiences regarding weight loss provide additional knowledge to the quantitative findings of study 2, and the results of the feedback questionnaire that found weight loss to be a benefit for 29.4% of the sample. They also highlight the contribution of external factors such as diet and medication to weight loss, which have not been considered by previous research when evaluating the impact of an ERS on weight loss. Furthermore, the experiences of P11 regarding changes in girth measurements but not overall body weight exemplifies how body composition measures such as BMI can generate incorrect conclusions about the impact of an intervention (Toomey et al., 2015). This measure may therefore also generate incorrect conclusions about the health status of an individual (Toomey et al., 2015), as improvements in outcomes such as visceral fat (Lee et al., 2005), BP, plasma lipids and biomolecular markers of CVD risk (Webb et al., 2016) have been demonstrated without significant weight loss. This therefore emphasises the importance of considering alternative measures, such as girth measurements, when evaluating an intervention (Toomey et al., 2015).

Improved mobility was a further frequent benefit, particularly for those with existing musculoskeletal or mobility issues. Examples included the progression from "shuffling" to a near normal walking gait (P5, male, 41, obesity), being able to walk up and down stairs unaided (P8, male, 65, obesity), and being able to walk confidently without a walking stick (P22, female, 70, MSK).

> *"I'm standing more upright, and me [sic] gait, that's a lot better, I'm more steady and balanced and the strides a lot better … I was more or less shuffling, and I was unsteady on my feet but now I'm walking and I'm actually walking and looking up and around rather than looking where I'm going" (P5, male, 41, obesity)*

Improved mobility in persons with existing musculoskeletal or mobility issues illustrates how participants can perceive success according to their own personal circumstances. Improved mobility has previously been reported in persons with existing issues (Stathi et al., 2004; Wormald & Ingle, 2004) and the general ERS population (Stathi et al., 2004; Wormald et al., 2006). These findings provide further context to the results of study 2, in that participation in an ERS can help to reduce the burden of musculoskeletal conditions (Arthritis Research UK, 2018).

Changes in physical fitness, strength and mobility were often directly related to improvements in outcomes such as daily energy levels and the ability to complete activities of daily living. Some noted how they previously felt fatigued day-to-day, but since increasing their PA and fitness levels, they felt more energised. This was evident in participants with health conditions such as fibromyalgia, which are characterised by feelings of fatigue, but also in those with other health conditions. Several participants also described how activities such as housework were easier to accomplish, due to improvements in fitness, strength and mobility.

> *"I do suffer with like the fatigue of the fibro, but yeah it's boosted me [sic] energy" (P3, female, 41, fibromyalgia)*

> *"I think I have more energy to do things, you know? I suppose if your fitness goes you do tend to lose the energy to do things" (P13, male, 54, mental*

> *health issue)*
>
> *"I find things easier, I used to struggle doing basic housework, you know cleaning windows where you're stretching and perhaps lowering, but now I can do them and not even think about it" (P6, female, 58, obesity)*

Improvements in fitness, strength and mobility facilitating activities of daily living was also observed by Wormald and Ingle (2004). These examples demonstrate that improvements in one physical health related outcome can have a ripple effect on the improvement of other physical health related outcomes, many of which have not been considered or quantified by previous evaluations, but may have important consequences. The ability to carry out activities of daily living, for example, is an important benefit, as this signifies that individuals may be able to remain independent in their own homes for longer without external support (Whitehead, Worthington, Parry, Walker, & Drummond, 2015).

The number of physical health related benefits reported in this study is in contrast with the select few markers included in quantitative evaluations, or suggested by evaluation frameworks. Furthermore, some of these benefits were more prevalent amongst participants with certain health conditions, such as mobility and musculoskeletal conditions, illustrating how perceived benefits can be related to a participant's circumstances (Wormald et al., 2006).

Reporting of outcomes such as improved physical fitness, weight loss and reduced pain supports the findings of the feedback questionnaire, which also found these benefits to be prevalent across a large sample of the scheme. Many of these outcomes, such as improved fitness, strength, mobility and ability to carry out activities of daily living have been highlighted in previous qualitative research as positive benefits of participation (e.g. Queen et al., 2015; Wormald & Ingle, 2004; Wormald et al., 2006). The wider literature also suggests that these are markers of health associated with risk of mortality (e.g. Aune et al., 2016; Kim et al., 2018), and therefore substantial burden to both the individual and society (e.g. Arthritis Research UK, 2018; Baker, 2018). Despite this, they have rarely been measured in quantitative evaluations, and their inclusion may contribute to a more representative evaluation of ERSs (Mills et al., 2012). The SEF (Cavill et al., 2012) does recommend the measurement of physical outcomes such as BP, cholesterol and fitness, but notes that evaluators may not have the resources to do so. This may explain why few studies have evaluated these outcomes, and a feedback questionnaire such as the one implemented by this scheme may be an alternative way for participants to self-report these outcomes.

7.4.3.5. Mental health related benefits

Improved mental health was evident in those referred with and without pre-existing mental health conditions, and amongst those of different genders and age groups. This finding therefore supports the quantitative findings of improved mental health found in study 2. The sub-themes in this category were improved mental health, and PA as an escape from mental health issues.

Several participants with existing mental health issues noted how their mood and depression had improved as a result of participation. The self-reported mechanisms behind this improvement varied by participant, but examples included: PA being a new interest to focus on; socialisation during activities following a period of social isolation, and; an improvement in physical health symptoms that had previously adversely affected mental health.

> *“I suppose if you’re doing things it stops your mind dwelling on things doesn’t it? It gives you other things to focus on doesn’t it? … I think I was on a bit of a downwards spiral of you know becoming more and more depressed, I think it’s given me more of an impetus to be more positive about things and do more*

> *things" (P13, male, 54, mental health)*
>
> *"I definitely feel better mentally than I did when I first started coming, because I was getting quite depressed when I first came, and that seems to have eased … when I was feeling like I was, if I then got aches and pains it was just another thing on top of it, and it sort of pulled me down, do you know what I mean?" (P26, female, 68, MSK)*

Four participants without known mental health conditions also made specific references to improved mental health using terms such as 'mood' and 'mental health'. For two participants this was related to opportunities for socialisation, with one having recently spent a long time isolated at home recovering from surgery.

> *"it was good for the mind as well as the body … I think it's definitely helped with me [sic] mood, I had a right bad temper but that's what happens when you're cooped up all day, you end up going nuts" (P23, male, 44, CVD)*

Improvements in mental health have previously been reported in qualitative research (Crone et al., 2005; Fenton et al., 2015; Stathi et al., 2004), however this study has gone further by identifying that these improvements occur both in those with and without pre-existing mental health issues. Crone et al. (2005) investigated the factors that linked PA with improved mental health in ERS participants, identifying self-acceptance as a core category. In contrast, social interaction appeared to be a key mediator of improved mental health in this study, potentially linked to previous social isolation and/or loneliness.

The second sub-theme was 'PA as an escape from mental health issues', which was only reported amongst younger aged participants (42-44 years). Three participants, two of which had known mental health issues, described how participation in PA was an escape from mental health issues such as stress, depression and post-traumatic stress disorder. This was particularly evident for P11 who experienced personal issues throughout their participation, and described their gym sessions as the only time they had for themselves. They felt a noticeable sense of relief from their daily stressors when they attended the gym:

> *"it's the only thing really that I do for me … I'll be sat there thinking 'I can't cope with this', but I know I've got that time in the gym that's mine … I walk through the door and leave my troubles behind like 'bye!' … it's my refuge now" (P11, female, 42, ME)*

No identified previous evaluations of ERSs have addressed PA as an escape from mental health issues. It is, however, known that PA can be used by individuals as a distraction and coping mechanism for mental health issues, and subsequently improved mood (Craft & Perna, 2004; Mikkelsen, Stojanovska, Polenakovic, Bosevski, & Apostolopoulos, 2017), known as distraction hypothesis (Bahrke & Morgan, 1978).

Quantitative evaluations of ERS have tended to find mixed results for changes in mental health (Campbell et al., 2015; Pavey et al., 2011). In contrast, improved mental health has been consistently reported throughout qualitative research (e.g. Fenton et al., 2015; Stathi et al., 2004; Wormald & Ingle, 2004). In this study, around a third of interview participants reported an improvement in mental health, which supports the quantitative findings of study 2, and wider literature demonstrating the mental health benefits of PA (Rebar et al., 2015; Schuch et al., 2016). It could be postulated that quantitative measures of mental health at the population level have not provided a true representation of impact, if the sensitive nature of questionnaires has led to bias in self-reporting (Tourangeau & Yan, 2007). However, this may not explain the mismatch between the findings of quantitative and qualitative studies, if willingness to disclose is also reduced in qualitative interviews (Bowling, 2005).

Improved mental health in those with and without existing mental health issues is an important outcome, especially given recent mandates to address the prevention and management of mental health issues in the population (Department of Health & Social Care, 2018a; NHS, 2016). In England, one in six adults meet the criteria for a common mental health condition (e.g. anxiety, depression) (McManus, Bebbington, Jenkins, & Brugha, 2016). They are the third leading cause of years lived with disability in the UK (James et al., 2018), and individuals with these conditions experience a higher mortality rate than the general population (Walker, McGee, & Druss, 2015). Individuals with co-morbid physical and mental health conditions, characteristic of this interview sample, are more likely to participate in adverse health behaviours, and experience worse health outcomes than those with physical health conditions only (Naylor et al., 2012; Walker & Druss, 2016). As a result, health care utilisation is increased, raising direct healthcare costs by £8-13 billion per year in England (Naylor et al., 2012). Any preventive or treatment role that PA can offer to those at risk of, or who already have mental health issues is therefore valuable if it can reduce this burden to both the individual and health care system. The findings of this theme therefore support arguments for greater recognition of the value of ERSs in the management of mental health issues, and for improved mental health as an indicator of success that should be routinely included in ERS evaluations (Crone et al., 2005; Morgan et al., 2016). Currently, the SEF suggests the measurement of mental health as part of a quality of life measurement (Cavill et al., 2012).

7.4.3.6. Adjunct benefits

The ‘adjunct benefits of participation’ theme comprised the highest number of sub-themes, most likely as this category demonstrated personally meaningful benefits. The most frequently reported adjunct benefits were improved confidence, having a sense of purpose, the support given by exercise professionals, enjoyment, making dietary improvements, and encouraging others to become physically active. Other, less frequently reported benefits included changes to: medication; health care utilisation; independence; quality of life; cognitive health, and; sleep quality (see Table 7.4).

Confidence was a frequently reported benefit in both the feedback questionnaire and interview samples. In the interviews, confidence was expressed in four different ways: self-confidence; confidence to be active; confidence in mobility, and; confidence in health. Three participants related improved self-confidence to the ability to leave the house and be around others after a period of social isolation. One participant related it to weight loss and improved self-image, whilst another related it to regaining self-confidence that was lost after a diagnosis of MS and decline in physical abilities.

> *"I think if someone is a little bit… has a lack of confidence, you do tend to sit at home and think 'I don't really want to go anywhere, I don't really want to be in a crowd or do anything', which can be quite bad for you … I was getting to that stage where I just didn't want to go anywhere at all, this brings you out of yourself" (P6, female, 58, obesity)*

Three participants also reported improved confidence to be active. This was related to, for example, the ability to be active with multiple health conditions, to be independently active, or to be active without concerns about being in physical pain.

> "*before I wasn't sure about if I could do it, like if I could do the exercises without it being painful, but since I've been doing it it's helped it … I've got more confidence to do that [the exercises]" (P12, female, 57, mental health issue)*

Improved confidence in mobility was a benefit for two participants, who had mobility issues. For P22, this was a personally meaningful benefit. Prior to referral they were dependent on their spouse, and unable to do activities such as shopping alone. After improving their mobility, and their confidence in this, they were now able to go shopping alone, demonstrating a new level of independence.

> *"It's given me a lot more confidence … I can try things that I wouldn't have tried 12 months ago … if I wanted to go and do some shopping I'd say to my husband 'can we go shopping?', so that he could assist me, and I used to stand and push the trolley around and he used to run and get everything we needed, now it's 'I'm going shopping, do you want to come?'" (P22, female, 70, MSK)*

Improved confidence was a frequently reported benefit from the feedback questionnaire. However, the qualitative findings have complemented these data by uncovering in what way confidence has improved. Confidence has rarely been identified in previous ERS evaluations, however these findings do support qualitative research that identified improved self-confidence, and the confidence to be physically active through ERS participation (Wormald & Ingle, 2004; Wormald et al., 2006). It is known that the development of confidence to be active is an important mediator of PA adherence (Ashford et al., 2010; Williams & French, 2011).

A further sub-theme, improved sense of purpose, was evident in older, retired participants, and younger, unemployed participants. Older participants in particular appeared to have a growing sense of purposelessness following retirement, feeling isolated at home or stuck in monotonous routines. This was improved by having a new interest in PA, activities to attend that interrupted usual daily routines, and by meeting new people.

> *"when I drove a bus every day, I was on the go all the time, but when you're retired, you feel as if society has had enough of me [sic], they don't want to*

> *know me" (P1, male, 68, mental health)*
>
> *"I don't work either, so I've gone from like running pubs to looking after my children, and I've been ill, and now it's like I've got an interest and I'm learning something, so that's good for me" (P11, female, 42, ME)*

A further two participants reflected upon observations of individuals who had no interests following their retirement and stayed at home, and how PA could prevent this.

> *"you read about these people, as soon as they've retired they've had no interests outside of the house, and they just sit there and stagnate in front of the TV, and they pop off in five years … I want as I say, I want to get to one hundred [years old]" (P7, male, 72, obesity)*
>
> *"I think it's really good that we've got something like this that encourages people to get more active and to join in you know? It's too easy to sit at home you know? Some people, they just don't do anything when they retire" (P28, female, 67, impaired glucose tolerance)*

A renewed sense of purpose from participation in an ERS is a frequent finding of previous qualitative research (e.g. Hardcastle & Taylor, 2001; Scott et al., 2015; Stathi et al., 2004; Wormald et al., 2006). However, this study has highlighted that purposelessness is not limited to older age groups. This is an important improvement, as having a sense of purpose is more than just a psychological resource, having an impact on chronic physiological stress (Zilioli, Slatcher, Ong, & Gruenewald, 2015), healthy ageing (Fogelman & Canli, 2015), protection from events such stroke and myocardial infarction (Kim, Sun, Park, Kubzansky, & Peterson, 2013a; Kim, Sun, Park, & Peterson, 2013b), and risk of all-cause mortality (Boyle, Barnes, Buchman, & Bennett, 2009).

Several participants appreciated the knowledge of the ERS staff and the individualised programmes they had created, but ultimately did not appear to be dependent upon receiving this support. However, two participants made specific references to valuing the perceived safety of supervised exercise, similar to the findings of Mills et al. (2012), which may lead to a potential dependency on the scheme rather than fostering independent PA.

> *"I think it is a big positive that you can come to a class like I say that's supervised, and you know they're qualified, and you know if you get in to difficulties someone is gonna [sic] be at hand … I feel reassured with that as well, and that's another reason why I like to come, because I feel safe" (P25, female, 60, MSK)*

The support given by the ERS staff is a less frequently reported benefit in the literature (e.g. Stathi et al., 2004; Wormald & Ingle, 2004; Wormald et al., 2006). This study supports observations that the qualities of the exercise professional, and the support they provide are important to participants who are adjusting to a new PA environment (Mills et al., 2012; Wormald et al., 2006). However, there is a potential risk of dependence upon this support that may reduce the likelihood of long-term independent PA (Mills et al., 2012; Wormald et al., 2006).

Some of the lesser reported benefits in this study include those that have been reported in previous literature, and those that have potential health and economic consequences. Benefits such as dietary improvements, encouraging PA in others, health care utilisation, and changes to medication have for example been reported in previous literature (Hardcastle & Taylor, 2001; Isaacs et al., 2007; Queen et al., 2015; Wormald et al., 2006).

> *"I sleep better, I've come off two of me [sic] drugs for pain, and I was supposed to have an operation but I don't have to have it now … my father he*

> *actually needs something like this so I'm trying to get him to come … me [sic] eldest… he's seen what I get from exercise so he's going out on his bike … and also I've changed me [sic] cooking… I now just cook from scratch with everything" (P11, female, 42, ME)*

Whilst only two interview participants made references to medication changes, this was evident in over 10% of the feedback questionnaire sample, indicating that this may be a frequent benefit of the scheme. Not only do medication changes signify an improvement in health that necessitates this change, they also signify economic benefits to individuals and the healthcare system in prescription costs. The estimated NHS spend on medications in England is currently £17.4 billion per year, rising by 5% each year (Ewbank, Omojomolo, Sullivan, & McKenna, 2018). There is a need to halt this increase in expenditure, using non-medical forms of intervention where evidence dictates, such as ERSs and other social prescriptions (Thomson et al., 2015).

The variation in adjunct benefits, reported in both the feedback questionnaire and interviews, highlights the multitude of ways in which a person can benefit from participation in an ERS. This supports the argument that participants who enter an ERS may have diverse backgrounds, and a variety of physical, mental, emotional and social issues, and therefore have many potential ways of benefitting from participation (Wormald et al., 2006). This was also illustrated by the observation that no single demographic such as age or health condition tended to influence the reporting of certain outcomes; but rather the participant's background and what was personally meaningful to them.

These reported outcomes are important benefits that demonstrate a personally relevant impact for each individual, but are those that would not have been captured by traditional quantitative evaluations, nor would they be prompted to by current evaluation frameworks (e.g. the SEF - Cavill et al., 2012), thereby underestimating the impact of the scheme. Some of these outcomes, such as withdrawal of medications and returning to work, are those that have cost implications that have not been considered by models of cost-effectiveness (e.g. Anokye et al., 2011). These findings therefore support the value of mixed methods research when evaluating the impact of schemes (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008). Knowledge of these wider benefits may support evaluators to develop a broader range of outcome indicators for inclusion in future research (Mills et al., 2012; NICE, 2014a) and facilitate the development of evaluation frameworks that prompt evaluators to consider the adjunct benefits of participation.

7.4.3.7. Limitations to physical activity and outcomes

Several participants discussed factors that limited PA progression and/or improvement in outcomes. Factors limiting PA progression are first discussed, followed by factors limiting outcomes.

Four participants noted how they wanted to do more activities but had limited free time. Interestingly, three of these participants were retired, citing lack of free time due to looking after their grandchildren. This conflicts with previous hypotheses that older aged adults have more free time available to incorporate PA in to their lifestyle (Tobi et al., 2012).

> *“sounds awful when you’re retired, but it’s fitting in things, cause [sic] you’ve no time to work … especially if you’ve got, like I’ve got grand children, so like we help out in the school holidays and half term and things like that” (P28, female, 67, impaired glucose tolerance)*

In addition to lack of free time, inconvenient timing of certain activities was a barrier for four participants. Two participants described clashes between PA classes and personal activities. One participant, for example, was remaining in a low level PA class because the next level class clashed with caregiving duties. A further two female participants, of middle to older age (57, 70 years), reported a desire to attend water based classes, but these classes were only provided in the evenings, and both participants described caution in leaving the house in the evening. Off-peak scheduling of activities has previously been identified as a facilitator of adherence to ERSs (Morgan et al., 2016).

> *"aqua fit is always on at night here, and I said you know 'you're not going to get ladies of my age, without transport, coming out in the evening, they'll come out when it's light'" (P22, female, 70, MSK)*

The most frequent limitation to PA, however, was health related symptoms. This was particularly common amongst older age participants, and related to specific health conditions. The impact of arthritis was the most prevalent barrier to PA. This group of participants perceived limitations to what activities they could do without aggravating arthritis related pain.

> *"I'm limited, limited in what I can do yeah, but there's not a lot I can do about that (P7, male, 72, obesity) [In reference to osteoarthritis and spondylitis]*
>
> *"Well I've got arthritis so I'm a bit limited you know in what I can do … so that's why I just do the leg exercises really" (P20, female, 72, MSK) [In reference to arthritis of the hand and wrist]*

Further examples included the impact of MS on fitness and balance, and therefore types of PA possible (P17, male, 67, MS), and the impact of depression on motivation to be active (P18, male, 57, diabetes). P24, who was active, albeit infrequently, had multiple conditions that impacted the ability to be active. Whilst cardiomyopathy limited the volume of PA possible, the biggest limitation for this participant was fibromyalgia related pain:

> *"after the first couple of minutes I started getting really bad chest pains, because I've got cardiomyopathy … I'm blaming all this on the fibro … sometimes these work for some people, but not for everyone, I think my demise is my my fibromyalgia" (P24, male, 59, fibromyalgia)*

Factors such as free time and health problems have been previously linked to reduced likelihood of adherence to an ERS (Morgan et al., 2016). However, in this study, these factors did not appear to impact adherence but rather the ability to progress the intensity, frequency or type of PA. This may be due to participants accepting their limitations and working within them, rather than claiming denial of such limitations leading to dropout (Scott et al., 2015). Furthermore, previous research has highlighted the potential for participants to feel restricted by health related symptoms, and therefore different participants may have different perceptions of how success can be defined as a result of these limitations (Mills et al., 2012; Tobi et al., 2012). The findings of this theme would tend to agree with these previous observations.

In contrast, the reporting of limited free time by retired, older aged adults, conflicts with previous hypotheses that they have more free time available to incorporate PA in to their lifestyle (Tobi et al., 2012). This may be related to both the SE context of Tameside, and national context, whereby today over 40% of grandparents provide regular childcare (Age UK, 2017), potentially related to significant increases in the cost of childcare for parents who wish to return to work (Family & Childcare Trust, 2018).

Participants also perceived the extent to which outcomes could be improved to be limited by health related symptoms, medication and age. Health related limitations were mostly reported by older age participants, and were related to the participant's specific health conditions. Examples included improvements in mobility limited by osteoarthritis (P7, male, 72, obesity), improvements in strength limited by Parkinson's disease (P15, male, 76, Parkinson's), improvements in fitness limited by MS (P17, male, 67, MS), and improvements in body weight limited by the impact of colitis (P20, female, 72, MSK).

> *"me [sic] strength with me [sic] Parkinson's disease is not very good, so I'm not as strong" (P15, male, 76, Parkinsons)*
>
> *"not as much as I would have liked, but there again I know what's causing that and that's the osteoarthritis of my spine and spondylitis of my neck (P7, male, 72, obesity) [In reference to whether they had experienced improvements in mobility]*
>
> *"I just want to not necessarily improve myself, my life capabilities, my MS thing, but I just don't want to fall back down … I would love to do more but I know I can't" (P17, male, 67, MS)*

A further two participants had hoped to have more success with weight loss, but felt that medications such as insulin and steroids prevented this. Lastly, P10 perceived improvements in outcomes such as fitness and BP to be limited by older age, and described a desire to maintain their current health and fitness as opposed to improving it.

> *"Am [sic] doing it for weight loss as well which I'm not having as much success with because I'm diabetic and I take insulin, and you put weight on … so I'm maintaining the weight, but I want to actually to lose it" (P6, female, 58, obesity)*
>
> *"I mean I'm 74 years of age, and you do deteriorate there's no question about it, there's no way I'm going to have the health and fitness of a 30 year old, no*

> *way, you know? But at least I can hopefully sort of stabilise my fitness and try and keep it at a certain level" (P10, male, 74, diabetes)*

These findings also agree with previous literature, in that participants may feel restricted by health related symptoms, and may therefore feel restricted in the extent to which outcomes can be improved (Mills et al., 2012; Tobi et al., 2012).

The findings of this theme illustrate how qualitative research can capture participant experiences, which provide context to quantitative findings. Data from study 2 on PA level and BMI, for example, do not provide context about restrictions that prevented these outcomes from being improved further. Limitations such as the impact of medication on change in weight loss, for example, have also highlighted limitations of previous quantitative research, which have not considered these factors when evaluating outcomes. Aspects of participant journeys through an ERS, such as health related limitations, have rarely been captured or considered by previous research. These findings do however add support to a study by Mills et al. (2012), which found perceived restrictions in the extent to which outcomes can be improved.

7.4.3.8. General feedback

Beneficial aspects of the scheme were also captured through the theme of 'general feedback'. Thirteen participants made reference to the helpful and supportive nature of the ERS staff. Nine made general comments about how participation in the scheme had been beneficial for them. Seven made reference to how they had enjoyed participating in the scheme and activities, and three perceived the leisure centres and PA classes to have a welcoming atmosphere.

> *"they're very supportive, they're very helpful and they go out of their way to help you" (P3, female, 41, fibromyalgia)*
>
> *"it's been everything I've wanted and more really, I didn't expect to get as*

> *much out of it as I have" (P6, female, 58, obesity)*
>
> *"as I say the last two weeks I've been at my son's … you miss it, you really look forwards to getting back, it wasn't a bind coming today you know I didn't think 'oh', I was really looking forwards to it" (P28, female, 67, impaired glucose tolerance)*

Only two negative feedback comments were made by younger aged participants (44, 48 years). These were in relation to cost, because participants believed that the scheme should have been provided free of charge by the NHS. P29 (female, 48, MSK) elaborated further that healthcare professionals should prescribe free PA instead of medications, because of the holistic benefits of PA.

> *"some of it I think should be provided maybe for free, to help people like me, because otherwise they're not gonna [sic] do it … prescribe that [ERS] so instead of taking all the meds, give them the exercise and stuff that they need, they'll kill two birds with one stone, you know your health, your social aspects, everything … instead of the medical profession prescribing drugs, prescribe this" (P29, female, 48, MSK)*

In general, the findings of this theme support the outcomes of the feedback questionnaire, which demonstrated high levels of satisfaction with the ERS, including enjoyment and helpfulness of staff. They also provide further context to the experiences of participating in the scheme. The most frequent sub-theme, the helpful and supportive nature of the ERS staff, has been highlighted by previous qualitative research as an important determinant of ERS adherence (Morgan et al., 2016). The findings also highlight how cost of participation can be a barrier to participation (Morgan et al., 2016), which could potentially create an inequality in access amongst socio-economically deprived groups.

7.5. Strengths and limitations of study 3

One of the primary strengths of this study was the mixed methods design. Using quantitative data from the feedback questionnaire alongside qualitative data from semi-structured interviews was complementary and provided a more complete analysis of the benefits of participation (Creswell, 2003; Creswell & Plano Clark, 2011). Both methods achieved relatively large sample sizes, and by using purposive sampling, the semi-structured interview sample was similar to the demographic spread of the 'Live Active' population. Triangulation of data enhanced the trustworthiness of the findings (Lincoln & Guba, 1985; Tracy, 2010). Additional steps were also taken to enhance the trustworthiness of the findings in this study, such as: prolonged engagement with the ERS, to become familiar with the research setting and context (Lincoln & Guba, 1985; Shenton, 2004); providing a thick description of the research setting, research approach, and the context of participants' experiences (Shenton, 2004; Tracy, 2010); peer debriefing between researcher and supervisors (Lincoln & Guba, 1985; Shenton, 2004); tracing interpretations back to source quotes (Lincoln & Guba, 1985), and; exploring congruence of findings with previous literature (Silverman, 2001).

One of the limitations of this study was that the 'Live Active' feedback questionnaire included tick-box options for a variety of potential outcomes, which may have encouraged the reporting of these particular benefits (Choi & Pak, 2005). Furthermore, the success of the interview depended upon the extent to which the participant voiced their true opinions (Denscombe, 2007; Sparkes & Smith, 2014). It may be possible that participants voiced opinions that reflected the scheme, or themselves, in a more positive light. However, this is seen as unlikely, as participants did openly express disappointment with aspects of the scheme and/or their own participation. Furthermore, participants may not have been willing to disclose information relevant to the questions if they felt it was sensitive in nature, such as discussion of mental health issues.

7.6. Study 3 Summary and conclusions

Positivist methods such as RCTs, which have dominated the field of ERS evaluation, cannot embrace the complexity of real world interventions such as ERSs, nor do they capture the holistic worth of schemes beyond traditional markers of health (Crone et al., 2005; Dugdill et al., 2005; Mills et al., 2012; Riddoch et al., 1998). Therefore, for more than a decade research has recommended that future evaluations take a pluralistic approach to evaluation to provide a more representative evaluation of a scheme's impact (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998). By taking a pluralistic approach in this study, a wide range of benefits to participation have been uncovered that collectively demonstrate a holistic benefit of this ERS. Therefore, evaluating the impact of this scheme based solely on the quantitative outcomes analysed in study 2 would have underestimated the effectiveness of the intervention (Wilson, 2015).

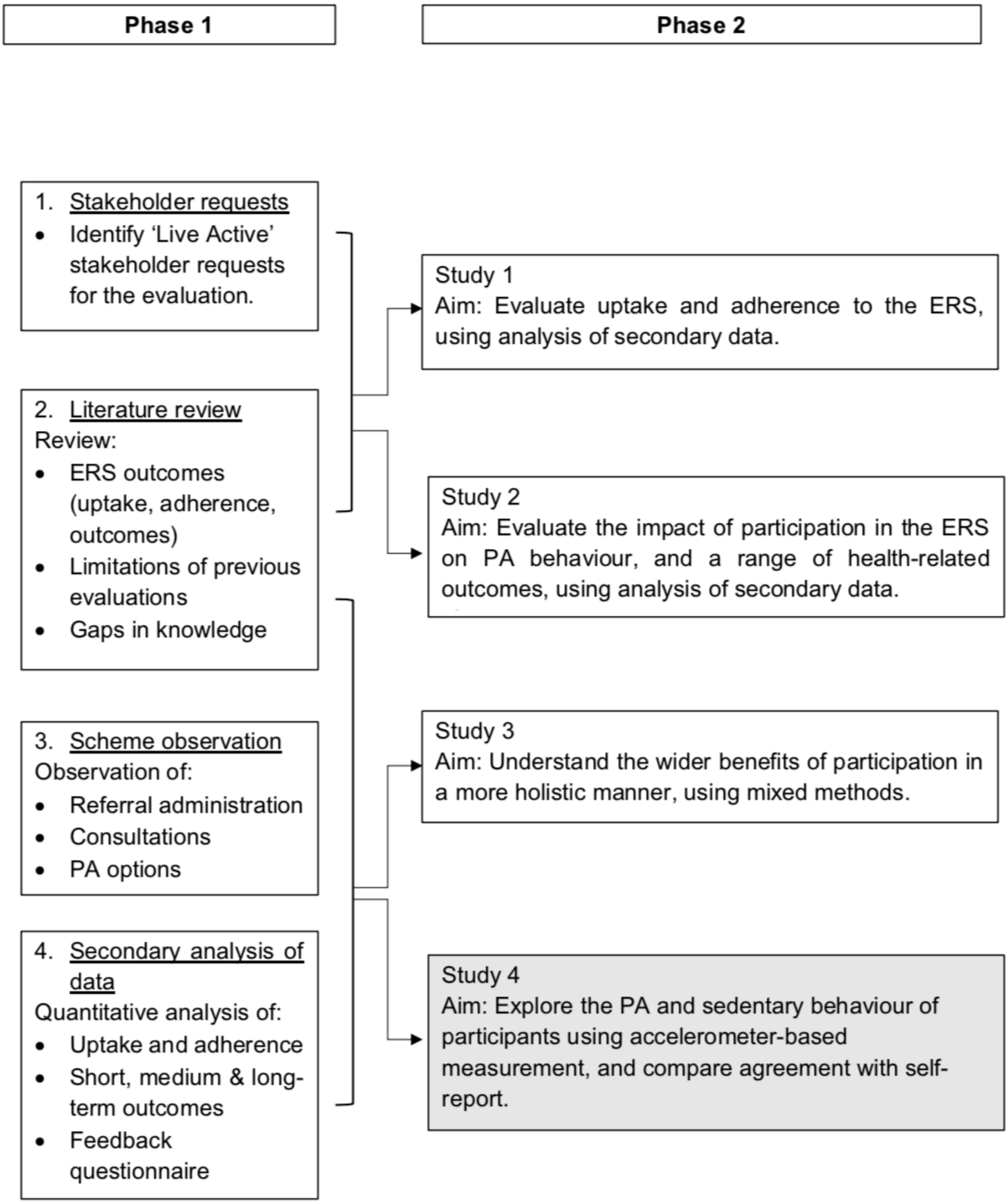
The results of the feedback questionnaire identified improved fitness, decreased medication use, improved sense of wellbeing and confidence, weight loss and socialisation as frequent benefits of participation. These were also frequent sub-themes of the semi-structured interviews, with the exception of 'sense of wellbeing'. However, participants also reported a broad range of additional benefits, demonstrating the multitude of personally meaningful ways in which one can benefit. An important observation is that participants reported a holistic benefit of participation. Rarely did they report solely physical health benefits, or adjunct benefits for example, but a range of benefits across the themes. The broad range of benefits reported may be explained by diverse populations that schemes such as this engage with, in which many of the participants may have a range of not just health, but complex emotional and social problems, and therefore have many potential ways of benefitting from participation (Wormald et al., 2006). This was also illustrated by the observation that reporting of the majority of outcomes was not influenced by age, gender, or health condition, but rather outcomes were reported across these categories and were often related to those that were personally meaningful to the participant.

Collectively, these findings demonstrate that participation in this ERS met the aims outlined by NICE (2014a, p. 7) that ERSs should "increase someone's physical activity levels on the basis that physical activity has a range of positive health benefits". Many of the benefits reported from the two data sources, such as social or adjunct benefits, would not be considered relevant health improving benefits by previous research in this field, traditionally concerned with change in physiological markers of health (Dugdill et al., 2005; Mills et al., 2012). However, evidence does suggest that these are outcomes associated with improved health and in some instances economic savings. Withdrawal of medications, for example, has obvious economic savings to the participant and healthcare system. The ability to carry out activities of daily living signifies that individuals can remain independent in their own homes for longer without external support (Whitehead et al., 2015). Improved social health is associated with a lower mortality risk (Holt-Lunstad et al., 2015), and lower rates of health care utilisation and therefore economic expenditure (Valtorta & Hanratty, 2012). Improved mental health is also associated with a lower mortality risk (Walker et al., 2015), and the total economic costs of mental health issues in the UK amount to >£94 billion per year (Organisation for Economic Co-operation & Development, 2018). Many of these outcomes are those which the Department of Health & Social Care, and the NHS, are actively seeking to improve, such as reduced medication expenditure (Ewbank et al., 2018), improved social health (Department for Digital, Culture, Media & Sport, 2018), and improved mental health (Department of Health & Social Care, 2018a; NHS, 2016). Considering the impact on a variety of outcomes spanning physical, mental and social health, where there is an imperative to urgently improve outcomes and reduce expenditure, these findings highlight the value of ERSs as a form of non-medical intervention, or 'social prescription' (Thomson et al., 2015).

The positive benefits reported in this study, and previous qualitative studies (e.g. Crone et al., 2005; Queen et al., 2015; Wormald et al., 2006), provide a compelling narrative for the impact of ERSs. Despite this, many of these benefits, such as social health or healthcare utilisation, have rarely been included in quantitative ERS evaluations (Dugdill et al., 2005), nor have evaluators been directed to consider these potential outcomes by evaluation frameworks (e.g. The SEF - Cavill et al., 2012). This may reflect the predominance of positivist research methods that do not value participant's experiences (Crone et al., 2005; Dugdill et al., 2005; Riddoch et al., 1998), or a disconnect in expectations from stakeholders such as commissioners and researchers, to that which is realised on the ground by participants (Henderson et al., 2018). The findings of this study therefore support arguments for more pluralistic evaluations of ERSs (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008; McNair et al., 2005; Riddoch et al., 1998), which consider what is meaningful to not only stakeholders such as researchers and commissioners, but also the participants themselves. In doing so, a more balanced evaluation of a scheme's impact can be conducted, which avoids underestimation of impact (Wilson, 2015). Knowledge of the wider benefits of ERSs reported in this study, and previous studies (e.g. Queen et al., 2015; Wormald et al., 2006), can also support the development of a range of outcome indicators for inclusion in future evaluations that include quantitative components (Mills et al., 2012; NICE, 2014a), and facilitate the development of evaluation frameworks that capture the holistic impact of ERS participation.

External evaluation of schemes, however, is not a viable reality for all ERSs, due to barriers such as finance (Gidlow et al., 2008). Therefore, if ERSs are to develop an evaluation culture amongst their staff (Gidlow et al., 2008) in order to conduct their own on-going evaluations, then evaluation methods must be easy to understand and to adopt into existing workflows. It may therefore be beneficial for scheme developers to implement a feedback questionnaire such as the one used in this study, that allows for the selection of a range of holistic benefits and open text reporting, to capture the full impact of participation at scale.

To conclude, the results of this study, considered alongside the results of quantitative analyses in study 2, have demonstrated a successful impact of the 'Live Active' ERS. These findings indicate that, as a form of non-medical intervention, ERSs have the potential to address many of the burdens posed by LTCs, which the government and NHS is actively seeking to address. Further qualitative research is now required to understand the mechanisms that have led to this scheme's success. Similar work is also necessary in other ERSs identified as effective, to generate an understanding of the common principles behind successful ERSs, which can then be scaled across schemes.



Overview of the research framework (study 4)

Chapter 8 Study 4: Device-based measurement of physical activity and sedentary behaviour in a sample of 'Live Active' participants, and comparison with self-reported behaviour

8.1. Study 4 – Introduction, aims and objectives

Accurate assessment of PA behaviour is fundamental if it is to be the primary outcome of an intervention (Pettee-Gabriel et al., 2012). However, as a complex behaviour, obtaining detailed and accurate information about PA from an individual can be difficult (Trost & O'Neil, 2014). The responses cannot be quantified as easily as other health behaviours such as smoking (Trost & O'Neil, 2014). Historically, PA levels in ERS participants have been measured using self-report methods such as questionnaires (Campbell et al., 2015; Pavey et al., 2011). These methods are ideal for a scheme such as exercise referral, as they are a low-cost and time efficient solution to gaining relatively in-depth data from participants (Haskell, 2012; Strath et al., 2013). However, as a self-report instrument, they are also subject to potential bias, such as poor memory and recall of PA history, which may result in poor validity (Haskell, 2012; Helmerhorst et al., 2012; Strath et al., 2013). In light of the potential bias of self-report questionnaires, and that to date all identified measurement of habitual PA in ERS evaluations has been through self-report questionnaires (Campbell et al., 2015; Pavey et al., 2011), a growing number of researchers have recommended that future ERS evaluations also measure PA using device-based methods such as accelerometers (Campbell et al., 2015; Hanson et al., 2013; Isaacs et al., 2007; Pavey et al., 2011).

In phase 1 of this thesis, secondary analysis of data indicated untypically high PA levels of participants in the 'Live Active' scheme, which were subsequently confirmed in study 2 (see section 6.5.2). Consideration of this observation, alongside the knowledge that no identified studies had evaluated the PA levels of ERS participants using device-based measurement, led to the development of the aims and objectives of this study.

Study 4 aim: Explore the PA and sedentary behaviour of participants using accelerometer-based measurement, and compare agreement with self-report.

Study 4 objectives:

4.1. Evaluate accelerometer-measured PA and sedentary behaviour of participants in week 1, to determine patterns of behaviour upon entry to the scheme.

4.2. Evaluate accelerometer-measured changes in PA and sedentary behaviour between weeks 1 and 12.

4.3. Compare time spent sitting per day according to i) the IPAQ-SF and ii) the activPAL™ accelerometer, at weeks 1 and 12.

4.4. Compare time spent in light and moderate-vigorous intensity PA lasting 10 minutes or more according to i) the IPAQ-SF and ii) the activPAL™ accelerometer, at weeks 1 and 12.

8.2. Literature review of physical activity measurement in ERSs

8.2.1. Physical activity measurement

Within the wider field of PA research, methods of PA measurement have been critically discussed for over a decade (Ainsworth, Cahalin, Buman, & Ross, 2015; Melanson, Freedson, & Blair, 1996; Warren et al., 2010). However, within the ERS literature or related policy, there has been little reference to this discussion, and no critical discussion of how PA behaviour should be assessed within ERSs, despite it being the primary outcome of schemes.

The convenience of self-report questionnaires, and availability as secondary data, has likely led to their predominance in ERSs and evaluations (Haskell, 2012; Strath et al., 2013). However, lack of critical discussion and guidance may have led to the variation in the PA questionnaires used by different ERSs, and therefore differences in how PA level is reported (e.g. MET mins of total PA, total minutes of PA, proportion meeting PA guidelines), making direct comparison between ERSs difficult (Campbell et al., 2015; Pavey et al., 2011).

In any PA intervention, the considered evaluation and selection of a PA measurement tool is paramount. PA measurement tools are varied and can be highly erroneous (Ainsworth et al., 2012), especially when they are used in populations for which the tool has not been designed, or not found to be valid and reliable (Ainsworth et al., 2012, 2015). The following section will discuss the advantages and limitations of PA questionnaires, from which the ERS evidence base is built upon.

8.2.2. Physical activity questionnaires

A variety of self-report PA instruments are available such as PA diaries and PA recall questionnaires (Ainsworth et al., 2015). As all of the identied ERS studies in the literature review measured PA through questionnaires (Campbell et al., 2015; Pavey et al., 2011), the following discussion will focus on PA questionnaires. These questionnaires typically require an individual to recall details about any PA performed during a specific time period, such as the frequency, intensity, duration and type of PA performed over the previous seven days (Ainsworth et al., 2015). Using this information, individuals can be classified in to groups such as inactive, moderately active and highly active (Strath et al., 2013; Warren et al., 2010).

PA questionnaires have been used in research for more than 50 years, including both surveillance and behaviour change intervention research (Ainsworth et al., 2015; Haskell, 2012). They are a relatively simple, cost-effective and time-efficient way to collect PA data from a large sample of individuals, in a manner that is well accepted and of low burden to individuals (Kowalski, Rhodes, Naylor, Tuokko, & MacDonald, 2012; Strath et al., 2013; Warren et al., 2010). Numerous questionnaires exist, many of which can be self or interviewer administered according to context of the study (Warren et al., 2010). Furthermore, several of the available PA questionnaires provide more information than could be gained from device-based measurement, such as the type of PA performed (Ainsworth, 2009).

Despite these benefits, PA questionnaires also have limitations that can impact precision of measurement (Ainsworth et al., 2015). When seeking to understand the accuracy of PA questionnaires, one must consider the reliability and validity of these instruments (Warren et al., 2010). Reliability refers to the extent to which an instrument gives consistent and repeatable measurements (Kelly, Fitzsimons, & Baker, 2016; Warren et al., 2010). Validity refers to the extent to which an instrument measures what it intends to (Warren et al., 2010), providing an exact representation of what happened (Kelly et al., 2016).

The PA questionnaire used in the 'Live Active' ERS, the IPAQ (IPAQ, 2005), is the most frequently used PA questionnaire worldwide (Lee et al., 2011; Silsbury et al., 2015; van Poppel et al., 2010). Systematic reviews have reported good reliability for the IPAQ, but mixed validity results when compared with instruments such as accelerometers (Cerin et al., 2016; Craig et al., 2003; Steene-Johannessen et al., 2016; van der Ploeg et al., 2010). There is a general consensus that individuals over-report time spent in moderate or vigorous intensity PA, and under-report time spent sedentary or in light intensity PA (Cerin et al., 2016; Chastin et al., 2018; Craig et al., 2003; Rosenberg et al., 2008; Steene-Johannessen et al., 2016; van der Ploeg et al., 2010). These measures may therefore provide an estimate of perceived PA, but not actual PA behaviour (Pettee-Gabriel et al., 2012). This has led researchers to recommend that PA derived from self-report PA questionnaires should be interpreted with caution (Steene-Johannessen et al., 2016), and referred to as self-reported PA to reflect that it may not be accurate (Troiano, 2009).

Recalling information such as PA history is a complex cognitive process (Baranowski, 1988) involving at least four steps: comprehension of the question asked; recall of the required information from memory; judgement of the link between question asked and answer retrieved, and; communication of the answer (Tourangeau, 1984). Bias in response can occur at any of these stages (Tourangeau, 1984), such as: poor memory and recall of information (Altschuler et al., 2009; Heesch et al., 2010); misinterpretation of terminology such as PA intensity (Altschuler et al., 2009; Finger et al., 2015; Sallis & Saelens, 2000; Scholes et al., 2016); difficulty classifying the intensity of non-aerobic activities such as resistance exercise (Pettee-Gabriel et al., 2012), and; social desirability (Adams et al., 2005; Motl et al., 2005).

It is known, for example, that adults can experience problems recalling the correct frequency, intensity, and duration of PA (Altschuler et al., 2009; Heesch et al., 2010). Previous research suggests that adults experience particular difficulty recalling unstructured activities (e.g. household related PA), and light intensity activities, leading to an under-reporting of these activities (Maddison et al., 2007; Sallis & Saelens, 2000).

However, previous research also suggests that adults can face difficulty in understanding the concepts of ‘light’, ‘moderate’ and ‘vigorous’ intensity PA, and classifying these activities correctly (Altschuler et al., 2009; Finger et al., 2015; Sallis & Saelens, 2000; Scholes et al., 2016). Therefore, individuals may recall a bout of PA, but misinterpret and report the intensity of the bout incorrectly. Light intensity PA may be reported as moderate intensity, and moderate intensity PA reported as vigorous intensity, resulting in the over-reporting of moderate and vigorous intensity

PA, and under-reporting of light intensity PA (Altschuler et al., 2009; Finger et al., 2015; Sallis & Saelens, 2000; Scholes et al., 2016).

Finally, social desirability refers to the tendency for individuals to present the most favourable image of themselves, which they believe to be more socially acceptable than their actual behaviour (Bowling, 2005). Therefore, an individual understands the question and retrieves the truthful answer, but purposefully adjusts the answer to report a more socially desirable response (Tourangeau, Rips, & Rasinski, 2000). Evidence suggests that social desirability can contribute to the over-reporting of PA, and that this may be more evident in structured PA questionnaires, compared to simple questionnaires such as those that ask individuals to quantify total levels of PA (Adams et al., 2005). However, not all studies have identified an impact of social desirability on PA reporting (Motl et al., 2005).

The following section will discuss the advantages and limitations of device-based measures of PA, which researchers have recommended are included in future ERS evaluations as an alternative or supplement to PA questionnaires (Campbell et al., 2015; Hanson et al., 2013; Isaacs et al., 2007; Pavey et al., 2011).

8.2.3. Device-based measurement of physical activity

Devices such as pedometers and accelerometers, often referred to in the literature as ‘objective measures’, assess one or more dimensions of PA, such as the frequency, intensity, duration or type of PA (Strath et al., 2013). The type of data available from these devices is varied, but often includes metrics such as step count, minutes of PA, and intensity of PA (Strath et al., 2013). Over the last decade there has been a significant increase in the proportion of studies using device-based measures to evaluate PA interventions among free-living adults, rising from 4.4% in 2006 to 70.6% in 2016 (Silfee et al., 2018).

Accelerometers, which are worn on the body, measure PA and sedentary behaviour by quantifying the acceleration and inclination resulting from movement (Chen, Janz, Zhu, & Brychta, 2012). The primary advantage of accelerometers is the precise manner in which they can quantify the dimensions of PA (frequency, intensity, time, type) over extended periods of time in free-living conditions (Butte, Ekelund, & Westerterp, 2012; Hills et al., 2014; Strath et al., 2013). Furthermore, accelerometers can quantify light intensity PA and sedentary behaviour (Lee & Shiroma, 2014), two components that individuals often underestimate by self-report (Cerin et al., 2016; Chastin et al., 2018; Steene-Johannessen et al., 2016). Researchers have therefore argued that accelerometers can quantify PA and sedentary behaviour with greater accuracy than self-report questionnaires (Hills et al., 2014; Reilly et al., 2008; Trost & O'Neil, 2014). This is supported by a review finding less variability across studies in the reliability and validity of accelerometers compared to PA questionnaires (Dowd et al., 2018). These advantages, alongside improvements in technology, cost and user acceptance have contributed to the recent shift towards prioritisation of device-based measurement in PA interventions (Silfee et al., 2018).

In a review of objective monitoring of PA behaviour, Freedson and Miller (2000) concluded that objective measures such as pedometers and accelerometers “certainly eliminate subjectivity of obtaining physical activity information” (p. 27-28). However, it is known that researchers and participants can influence accelerometer data collection, and subsequent analyses, whether consciously or inadvertently (Pedišić & Bauman, 2015). This introduces a degree of non-objectivity to accelerometers (Pedišić & Bauman, 2015), rejecting the notion that data from these devices are superior because they are objective measures (Haskell, 2012).

Researchers, for example, must make a number of subjective decisions about research and data analysis protocols when using accelerometers. When desiging a research protocol they must decide what an appropriate wear period is (number of days), whether the monitor should be worn continuously or during waking hours only, and when to exclude participants based on minimum acceptable wear time (Edwardson et al., 2017). During the data processing and analysis, reseachers must also decide how to identify wake and sleep times, and non-wear times (Edwardson et al., 2017). However, perhaps the most important decision, is how researchers choose to define ‘cut-off points’ that are used to determine the time spent in light, moderate and vigorous intensity PA (Troiano, McClain, Brychta, & Chen, 2014). For accelerometers that determine intensity by cadence, ≥100 steps/minute has been associated with moderate intensity PA (≥3 METs) (Tudor-Locke et al., 2018). Therefore, the findings of PA measurement research have a high dependence on these cut-off points, if the aim is to monitor time spent in PA of various intensities (Pedišić & Bauman, 2015). There has been a growing critical discussion of cut-off point thresholds, which vary between studies, and have not reached an agreed consensus (Pedišić & Bauman, 2015). There is a concern that some cut-off points may misclassify the intensity of PA (Trost, 2007; Trost & O’Neil, 2014), and may not be representative of the same PA intensity across individuals (Hills et al., 2014).

Participants can also introduce non-objectivity to accelerometers if they are able to influence the data collection (Pedišić & Bauman, 2015). This includes intentional non-wear time, whether temporarily or permanently, reducing wear time and therefore measurement of behaviours (Pedišić & Bauman, 2015). Although not a consistently reported problem, participants may shake or move the device to falsely increase PA counts, or may intentionally or unintentionally increase habitual PA as a result of being monitored (Hills et al., 2014; Pedišić & Bauman, 2015). Devices that do not provide feedback to the participant can reduce the likelihood of this interference (Hills et al., 2014).

There are also some generalised limitations of accelerometers that may impact the assessment of PA and sedentary behaviour, often due to the location at which accelerometers are worn (usually the lower body), and their inability to detect changes in load (Lee & Shiroma, 2014). Most accelerometers, for example, are unable to accurately measure upper-body or ambulatory activities such as weight lifting and cycling, and often do not account for the increased energy cost associated with activities such as climbing stairs or carrying heavy items. Therefore, accelerometers can underestimate total PA and energy expenditure (Pedišić & Bauman, 2015; Troiano et al., 2014; Trost, 2007; Trost & O'Neil, 2014).

Despite these limitations, researchers have argued that, in comparison with self-report questionnaires, accelerometers provide a more accurate assessment of PA and sedentary behaviour (Hills et al., 2014; Reilly et al., 2008; Trost & O'Neil, 2014). The activPAL™, the accelerometer used in this study, is considered to be the most accurate accelerometer for the assessment of sedentary behaviour (Grant, Ryan, Tigbe, & Granat, 2006; Kim, Barry, & Kang, 2015; Kozey-Keadle, Libertine, Lyden, Staudenmayer, & Freedson, 2011). It has also been validated for the classification of upright, standing and walking activities in a variety of populations, in free-living conditions (Dahlgren, Carlsson, Moorhead, Häger-Ross, & McDonough, 2010; Dowd, Harrington, & Donnelly, 2012; Grant, Dall, Mitchell, & Granat, 2008; Harrington, Welk, & Donnelly, 2011). However, in light of the limitations of accelerometers discussed in this section, it has been recommended that accelerometers should supplement rather than substitute PA questionnaires (Haskell, 2012; Pedišić & Bauman, 2015).

8.2.4. Device-based measurement in ERSs

To date, all ERS studies identified in the literature review have measured PA through self-report questionnaires, with no known studies quantifying habitual PA using devices such as accelerometers (Campbell et al., 2015; Pavey et al., 2011, excluding

Giné-Garriga et al., 2017 published as a trial protocol). The sedentary behaviour of ERS participants has also been underexplored, through self-report or device-based methods (excluding Giné-Garriga et al., 2017, study in progress).

Three identified studies have used accelerometers within ERS research, two of which were not for the purpose of measuring habitual PA, and the third of which cannot be generalised to a UK ERS context. First, Hawkins, Oliver, Wyatt-Williams, Scale and van Woerden (2014) explored the feasibility of using accelerometers within an ERS as a motivational tool, but did not report any PA measurements, instead evaluating the usability and tolerance of accelerometers by participants. Second, Webb et al. (2016) only used accelerometers during ERS exercise sessions, therefore giving a measurement of PA done during exercise sessions as opposed to total weekly PA. Third, Gallegos-Carrillo et al. (2017) measured the habitual PA of ERS participants at three time points. However, this was an ERS based in Mexico that cannot be generalised to a UK ERS context (e.g. referral process and programme). The authors reported non-significant increases in accelerometer-measured PA compared to a comparison group, and over-reporting of PA. A further study by Gabrys, Michallik, Thiel, Vogt and Banzer (2013) measured total weekly PA using accelerometers. However, due to the characteristics of the intervention it cannot be considered to be an ERS as described, and has been excluded from previous systematic reviews of ERSs for this reason.

8.2.5. Study 4 - summary

In summary, questionnaires are a low-cost solution to gaining in-depth data from large samples, but device-based measures may provide more comprehensive and valid measurements, with the suggestion that both are captured where feasible to provide a complete measurement of PA (Haskell, 2012). In consideration that all measurement of habitual PA in identified ERS evaluations has been through self-report questionnaires, a growing number of researchers have recommended that future evaluations also measure PA using devices such as accelerometers (Campbell et al., 2015; Hanson et al., 2013; Isaacs et al., 2007; Pavey et al., 2011). Furthermore, the sedentary behaviour of ERS participants has also been underexplored, through self-report or device-based methods (Giné-Garriga et al., 2017).

8.3. Study 4 Methods

This section provides a description of the methods used to address the aims and objectives detailed in section 8.1.

8.3.1. Participant recruitment

Participants were recruited in their first week of participation in the scheme, during their initial consultation, using convenience sampling. Participants were informed of the research study by the exercise professional, and if they expressed an interest in participating, they were formally invited to participate in the study by the researcher. Eligible participants were those who were aged 18 years or over, and could mobilise and walk independently. Due to the geographical spread of Active Tameside leisure centres, participants were predominantly sampled from three leisure centres experiencing the highest number of initial consultations.

Convenience sampling is a form of non-probability sampling, and involves the recruitment of participants from a population based on attributes such as being easily accessible, readily available and willing to participate (Salkind, 2010; Sedgwick, 2013; Teddlie & Yu, 2007). The advantages are that compared to other strategies, it is easy to implement and less time-intensive (Bornstein, Jager, & Putnick, 2013; Gravetter & Forzano, 2011; Sim & Wright, 2000). However, convenience sampling is considered to be a weak form of sampling, as there is no randomisation in the sampling process, limiting the inferences that can be made from the data (Bornstein et al., 2013; Gravetter & Forzano, 2011; Weathington, Cunningham, & Pittenger, 2010). The predominant concern is that convenience sampling creates biased samples that are not representative of the study population. Therefore, the results of the convenience sample may differ significantly to those of the study population. If so, the results then cannot be generalised to the study population, lowering the external validity of the study (Bornstein et al., 2013; Gravetter & Forzano, 2011; Sedgwick, 2013).

However, in some circumstances more rigorous forms of sampling, such as random sampling, are not possible, yet the researcher can still make efforts to recruit a reasonably representative convenience sample (Gravetter & Forzano, 2011; Sim & Wright, 2000). Convenience sampling was a practical sampling technique for this study, since the frequency of initial consultations varied by leisure centre. The frequency of consultations could also vary within leisure centres, by day, week and month. It was therefore more time-efficient to recruit participants from leisure centres with the highest number of consultations on a given day. Participants may also have not attended their initial consultation, or may have had mobility problems that were not specified on their referral form. Therefore, barriers could have been faced when using alternative techniques such as random sampling, as the participant may have been absent or ineligible.

8.3.1.1. Sample size calculation

Using Cohen's (1992) estimation of effect sizes, for a significance test with alpha at 0.05, when the effect size is large (0.50), a sample size of 28 would be required. At the same alpha level, but when the effect size is medium (0.30), a sample size of 85 would be required. Previous studies that have found meaningful results when comparing the outcomes of the IPAQ and activPAL™ have used sample sizes within this range of 28 and 85 participants (e.g. Busschaert et al., 2015; Chastin, Culhane, & Dall, 2014; Fitzsimons, Kirk, Murphy, & Mutrie, 2012; Vasudevan et al., 2013). In consideration of power calculations and previous research, the target sample size was 57 participants (lower range 28, upper range 85), allowing for a dropout rate of 30%, which was comparable to the dropout rate of the ERS at the time. This target was within the range of previous research studies, and was reasonable considering the logistics of data collection, and the comprehensive data processing and analysis requirements of activPAL™ and IPAQ data.

8.3.2. Research protocol

Following the initial consultation, if participants agreed to participate in the study, they were provided with an activPAL™ (day 1), with data recording programmed to commence at midnight. They were asked to wear the activPAL™ continuously for the following seven consecutive days (days 2-8), removing only for water-based activities (bathing, swimming). Data recording terminated at midnight on the 9th day, and therefore participants were asked to remove the activPAL™ on the morning of the 9th day and return it to the researcher. Upon returning the activPAL™ on the 9th day, the participant completed the IPAQ-SF, to provide a self-report of PA levels over the previous seven days for comparison with the accelerometer-measured data collected over this period. During the participant’s week 12 consultation, they were again invited to wear the activPAL™ for seven days. If they agreed to participate, the same protocol was followed.

Demographic information about participants was routinely collected by the exercise professional during the first consultation, and stored on the ERS database. Demographic information for each participant was therefore exported from this database, with consent from the participant.

8.3.2.1. The IPAQ

The 7-day IPAQ-SF was the self-report PA questionnaire routinely used in the ‘Live Active’ ERS (appendix 4). It is a nine item form that records the frequency and duration of PA lasting ≥10 minutes, over three intensity levels: 1) vigorous intensity; 2) moderate intensity, and; 3) walking (over the previous seven days). The fourth question, sitting time, is a single item measure that asks participants to estimate for how long (in hours and minutes) they sit each day, on average.

8.3.2.2 The activPAL™

The activPAL™ is a small, light-weight (15 g) accelerometer-based device that is attached to the front midline anterior aspect of the thigh (PAL Technologies Ltd, Glasgow, Scotland). The right thigh was used in this study, as is best practice, with the exception of participants who had right leg movement limitations, or for who the right leg was irritated by the attachment (Edwardson et al., 2017).

Participants were provided with written instructions and a demonstration of how to position and attach the activPAL™ (appendix 11, adapted from Edwardson et al., (2017) supplementary materials). Participants were also provided with a paper diary to record sleep and wake times, to assist with isolating key periods of interest, such as waking hours. There was also a section to record time and reason for removal or re-attachment of the activPAL™, and other comments (appendix 11).

8.3.3. Data processing

8.3.3.1. IPAQ

Time spent in vigorous intensity PA, moderate intensity PA, and walking activities, per week, were retained as continuous variables (minutes). To calculate total PA per week, the time spent in these activities was summed. Time spent sitting per day was a single item measure, provided as a continuous figure (minutes per day).

8.3.3.2. activPAL™

Data from the activPAL™ were downloaded and processed using the manufacturer's software; activPAL™ v7.2.32 (PAL Technologies). Using proprietary algorithms, the software classified data from the activPAL™ into sedentary (sitting/lying), standing and stepping events (in 15 second epochs). This information was exported as a Microsoft Excel file, known as the 'events file', which detailed the date, time and duration of all sedentary, standing and stepping events, including the cadence of stepping events. The 'events file' was then processed using a MATLAB code, to group these events together based on whether they were sedentary, standing or stepping events. The Microsoft Excel file generated from this code allowed stepping events to be isolated and analysed based on duration of stepping bout, and/or intensity (cadence) of stepping bout. This allowed, for example, stepping events lasting ≥10 minutes at moderate-vigorous intensity to be isolated and summed.

All analyses were conducted on events occurring during waking hours. These hours were isolated by identifying, for each day, time awake as the time of the first standing event after a long continuous period of non-upright posture (usually between 12am-9am), and time asleep as the last standing event before a long continuous period in non-upright posture (usually after 9 pm). Analyses were therefore conducted on events occurring between time awake and time asleep, each day. Identification of time awake and time asleep was facilitated by the information provided in participants' written diaries. Non-wear time events identified from the written diary were removed from the analysis, such as removal for water-based activities.

To evaluate PA levels, duration of stepping events were summed across all waking hours according to bout duration (any duration, ≥10 minutes), intensity (any cadence, ≥100 steps per minute), or both (any duration and any cadence, any duration and ≥100 steps per minute, ≥10 minutes and any cadence, ≥10 minutes and ≥100 steps per minute). This allowed for the calculation of total PA, total low intensity PA, total MVPA, and PA occurring at each intensity level for ≥10 minutes, per week. Previous research has suggested that a cadence of ≥100 steps per minute is equivalent to moderate intensity PA (Tudor-Locke et al., 2018). Furthermore, stepping events lasting ≥10 minutes were identified for the purpose of comparison with data from the IPAQ-SF, which asks participants to report the frequency and duration of PA lasting ≥10 minutes.

To evaluate sedentary behaviour, the duration of all sedentary events during waking hours were summed per day (minutes). Average sitting time per day (minutes) was calculated, and reported, by dividing total sedentary time by total wear time (days).

Researchers have tended to use seven day wear time protocols when assessing PA and sedentary behaviour using accelerometers (Matthews, Hagströmer, Pober, & Bowles, 2012). However, there is not a general consensus that this wear time protocol is optimal, with some arguing that fewer, or more days are required (e.g. Bergman, 2018; Trost, McIver, & Pate, 2005). Furthermore, researchers cannot be assured that participants will comply with wear time protocols, in some cases providing fewer days of monitoring than requested (Sasaki et al., 2018). In a review of accelerometer measurement of PA and sedentary behaviour, Trost et al. (2005) suggested that 3-5 days of accelerometer monitoring are required to estimate habitual activity in adults. This minimum wear time period has also been suggested for the assessment of PA and sedentary behaviour in older adults (Hart, Swartz, Cashin, & Strath, 2011; Sasaki et al., 2018).

In consideration of this knowledge, the minimum acceptable wear time for the comparison of accelerometer-measured and self-reported PA and sedentary behaviour was seven days, due to the IPAQ being a seven day recall questionnaire. The minimum acceptable wear time for the assessment of short-term change in accelerometer-measured PA and sedentary behaviour was five days. Where participants had different wear times at weeks 1 and 12 (i.e. 6 days vs 7 days), wear time in days was matched (i.e. 6 days vs 6 days), utilising data from the same days of the week where possible.

8.3.4. Statistical analysis

Statistical analyses were performed using IBM SPSS statistics version 23.0 (SPSS Inc., Illinois, USA) and Microsoft Excel version 15.2. Sedentary and PA behaviour at weeks 1 and 12 were explored using descriptive statistics. Changes in these outcomes between weeks 1 and 12 were evaluated using paired t-tests.

Agreement between the IPAQ-SF and activPAL™ outcomes were assessed using Kendall's tau-b correlation and Bland-Altman analysis (Bland & Altman, 1986). Data were first checked for normality of differences and homoscedasticity. Bland-Altman plots were then produced, detailing the mean bias and limits of agreement, with 95% confidence intervals.

For the purpose of exploring agreement between the IPAQ-SF and activPAL™, and for comparison of the results with those of previous findings (study 2), or previous literature, all sedentary behaviour outcomes are referred to as average time spent sitting per day in minutes, and all PA outcomes are referred to as minutes of PA per week.

8.4. Study 4 - results

This section reports the results of data analyses, aiming to address the aims and objectives detailed in section 8.1. It begins by describing levels of sedentary behaviour (minutes per day) and PA (minutes per week) at week 1, and week 12, with a comparison of accelerometer-measured and self-reported sedentary behaviour and PA at both time points. It finishes by exploring the change in accelerometer-measured sedentary behaviour and PA between weeks 1 and 12.

8.4.1. Participants

In total, 38 participants agreed to take part in the study. Of these, 28 met the required minimum wear time during week 1 for a comparison of accelerometer-measured and self-reported outcomes (seven days). Due to a high level of dropout from the scheme (n=12) or the study (n=15), only 11 participants agreed to repeat the protocol at week 12. Of these, nine met the required minimum wear time for a comparison of accelerometer-measured and self-reported outcomes (seven days), and all 11 participants met the required minimum wear time for a comparison of accelerometer-measured PA and sedentary behaviour at weeks 1 and 12 (five days).

8.4.2. Sedentary behaviour at week 1

Participants included in the week 1 analysis (n=28) had a mean ±SD age of 58.9 ±13 years. Sixty percent were male (n=17), and the majority were not working due to retirement or other reasons (53.6%, n=15). A further 25% were employed full-time (n=7), 14.3% were unemployed (n=4), and 7.1% were employed part-time (n=2).

Considering accelerometer-measured sedentary behaviour, at week 1, during waking hours, participants (n=28) spent an average of 568.3 ±86.8 minutes per day sedentary (mean ±SD). When categorised into time bands, the results showed that the majority of participants spent between 8-12 hours per day sedentary (8-10 hours, n=14, 50%; 10-12 hours n=9, 32%). There were no significant differences in average sedentary time per day between males and females (593.5 ±62.9 vs 529.2 ±106, t(26) = 2.019, *p* >0.05), or occupational groups (*F*(3, 24) = 2.775, *p* >0.05). There was also no significant correlation with age (r=0.067, *p* >0.05).

When comparing average accelerometer-measured sedentary time per day (568.3 ±86.8mins per day) with average self-reported sedentary time per day (483.2 ±151.1

mins per day), Kendall's tau-b showed a significant correlation between the two measures (τ_b=0.278, p ≤0.05). The Bland-Altman plot showed that participants under reported sedentary time by an average of 85 minutes per day (Figure 8.1, below). Regression of the mean and difference indicated that error reduced as average sitting time per day increased (r=0.53, p ≤0.01). It is noted that participants appeared to estimate sitting time to the nearest hour per day, which may have contributed to the striations present in the plot. All but one of the differences fell within the limits of agreement (-203.8 – 373.9 mins per day). However, these limits of agreement were considerably wide, and therefore agreement between the two methods was not considered to be acceptable.

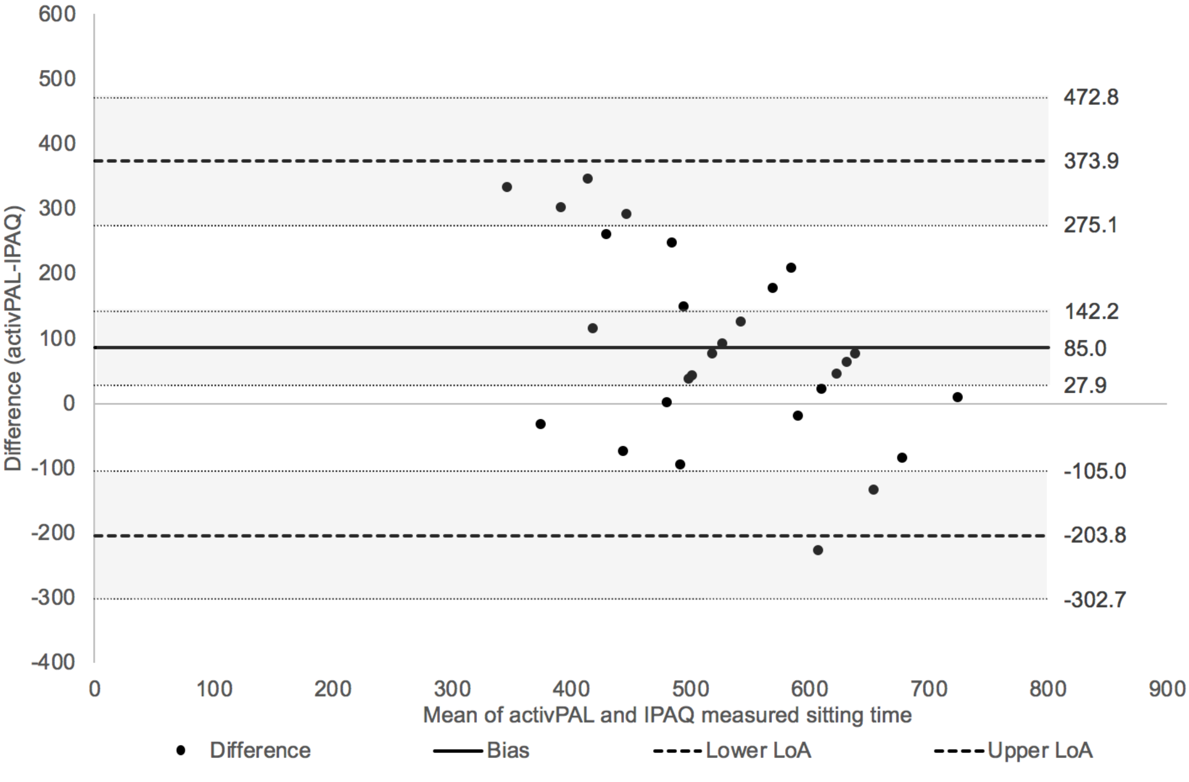


Figure 8.1. Bland-Altman plot of the difference in average sedentary time per day (mins) when measured using accelerometry or self-report at week 1. The mean bias and 95% limits of agreement are indicated, with respective 95% confidence intervals highlighted.

8.4.3. Physical activity at week 1

Considering accelerometer-measured PA, participants accumulated a mean of 601.2 ±258.2 minutes of stepping, across seven days, of which 154.1 ±158 minutes were at moderate-vigorous intensity (≥100 steps per min) (Table 8.1, below). When considering the duration and intensity of stepping bouts taken each day, on an average day, most stepping bouts were either short in duration, lasting 0-30 seconds (41.1%) or 30-60 seconds (14.6%), or longer in duration lasting ≥5 minutes (21.1%) (Figure 8.2). On an average day, stepping bouts were also mostly low intensity (64.4% between 20-100 steps per min), with 35.6% considered moderate-vigorous intensity (≥100 steps per min) (Figure 8.3). Therefore, PA tended to be accumulated throughout the day in short bouts of low intensity PA.

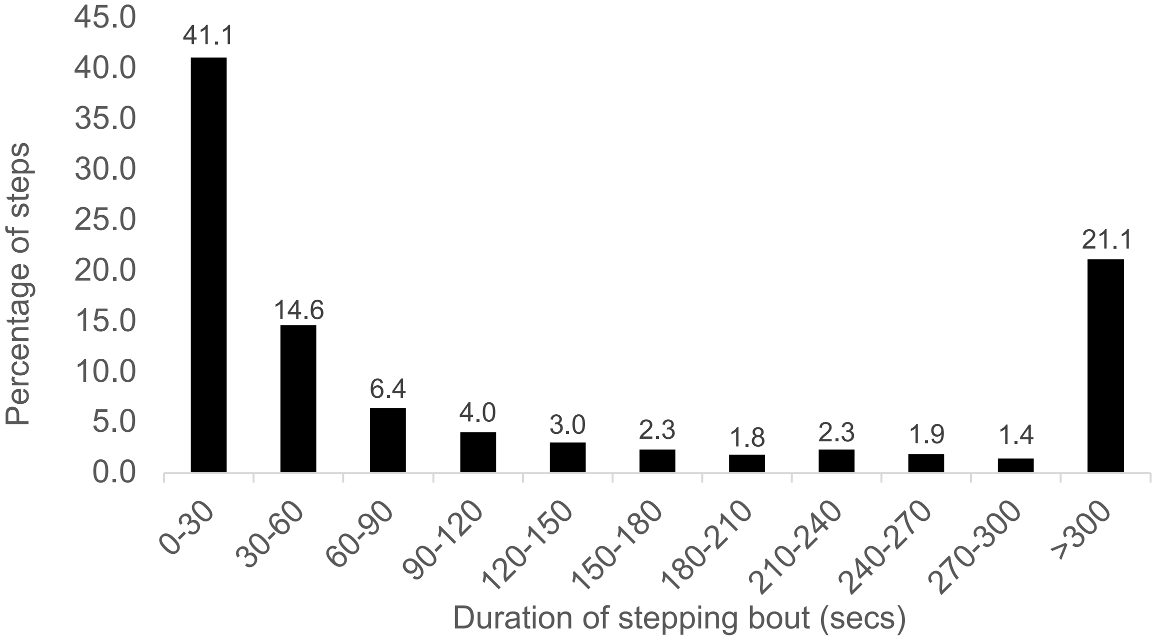


Figure 8.2. The percentage of steps that occurred in each bout length, on an average day, at week 1 (accelerometer-measured).

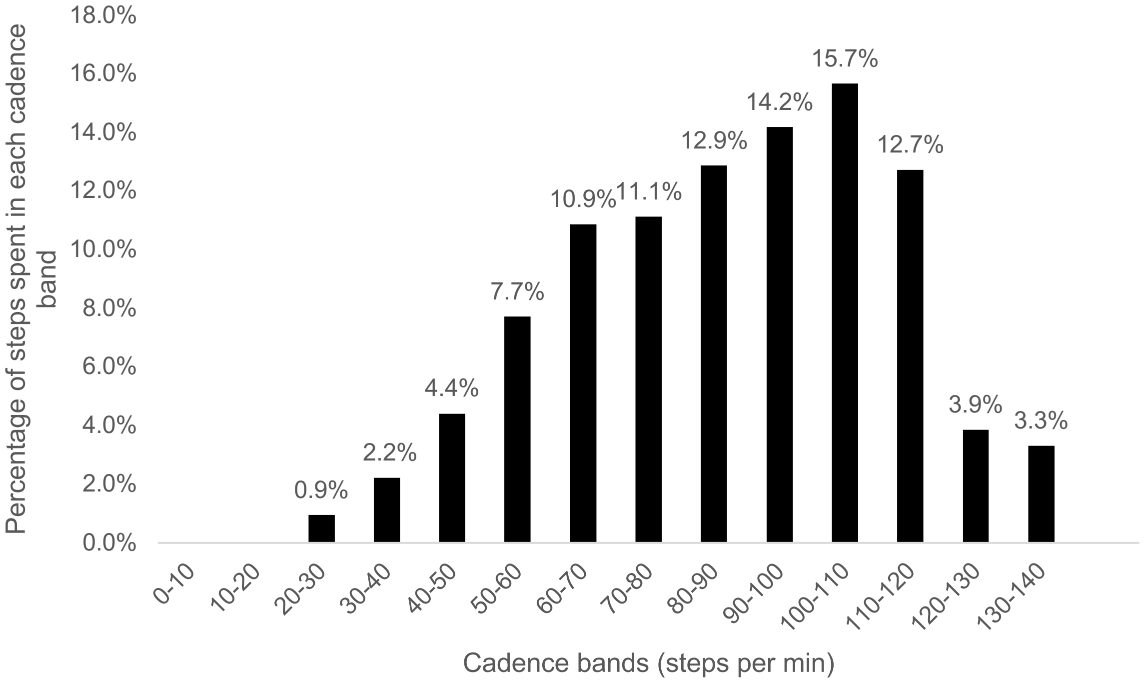


Figure 8.3. Intensity of stepping bouts at week 1, on an average day (accelerometer-measured).

The UK PA guidelines recommend that PA is accumulated in bouts lasting 10 minutes or more (DoH, 2011), therefore the duration of stepping bouts were further explored. During seven days of wear time, 18/28 participants (64%) achieved ≥1 stepping bouts lasting ≥10 minutes, accumulating a mean of 51.1 ±86.5 minutes per week in stepping bouts lasting ≥10 minutes. When considering both intensity and duration, 15/28 participants (54%) achieved ≥1 moderate-vigorous intensity stepping bouts lasting ≥10 minutes, accumulating a mean of 42.5 ±85.8 per week in these bouts (Table 8.1, below). Therefore, 10 participants did not achieve a single ≥10 minute bout of stepping during seven days of wear time.

Table 8.1. Accelerometer-measured mean minutes of PA (per week) accumulated at week 1, according to the duration and intensity of the PA bout.

	Mean minutes accumulated (per week)	SD
Total PA (any intensity, any duration)	601.2	258.2
Total PA (≥moderate intensity, any duration)	154.1	158
Bout duration ≥10 mins, any intensity	51.1	86.5
Bout duration ≥10 mins, ≥moderate intensity	42.5	85.5

When comparing acclerometer-measured PA with self-reported PA (Table 8.2), Kendall's tau-b showed a significant correlation for total PA (τ_b=0.399, *p* ≤0.01), and MVPA (τ_b=0.303, *p* ≤0.05), but not low intensity PA (τ_b=0.142, *p* >0.05). The Bland Altman plots showed that participants over-reported total PA by 186.9 minutes per week, low intensity PA by 77.5 minutes per week, and and MVPA by 109.5 minutes per week. After the removal of an extreme outlier, over-reporting of total PA reduced to 123.8 minutes per week (limits of agreement: -334 – 86.3 mins per week), over-reporting of low intensity PA increased slightly to 80.3 minutes per week (limits of agreement: -289.1 – 128.5 mins per week) (Figure 8.4, below), and over-reporting of MVPA reduced to 43.5 minutes per week (limits of agreement: -172.5 – 85.5 mins per week) (Figure 8.5, below). Regression of the means and differences indicated that error increased as PA increased (total PA r=0.49, *p* ≤0.01, low intensity PA r=0.88, *p* >0.05, MVPA r=0.20, *p* >0.05). The majority of the differences fell within the limits of agreement for each respective plot. However, these limits of agreement were considerably wide, and therefore agreement between the two methods was not considered to be acceptable.

Table 8.2. Accelerometer-measured and self-reported PA (mins per week), accumulated in bouts ≥10 minutes, at week 1 (outlier removed).

	Accelerometer-measured		Self-reported	
	Mean	SD	Mean	SD
Total PA	53	86	176.9	132.6
Total low intensity PA	8.9	27.1	89.3	104.8
Total MVPA	44.1	85.4	87.6	103

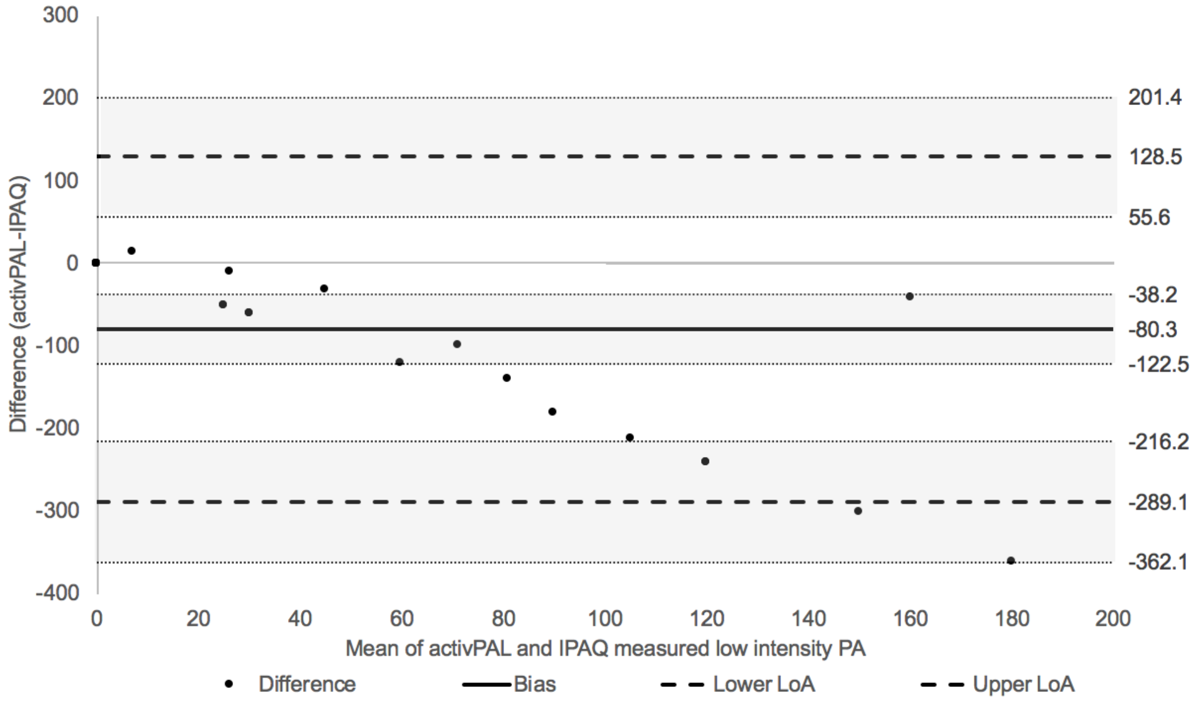


Figure 8.4. Bland-Altman plot of the difference in total low intensity PA (mins per week) when measured using accelerometry or self-report at week 1. The mean bias and 95% limits of agreement are indicated, with respective 95% confidence intervals highlighted.

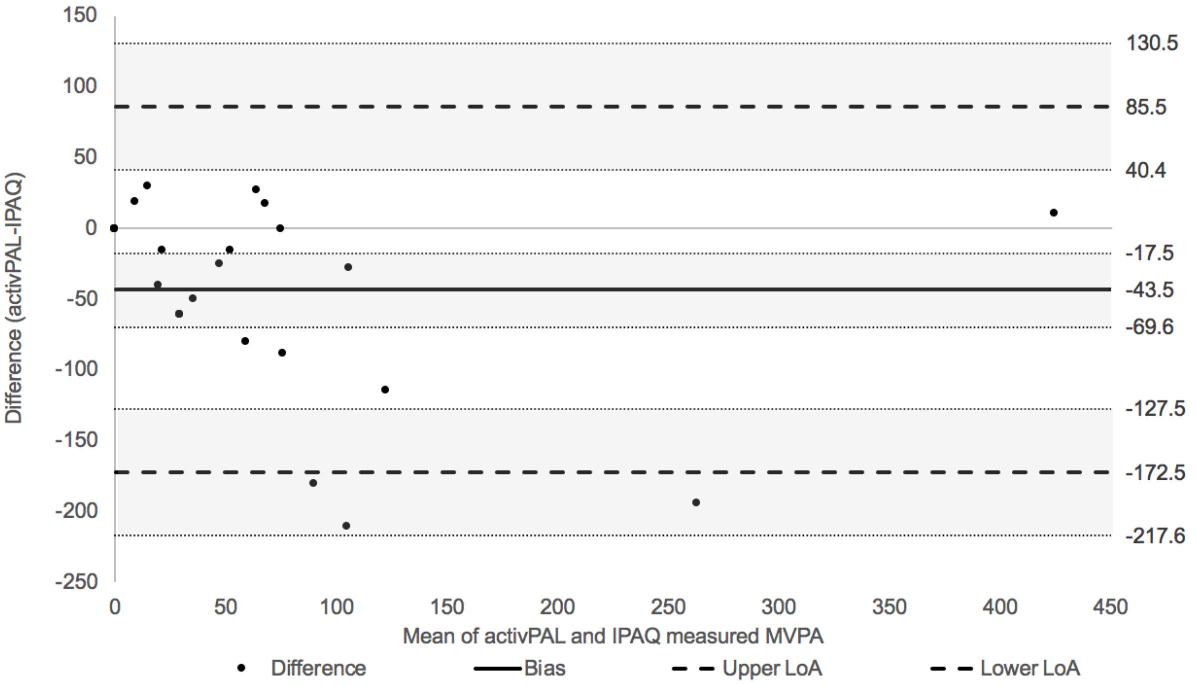


Figure 8.5. Bland-Altman plot of the difference in total MVPA (mins per week) when measured using accelerometry or self-report at week 1. The mean bias and 95% limits of agreement are indicated, with respective 95% confidence intervals highlighted.

8.4.4. Sedentary behaviour at week 12

Participants included in the week 12 analysis (n=9) had a mean ±SD age of 57.8 ±17 years. Sixty-seven percent were female (n=6), and the majority were not working due to retirement or other reasons (66.7%, n=6). A further 22% were employed full-time (n=2), and 11% were employed part-time (n=1).

Considering accelerometer-measured sedentary behaviour, at week 12, during waking hours, participants (n=9) spent an average of 520.1 ±89.2 minutes per day sedentary. When categorised in to time bands, the majority of participants spent between 6-10 hours per sedentary (6-8 hours, n=4, 44%; 8-10 hours, n=4, 44%). There were no significant differences in average sedentary time per day between males and females (545.3 ±42.8 vs 507.4 ±106.2, t(7) = 0.583, *p* >0.05), or occupational groups (*F*(2, 6) = 0.348, *p* >0.05). There was also no significant correlation with age (r=0.373, *p* >0.05).

When comparing average accelerometer-measured sedentary time per day (520.1 ±89.2 mins) with average self-reported sedentary time per day (360 ±140.7 mins), Kendall’s tau-b showed no significant correlation between the two measures (τ_b=-0.059, *p* >0.05). The Bland-Altman plot showed that participants under reported sedentary time by an average of 160.1 minutes per day (Figure 8.6, below). Regression of the mean and difference indicated that error reduced as average sitting time per day increased (r=0.53, *p* ≤0.01). All of the differences fell within the limits of agreement (-159.3 – 479.4 mins per day). However, these limits of agreement were considerably wide, and therefore agreement between the two methods was not considered to be acceptable.

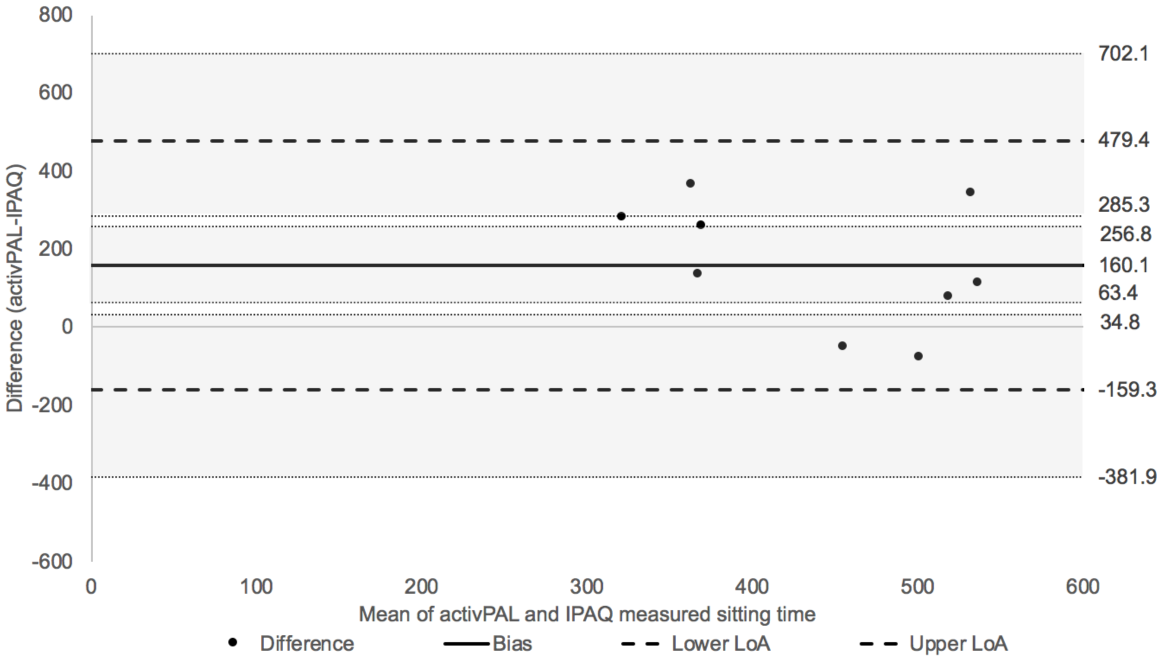


Figure 8.6. Bland-Altman plot of the difference in average sedentary time per day (mins) when measured using accelerometry or self-report at week 12. The mean bias and 95% limits of agreement are indicated, with respective 95% confidence intervals highlighted.

8.4.5. Physical activity at week 12

Considering accelerometer-measured PA, participants accumulated a mean of 684.6 ±170.4 minutes of stepping, across seven days, of which 189.7 ±103.2 minutes were at moderate-vigorous intensity (≥100 steps per min) (Table 8.3, below). When considering the duration and intensity of stepping bouts taken each day, on an average day, most stepping bouts were either short in duration, lasting 0-30 seconds (36.1%), or 30-60 seconds (12.9%), or longer in duration lasting ≥5 minutes (24%) (Figure 8.7). On an average day, stepping bouts were also mostly low intensity (62.7% between 20-100 steps per min), with 37.3% considered moderate-vigorous intensity (≥100 steps per min) (Figure 8.8). Therefore, PA tended to be accumulated throughout the day in short bouts of low intensity PA.

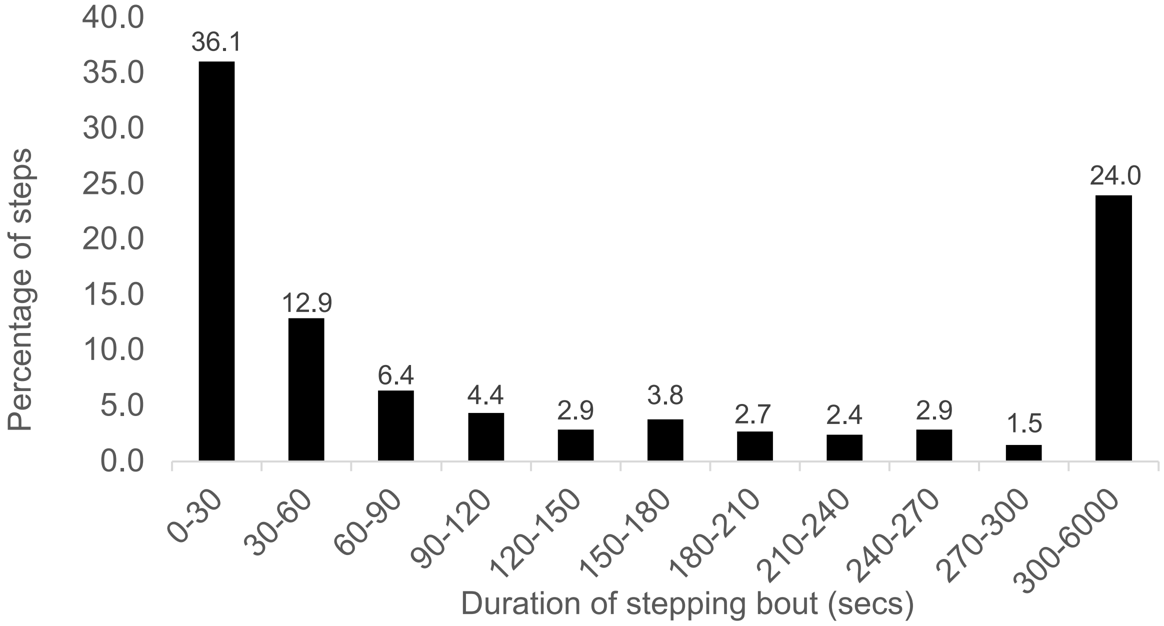


Figure 8.7. The percentage of total steps that occurred in each bout length, on an average day, at week 12 (accelerometer-measured).

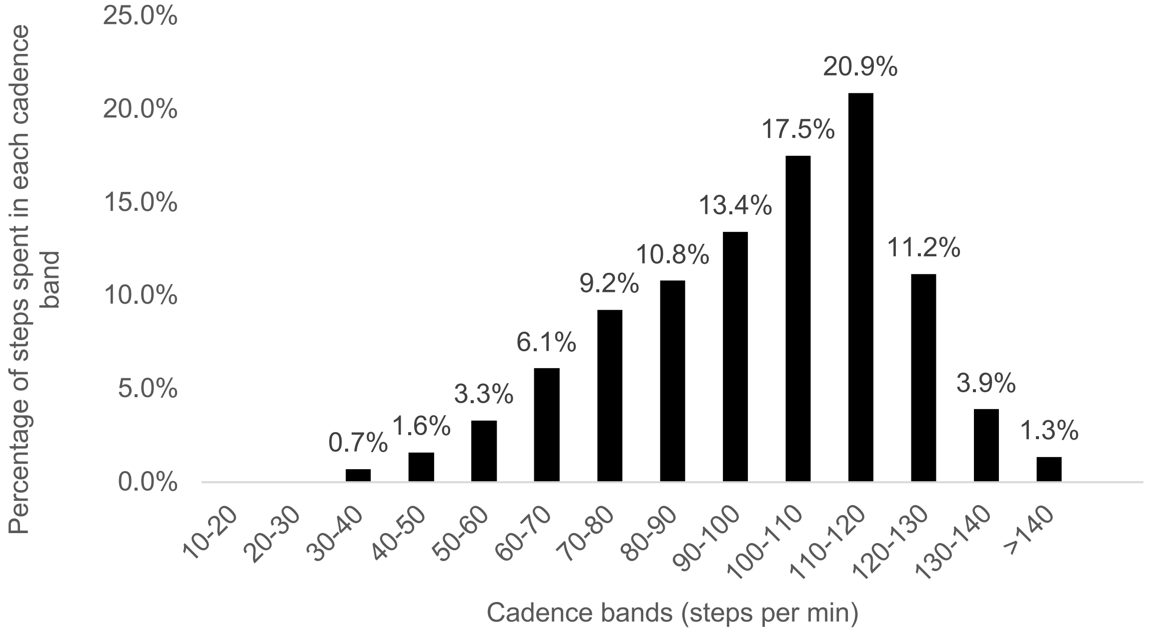


Figure 8.8. Intensity of stepping bouts at week 12, on an average day (accelerometer-measured).

When stepping bouts were further analysed taking duration of bout in to consideration, during seven days of wear time, 8/9 participants (88.9%) achieved ≥1 stepping bouts lasting ≥10 minutes, accumulating a mean of 63.1 ±45.7 minutes per week in stepping bouts lasting ≥10 minutes. When considering both intensity and duration, the same 8/9 participants achieved ≥1 moderate-vigorous intensity stepping bouts lasting ≥10 minutes, accumulating a mean of 52.3 ±36.3 minutes per week in these bouts (Table 8.3, below).

Table 8.3. Accelerometer-measured mean minutes of PA (per week) accumulated at week 12, according to the duration and intensity of the PA bout.

	Mean minutes accumulated	SD
Total PA (any intensity, any duration)	684.6	170.4
Total PA (≥moderate intensity, any duration)	189.7	103.2
Bout duration ≥10 mins, any intensity	63.1	45.7
Bout duration ≥10 mins, ≥moderate intensity	52.3	36.3

When comparing accelerometer-measured PA with self-reported PA (Table 8.4), Kendall's tau-b showed no significant correlations for total PA (τ_b=0.222, p >0.05), low intensity PA (τ_b=0.218, p >0.05), or MVPA (τ_b=-0.093, p >0.05). The Bland-Altman plots showed that participants over-reported total PA by 96.9 minutes per week (limits of agreement: -370.1 – 176.4 mins per week), total low intensity PA by 64.2 minutes per week (limits of agreement: -189.6 – 61.1 mins per week) (Figure 8.9, below), and total MVPA by 32.7 minutes per week (limits of agreement: -246.9 – 181.6 mins per week) (Figure 8.10, below). Regression of the means and differences indicated that error increased as PA increased (total PA r=0.86, p ≤0.01, low intensity PA r=0.77, p ≤0.05, MVPA r=0.79, p ≤0.05). The majority of the differences fell within the limits of agreement for each respective plot. However, these limits of agreement were considerably wide, and therefore agreement between the two methods was not considered to be acceptable.

Table 8.4. Accelerometer-measured and self-reported PA (mins per week), accumulated in bouts ≥10 minutes, at week 12.

	Accelerometer-measured		Self-reported	
	Mean	**SD**	**Mean**	**SD**
Total PA	63.1	45.7	160	152.1
Total low intensity PA	10.8	23.8	75	64.8
Total MVPA	52.3	36.3	85	105

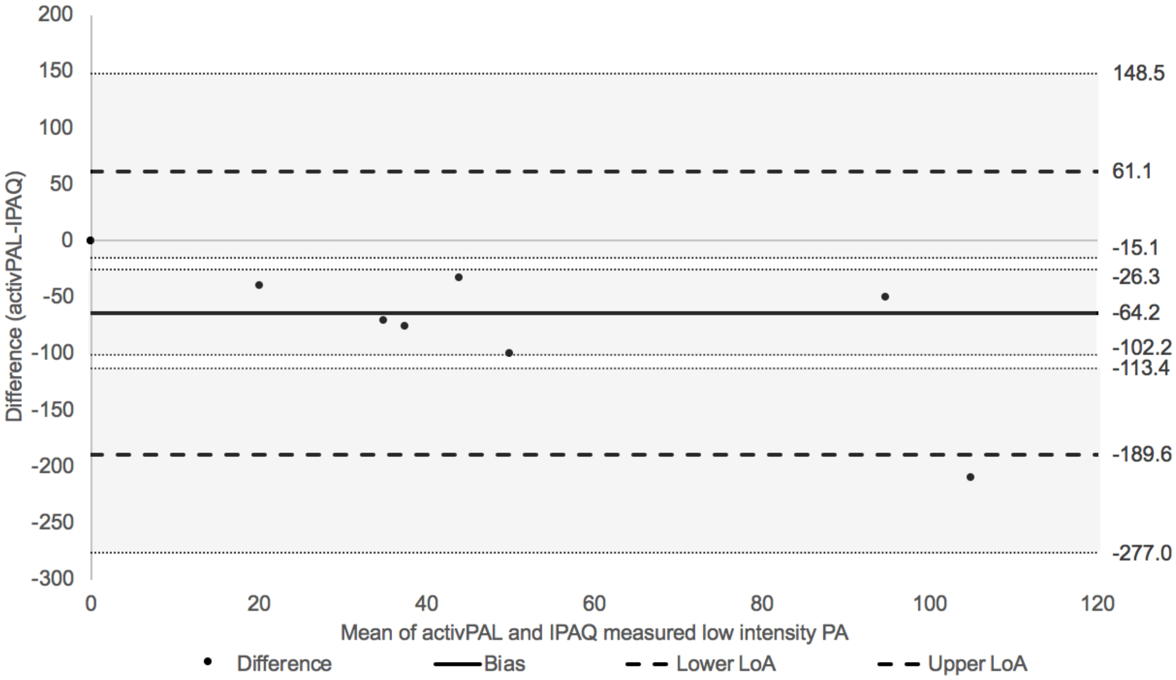


Figure 8.9. Bland-Altman plot of the difference in total low intensity PA (mins per week) when measured using accelerometry or self-report at week 12. The mean bias and 95% limits of agreement are indicated, with respective 95% confidence intervals highlighted.

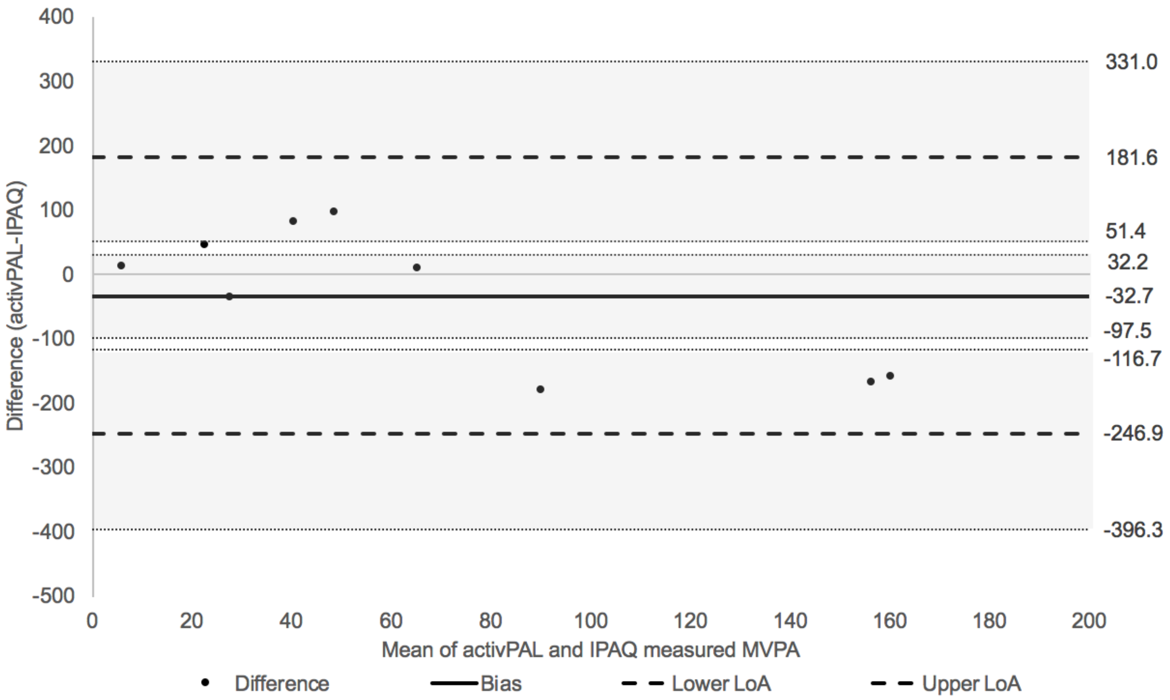


Figure 8.10. Bland-Altman plot of the difference in total MVPA (mins per week) when measured using accelerometry or self-report at week 12. The mean bias and 95% limits of agreement are indicated, with respective 95% confidence intervals highlighted.

8.4.6. Week 1 vs week 12 comparison

The following analyses compare the data of 11 participants who had data available at both weeks 1 and 12.

There was a non-significant decline in accelerometer-measured mean sedentary time per day (mins) between weeks 1 and 12 (553.6 ±84.8 vs 539.3 ±113.5, t(10) = 0.834, *p* >0.05). When exploring individual changes in sedentary behaviour, the data showed variation in change; eight participants reduced their mean sedentary time per day, whilst three participants increased sedentary time (Figure 8.11, below). At both time points, participants were mostly sedentary for 8-10 hours per day (54.5% vs 45.5%), with no significant difference in sedentary time between the two time points when grouped in to time bands (6-8 hours, 8-10 hours, 10-12 hours, >12 hours per day) ($\chi^2(6) = 10.3$, *p* >0.05).

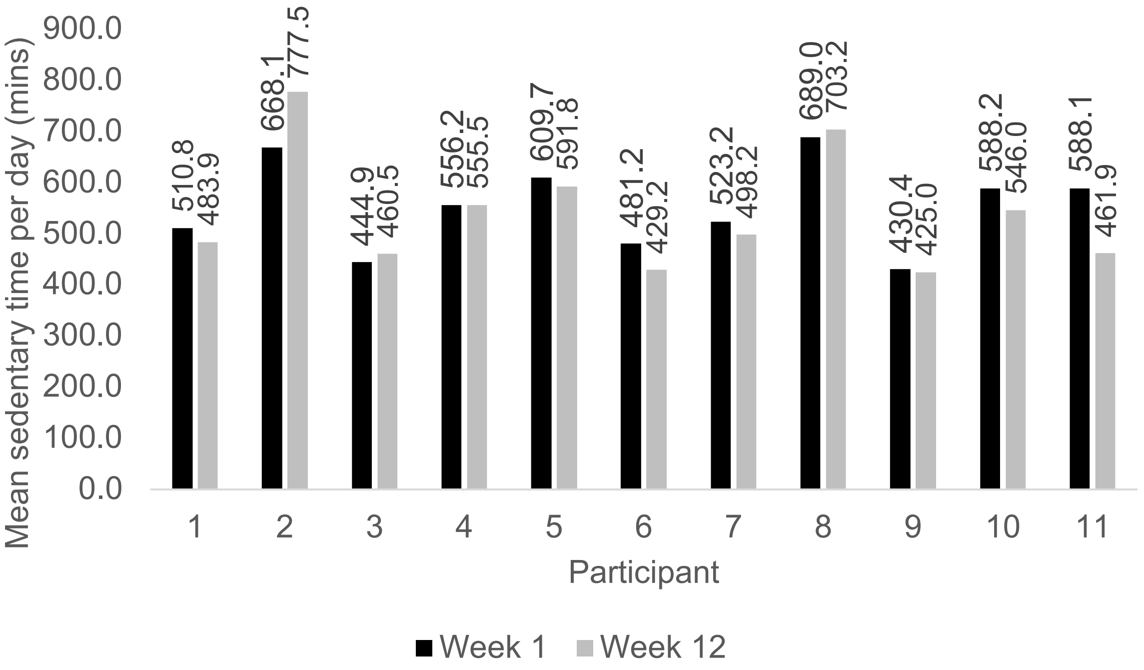


Figure 8.11. Accelerometer-measured mean sedentary time per day at weeks 1 and 12, illustrated for each participant.

Total accelerometer-measured stepping time at week 1 was 544.8 ±213.3 minutes per week, rising non-significantly to 597.5 ±240.3 minutes per week at week 12 (t(10) = -2.047, *p* >0.05). At week 1, 125.6 ±101.1 minutes of this stepping was at moderate-vigorous intensity, rising non-significantly to 153.6 ± 118 minutes at week 12 (t(10) = -1.378, *p* >0.05) (Table 8.5).

Table 8.5. Total accelerometer-measured mean ±SD minutes of PA (per week) accumulated at weeks 1 and 12, according to the duration and intensity of the PA bout.

	Week 1		**Week 12**			
	Mean	**SD**	**Mean**	**SD**	**Difference**	**SD**
Any bout duration, any intensity	544.8	213.3	597.6	240.3	+52.7	85.4
Any bout duration, ≥moderate intensity	125.6	101.1	153.6	118	+28	67.4
Bout duration ≥10 mins, any intensity	41.6	45.9	50.6	48.5	+9	31.8
Bout duration ≥10 mins, ≥moderate intensity	23.1	30.2	41.8	38.9	+18.7	40.1

On an average day, most stepping bouts at week 1 were either short in duration, lasting 0-30 seconds (40.3%) or 30-60 seconds (13.9%), or longer in duration lasting ≥5 minutes (21%). At week 12 a similar pattern was observed, but with a small decline in shorter duration bouts, and a small increase in longer duration bouts (0-30 seconds (37.2%), 30-60 seconds (13.7%), ≥5 mins (22.3%)). (Figure 8.12, below).

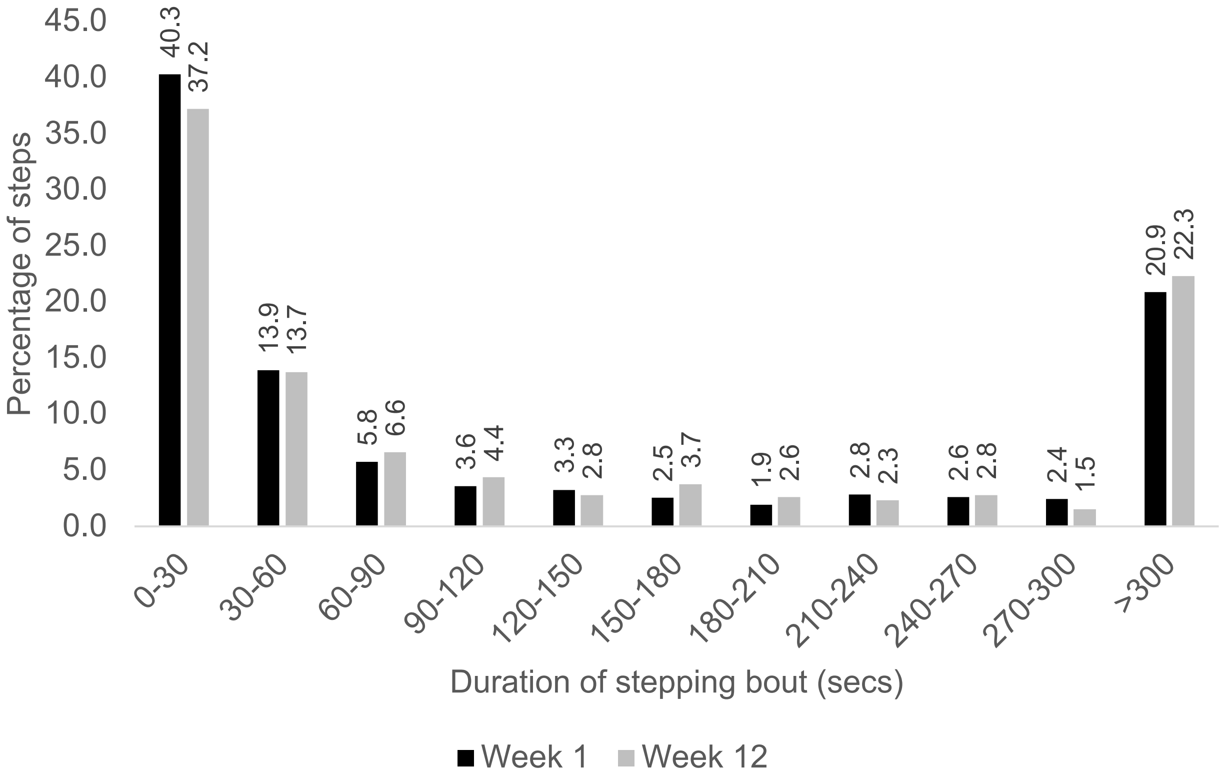


Figure 8.12. The percentage of total steps that occurred in each bout length, on an average day, at weeks 1 and 12 (accelerometer-measured).

On an average day, most stepping bouts at week 1 were low intensity (68.1% between 20-100 steps per min), with 31.9% considered moderate-vigorous intensity (≥100 steps per min). By week 12, low intensity stepping had reduced by 3.1% to 65%, and moderate-vigorous intensity stepping had increased to 35% (Figure 8.13, below).

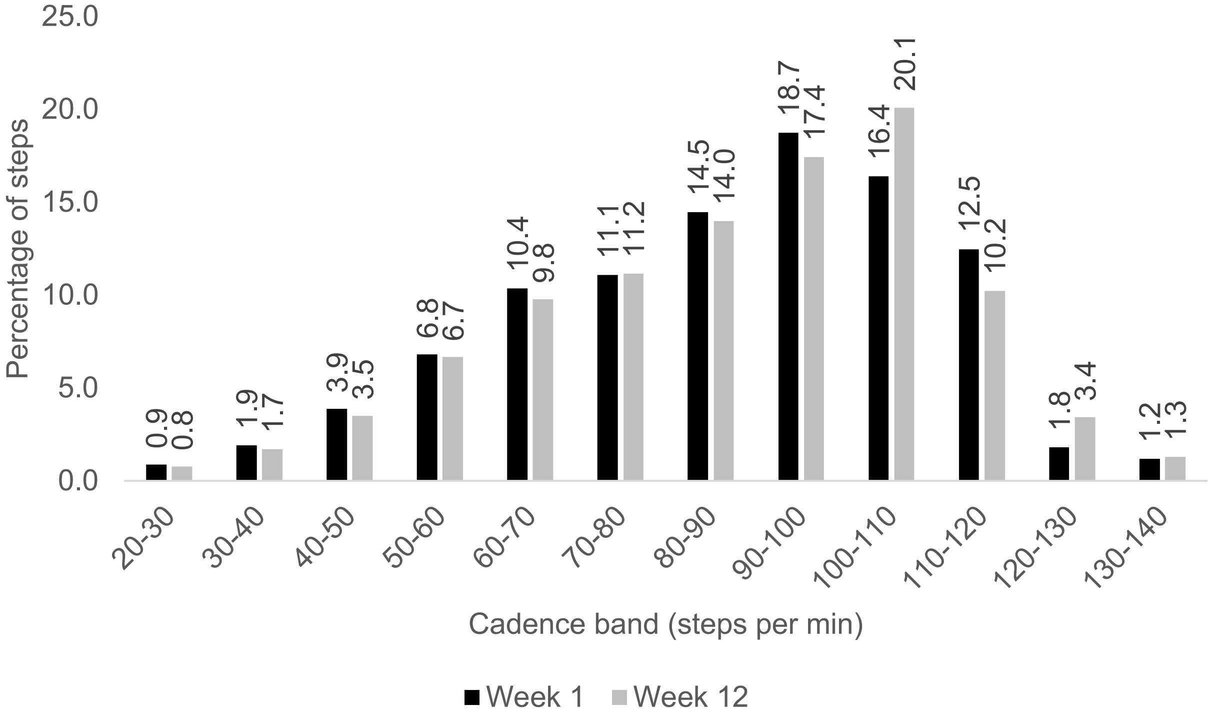


Figure 8.13. The percentage of steps that occurred in each cadence band, on an average day, at weeks 1 and 12 (accelerometer-measured).

Change in intensity of PA can also be observed in Figure 8.14 below, which illustrates that 6/11 participants (5, 6, 8, 9, 10, 11) increased total MVPA per week between weeks 1 and 12, with five of these participants (5, 6, 8, 10, 11) accumulating ≥150 minutes of moderate-vigorous intensity stepping per week at week 12. It is interesting to note that three out of these six participants (6, 8, 11) already achieved ≥150 minutes of MVPA per week at week 1. Figure 8.10 also illustrates that five participants (1, 2, 3, 4, 7) showed a decline in MVPA per week between weeks 1 and 12. Three of these participants (2, 3, 7) also showed a decline in total PA per week, whilst two particpants (1, 4) increased low intensity PA per week.

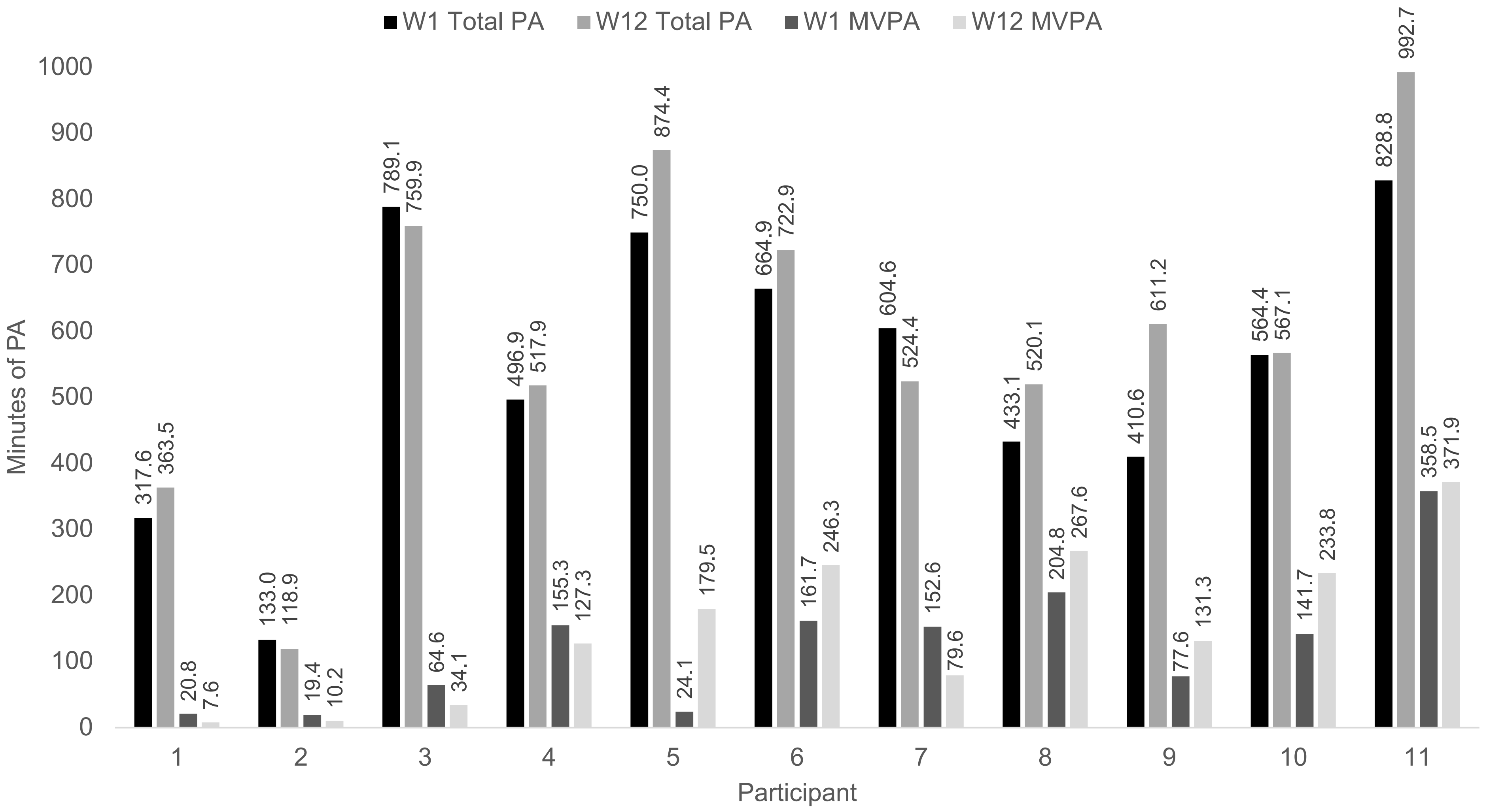


Figure 8.14. Accelerometer-measured total PA, and total MVPA (per week), at weeks 1 and 12, illustrated for each participant.

When stepping bouts were further analysed taking duration of bout in to consideration, 8/11 participants (72.7%) achieved ≥1 stepping bouts lasting ≥10 minutes at both time points, accumulating a mean of 41.6 ±45.9 minutes per week in stepping in bouts lasting ≥10 minutes at week 1, rising non-significantly to 50.6 ±48.5 minutes per week at week 12 (t(10) = -0.931, *p* >0.05) (Table 8.5). Therefore, 3/11 participants did not achieve a single bout of PA lasting ≥10 minutes at either time point.

When considering both duration and intensity, 7/11 participants (63.6%) achieved ≥1 bouts of moderate-vigorous intensity stepping lasting ≥10 minutes at week 1, rising to 8/11 participants (72.7%) at week 12. Time spent stepping, in bouts lasting ≥10 minutes at moderate-vigorous intensity, increased non-significantly from 23.1 ±30.2 minutes per week at week 1 to 41.8 ±38.9 minutes per week at week 12 (t(10) = -1.544, *p* >0.05) (Table 8.5).

8.5. Study 4 – Discussion

8.5.1. Study 4 - summary of findings

The aims of this study were to explore the PA and sedentary behaviour of participants in the ‘Live Active’ scheme using accelerometer-based measurement, and to compare agreement with self-report. To date, the sedentary behaviour of ERS participants has been underexplored, through self-report or device-based methods (excluding Giné-Garriga et al., 2017, in progress). Furthermore, all known measurement of habitual PA in ERS evaluations identified in the literature review has been through self-report questionnaires, with a growing number of researchers recommending that future evaluations also measure PA using devices such as accelerometers (Campbell et al., 2015; Hanson et al., 2013; Isaacs et al., 2007; Pavey et al., 2011). This study therefore contributes new knowledge, by reporting accelerometer-measured PA and sedentary behaviour of ERS participants, in addition to self-reported behaviour.

The findings showed that participants were highly sedentary when they entered the scheme, and were intermittently active at a low intensity, below the recommended PA guidelines. After 12 weeks of participation in the scheme, participants remained highly sedentary, increasing total PA non-significantly. Therefore, on average, participants were not physically active within the recommended PA guideline levels by week 12. However, these non-significant increases in PA could still be considered meaningful, with participants achieving, on average, an additional 53 minutes of PA at week 12. However, at the individual level, variation in PA between participants could be observed, with several participants showing a decline in total PA whilst several participants showed an increase (Figure 8.14).

The findings also showed that participants under-reported sedentary behaviour and over-reported PA. There is a general consensus that individuals over-report time spent in moderate or vigorous intensity PA, and under-report time spent sedentary or in light intensity PA (Cerin et al., 2016; Chastin et al., 2018; Steene-Johannessen et al., 2016). The findings therefore agree with this consensus, that individuals under-report sedentary behaviour and over-report MVPA. However, the participants in this study also over-reported light intensity PA, which is in contrast with this consensus.

8.5.2. Change in sedentary behaviour and physical activity level

The accelerometer-measured sedentary behaviour of participants upon entry to the scheme was higher than the national average for adults. Data from the latest Health Survey for England to report sedentary behaviour showed that adults spend between 4.6 – 5.3 hours per day sedentary (HSCIC, 2017). In this study, when considering participants with data available at both time points, average accelerometer-measured sedentary behaviour was 9.2 hours at week 1, and 9 hours at week 12, therefore substantially higher than the national average. However, data from the Health Survey for England was self-reported, and this most likely contributed to the substantial difference if participants under-reported sedentary behaviour. In a previous Health Survey for England (HSCIC, 2009), in which sedentary behaviour was assessed using accelerometery, average sedentary time was 9.7 – 9.9 hours per day, similar to the findings of this study.

There were no significant differences in accelerometer-measured PA between weeks 1 and 12. There were, however, non-significant increases in total PA (+53 mins), total MVPA (+28 mins), and total MVPA accumulated in bouts lasting at least 10 minutes (+19 mins). Exploration of the data also revealed inter-individual variation in change in PA level, with 8 out of 11 participants increasing total PA, but to varying extents (Figure 8.14). These results are therefore in contrast with the findings of study 2 that found significant increases in self-reported PA. Had the increases in PA observed in this study been statistically significant, they would have been considered meaningful increases. Therefore, future research is required to confirm these changes in PA using a larger sample size, as the small sample size in this study, and wider limitations of PA measurement, have limited the ability to draw generalisations from this study to the significantly larger sample in study 2.

With the exception of the excluded studies highlighted in section 8.2.4, no other identified studies have explored the sedentary and PA behaviour of ERS participants using device-based measurement. Therefore, the results of this study cannot be directly compared with those of a similar ERS evaluation. The lack of studies exploring the sedentary behaviour of ERS participants may be due to the focus of ERS evaluations on change in PA level, rather than sedentary behaviour. It could be anticipated that participants in an ERS would replace sedentary time with PA.

Therefore, it is not surprising that there was no significant change in sedentary behaviour, as there was no significant change in PA level. However, in a systematic review by Prince, Saunders, Gresty and Reid (2014), the authors concluded that interventions that focused solely on reducing sedentary behaviour had a greater impact than those that focused on either PA, or sedentary behaviour and PA at the same time. The focus of an ERS on PA change may therefore also explain the small, non-significant impact on sedentary behaviour.

The non-significant difference in PA level remains unexpected, given that the aim of an ERS is to increase PA levels (NICE, 2014a), and that the results of study 2 supported the achievement of this aim. However, in the interpretation of these findings, one must consider the limitations of PA measurement methods. These include the potential over-reporting of PA in study 2, and the limited ability of device-based methods to detect activities such as resistance exercises in this study. The non-significant differences observed in this study could be attributed to limitations such as these, in which a true reflection of PA is not obtained, which is discussed in more detail in section 8.5.4.

8.5.3. Comparison of accelerometer-measured and self-reported outcomes

In this study, participants under-reported sedentary behaviour, and over-reported light intensity PA and MVPA. These findings are in agreement with previous research (Cerin et al., 2016; Chastin et al., 2018; Craig et al., 2003; Rosenberg et al., 2008; Steene-Johannessen et al., 2016; van der Ploeg et al., 2010), with the exception of the over-reporting of light intensity PA.

It is not possible to determine if a particular source of bias contributed to these findings, but difficulty recalling PA and sedentary behaviour (Altschuler et al., 2009; Heesch et al., 2010), understanding the classification of PA intensity (Finger et al., 2015; Scholes et al., 2016), or social desirability (Adams et al., 2005; Motl et al., 2005) could have contributed. Furthermore, the IPAQ-SF asks participants to recall bouts of PA lasting 10 minutes or more, and participants may have self-reported PA accumulated from shorter duration bouts. This could have contributed to the discrepancy between instruments, as only accelerometer-measured bouts of PA lasting 10 minutes or more were included in this analysis.

It is also important to consider how limitations of accelerometer measurement and analysis may have contributed to the discrepancy between self-reported and accelerometer-measured PA. In this analysis, accelerometer-measured PA bouts were isolated to those lasting ≥10 minutes, with no interruptions, for comparison with the IPAQ-SF that asks participants to recall bouts of PA lasting ≥10 minutes. This does not account for brief interruptions in continuous PA, such as a participant interrupting a brisk walk for 10-seconds to cross a road safely. In this instance, the accelerometer records walking events before and after the interruption as two separate bouts of activity. This is an important limitation of this analysis, as it may not have accurately reflected time spent in continuous bouts of PA (≥10 mins), if brief but necessary interruptions in PA have occurred (Barry, Galna, Lord, Rochester, & Godfrey, 2015).

The findings from this study raise questions about the validity of the self-reported PA and sedentary behaviour data reported in study 2. However, in light of several limitations (discussed in section 8.5.4. below), it is difficult to generalise these findings to those reported in study 2, or to draw strong implications for methods of PA measurement in ERSs. Despite participants over-reporting PA at weeks 1 and 12, the accelerometer-measured PA levels of participants at these time points can still be considered meaningful levels of PA associated with health benefits (Arem et al., 2015; Hupin et al., 2015; Wen et al., 2011) (week 1 – 53 mins, week 12 – 63 mins of

PA accumulated in bouts of atleast ≥10 minutes). Furthermore, in study 2, there were significant improvements in the majority of health-related outcomes, which indicates that health enhancing levels of PA did increase.

These conclusions lead to the agreement that future studies should interpret self-reported PA as an estimate of perceived PA, rather than actual PA behaviour (Pettee-Gabriel et al., 2012), and that this data should be referred to as self-reported PA to reflect that it may not be accurate (Troiano, 2009). In future ERS evaluations, it would also be beneficial to interpret changes in self-reported PA levels alongside changes in health-related outcomes, to help determine if health enhancing levels of PA have indeed increased.

8.5.4. Limitations of study 4

This is one of the first known studies to explore the PA and sedentary behaviour of ERS participants using accelerometers. However, it is important to note a number of limitations of this study. The small sample size in this study has limited the ability to draw generalisations about change in PA and sedentary behaviour in this ERS. The small sample size was predominantly due to a reluctance amongst 'Live Active' participants to wear a leg worn accelerometer. There was also a high dropout from the week 12 follow-up due to participants dropping out of the scheme (n=12), or not wishing to repeat the protocol (n=15). This demonstrates the difficulty that can be faced when using devices such as accelerometers in research, and may explain why few studies have attempted to use device-based measurement in an ERS.

There are also limitations of accelerometer measurement and analysis, which may have significantly impacted the outcomes reported in this study. As discussed in the previous section, an important limitation is the impact of brief interruptions in PA on continuous bouts of PA. This is an important limitation in this study, as the results may not accurately reflect time spent in continuous bouts of PA (≥10 mins) if brief interruptions in PA have occurred (Barry et al., 2015). This is an important limitation of the reporting of PA accumulated in bouts of ≥10 minutes in this study, and the comparison of self-reported with accelerometer-meausured PA, which only took accelerometer-meausured bouts lasting ≥10 minutes into consideration.

A further technical limitation is the inability of the accelerometer used in this study to accurately measure upper-body or ambulatory activities such as weightlifting and cycling, in addition to water-based activities such as swimming. Resistance exercises, for example, may have been mistaken for a stationary standing or sitting event. Water based events such as swimming would not have been recorded. Therefore, total accelerometer-measured PA may have been underestimated for those participating in these activities, contributing to the non-significant increases in PA, and the over-reporting of PA (Pedišić & Bauman, 2015; Troiano et al., 2014; Trost, 2007; Trost & O'Neil, 2014).

Self-report questionnaires are also subject to limitations, especially when they are not used according to their guidelines (Ainsworth et al., 2012, 2015). The IPAQ was developed for population PA surveillance, for adults aged between 15-69 years. However, as observed in this evaluation, it is being routinely used with adults aged over 69 years, and as an evaluation tool in small scale intervention studies (IPAQ, 2005). Inappropriate use of measurement tools can lead to type 2 errors, with the rejection of interventions that appear not to improve behaviour, but in reality are successful (Kelly et al., 2016). Furthermore, the IPAQ-SF asks participants to recall bouts of PA lasting ≥10 minutes. This may have introduced a limitation to the comparison of self-reported and accelerometer-measured PA, if participants have reported bouts of PA lasting <10 minutes in their estimation of PA.

It is also important to consider intra-individual variation in habitual PA level, which can be influenced by factors such as the environment (Levin, Jacobs, Ainsworth, Richardson, & Leon, 1999; Matthews, Ainsworth, Thompson, & Bassett, 2002; Pedersen, Danquah, Petersen, & Tolstrup, 2016b). As discussed in study 1, for example, poor weather can impact PA behaviour on a short-term, or seasonal basis (Tucker & Gilliland, 2007). Furthermore, as found in study 3, participants can experience temporary interruptions to habitual PA due to factors such as illness. Therefore, at the individual level, PA behaviour can vary from day-to-day, week-to-week, and seasonally, potentially requiring more than 7-days of monitoring to establish true habitual PA (Bergman, 2018).

It is therefore recommended that future research assesses PA in ERS participants using device-based measurement, on the premise that substantial limitations can be overcome. This includes, for example, ensuring that a large sample size is feasible, that participants can comply with wear time protocols, and that devices can measure PA types prescribed by the ERS including resistance, ambulatory and water-based activities.

8.5.5. Conclusions of study 4

The findings of this study indicate that participation in the 'Live Active' scheme, for this sample, did not lead to statistically significant improvements in accelerometer-measured sedentary behaviour and PA in the short term. However, the observed increases in PA could still be considered meaningful. These results are, though, substantially limited by the small sample size achieved, and the wider limitations of measurement instruments. This study has demonstrated that self-report measures have limitations, but that accelerometers also have significant limitations. This leads to the agreement that there is no single best method for the measurement of PA behaviour without error (Bergman, 2018). This has important implications for PA promotion schemes such as ERSs, which demonstrate level of success by change in

PA behaviour (Campbell et al., 2015; Pavey et al., 2011). Despite the importance placed upon change in this outcome, there has been a noticeable lack of critical discussion about methods of assessing PA in ERSs. Given the limitations of PA measurement methods, it is suggested that stakeholders (e.g. researchers, commissioners, policy makers) place an equal emphasis upon the evaluation and achievement of improvement in health outcomes, as well as improvement in sedentary behaviour and PA, given that the aim of an ERS is to increase PA levels in order to improve health (NICE, 2014a). A similar approach was recently illustrated in an ERS evaluation by Stewart et al. (2017), in which the authors did not report PA level, but evaluated the impact of the ERS based solely on change in a range of health outcomes.

Chapter 9 Discussion

This final chapter provides a summary of the aims and objectives of the thesis, the main findings and contributions to knowledge, and the strengths and limitations of the thesis. Recommendations for research, practice and policy are made based on the literature review and the findings of this thesis, followed by a final conclusion.

9.1. Summary of the aims and objectives of the thesis

The overall aim of this thesis was to evaluate the effectiveness of the ‘Live Active’ ERS at increasing PA levels and improving health outcomes in its participants. This research was important, because despite there being over 600 ERSs in operation in the UK (Pavey et al., 2011), with the ERS model being implemented abroad (see for e.g. Gallegos-Carrillo et al., 2017; Pedersen et al., 2016a; Sørensen et al., 2011), the evidence base for the impact of ERSs is relatively small and equivocal (Campbell et al., 2015; Pavey et al., 2011), with the recommendation that disinvestment in ERSs is considered (Pavey et al., 2011). The first two studies in this thesis were longitudinal observational studies, which used secondary quantitative data collected by the ERS over a 2-year period. The first study aimed to understand engagement levels in the scheme, reasons for non-engagement, and which individual or referral characteristics (e.g. ethnicity, SES, referrer organisation) could predict uptake and adherence. The second study aimed to evaluate the medium to long-term impact of participation on PA levels and health-related outcomes, and if they differed between groups with different primary referral conditions. The third study aimed to understand the perceived benefits of participation in the ERS using mixed methods, to provide a more balanced evaluation of the scheme by capturing participants’ experiences. The fourth and final study aimed to evaluate the short-term changes in sedentary behaviour and PA, measured using accelerometers, and how these outcomes compared with self-reported sedentary behaviour and PA.

9.2. Summary of the main findings

As the findings of each study have been discussed in detail in each respective chapter, the main findings and contributions to knowledge will be synthesised in this section. Study 1 found that engagement levels in the 'Live Active' scheme were higher than those reported by previous evaluations, and synthesised in a systematic review by Pavey et al. (2012) (average uptake rate 66-81%, average adherence rate of 43%-49%), with a high uptake (81%) and adherence (56.9%) level. It is not possible to note with certainty the reasons for higher than expected engagement in this ERS. It is possible that engagement was related to aspects not explored in this study, but known to influence adherence, such as the longer than average scheme duration (Rowley et al., 2018), the subsidised price of activities, the personality and approach of the exercise professionals (Morgan et al., 2016), or because the scheme was developed based upon knowledge of the local context (Tobi et al., 2017).

This study added to emerging observations that the majority of dropout (52%) occurs in the first six weeks of participation (Kelly et al., 2017; Moore et al., 2013). Therefore, this is a critical period for preventing dropout. However, the reasons for early dropout were similar to those reported by participants who dropped out after the first six weeks, and may be better explored through qualitative research. Nonetheless, the finding of frequent early dropout has identified an area for future research to inform improvements in practice.

Study 1 also highlighted the value of exploring the reasons for non-engagement when interpreting engagement levels. The results showed that non-engagement occurred for a variety of reasons such as difficulty contacting participants (hypothesised to be associated in-part with lack of readiness to change), perceived health status, and health inequalities such as finance to pay for activities. However, non-engagement was not always a negative event, as previously hypothesised

(Harrison et al., 2005a), with some participants dropping out due to becoming independently active and not requiring further support. The results also showed how inappropriate referrals can inflate non-engagement levels (e.g. >8% of non-uptake), as previously observed (Johnston et al., 2005), and it is therefore important for ERSs to identify and seek to reduce inappropriate referrals. These findings indicate that non-engagement does not always necessarily reflect an ineffective ERS. The reporting of reasons for non-uptake and dropout from a large sample is novel, and it is recommended that ERSs regularly review these reasons, as they may identify areas for service improvement to maximise engagement.

More extensive participant profiling by the scheme allowed for the inclusion of a wider range of predictor variables in multivariate analyses. Uptake was associated with SES (Q2, Q4) and season of referral (summer). Adherence was predicted by entering the scheme in autumn, spring or summer, or being a non-smoker at the time of the initial consultation. Few studies have explored the impact of smoking status (Kelly et al., 2017; Ward et al., 2010), and further research is required to understand if smokers require additional support to become physically active. The association between season and engagement is also novel, with no identified published ERS evaluations examining this variable. This may support observations that seasonal variations in health status (Dopico et al., 2015), and the weather, may be a barrier to PA (Tucker & Gilliland, 2007). Collectively, these associations have prompted this ERS to consider what additional steps can be taken to reduce the likelihood of these factors negatively impacting engagement in the future.

One of the most important findings of this thesis, reported in study 2, was the significant medium to long-term improvements in self-reported PA level and the majority of health-related outcomes. These findings are important as previous systematic reviews, which are based on a small number of RCTs but inform policy, have found no consistent evidence for an improvement in PA and health outcomes, generating a consensus that ERSs are ineffective (Campbell et al., 2015; Pavey et al., 2011). The studies included in these reviews have also tended to explore the short-term impact of ERSs, with few studies reporting the long-term outcomes of an ERS, or the impact on multiple health-related outcomes spanning physical health, mental wellbeing and HRQoL. Although causality cannot be established without a comparison group, the results of this study suggest that the 'Live Active' scheme was successful at initiating and sustaining long-term PA behaviour change, in turn positively benefitting health, therefore meeting the aims of an ERS (NICE, 2014a).

A further original contribution to knowledge was the sub-group analysis of these outcomes based on primary referral condition, which indicated that participation was most effective for those referred for musculoskeletal conditions. It is not known why this group appeared to benefit most, but differences between referral groups are to be expected given the different aetiologies of these conditions, and therefore the different PA programmes provided, and the different ways in which one can benefit from participation (Rowley et al., 2018). The findings also showed a positive outcome of participation for those referred for mental health issues, who improved mental wellbeing scores more than any other group. Therefore, these findings have highlighted the value of referral to this ERS for these groups, in which there is a current imperative to improve health-related outcomes (Arthritis Research UK, 2018; Department of Health & Social Care, 2018a). As a result, there is scope for the targeted referral of participants with musculoskeletal conditions and mental health issues. However, adaptations to practice for particular groups must also be identified, to improve condition-related outcomes (e.g. obesity and BMI).

No identified studies have completed such an extensive sub-group analysis, reporting differences across this number of primary referral conditions or health outcomes. These results support emerging findings that differences in outcomes may exist between groups, which are masked when results are aggregated at the population level (Murphy et al., 2012; Rowley et al., 2018). Furthermore, population level analyses are limited in their usefulness, and do not provide direction for targeted service improvement. In this study, for example, participants referred specifically for having a high BMI did not significantly improve BMI. There is an urgent imperative to reduce levels of obesity in the UK (Department of Health & Social Care, 2018a), and this has therefore highlighted an important area for service improvement with this group. It is therefore recommended that future evaluations consider the inclusion of a sub-group analysis, to understand who benefits from participation and in what way, and to explore the long-term impact of participation beyond 12 months.

Study 3 made an original contribution to knowledge about the benefits of participation in an ERS using a mixed methods approach. The findings of qualitative interviews and feedback questionnaire data collectively demonstrated a holistic benefit of participation. Improved social health was the most apparent benefit, which was one of the more novel findings of this study, and promising given the recent mandate to reduce social isolation and loneliness in the UK (Department for Digital, Culture, Media & Sport, 2018). Benefits were also related to PA, physical health, mental health, and adjunct benefits. The positive benefits reported in this study provided a compelling narrative for the impact of this ERS. Considering the reported impact of participation, spanning physical, mental and social health, where there is an imperative to urgently improve outcomes and reduce expenditure, these findings highlight the value of ERSs as a form of non-medical intervention, or ‘social prescription’ (Thomson et al., 2015).

One of the most striking observations was that many of the benefits reported in this study were not routinely evaluated by the scheme (study 2), nor included in previous quantitative ERS evaluations (Campbell et al., 2015; Pavey et al., 2011), or evaluation frameworks such as the SEF (Cavill et al., 2012). Consequently, evaluating the impact of this scheme based solely on the outcomes of study 2 would have underestimated the effectiveness of the intervention (Wilson, 2015). Therefore, ERSs have been considered by some to be an ineffective use of public funds (Pavey et al., 2011), when in fact they may have the potential to reduce the burden of LTCs in the UK, and issues such as social isolation, if more participants are referred. The results of this study have led this ERS to reconsider what outcomes are collected during consultations to demonstrate impact, such as an update to the content of the feedback questionnaire, in order to capture a wider range of benefits.

These results support arguments for more pluralistic evaluations of ERSs, which can provide a more balanced evaluation of a scheme’s impact (Crone et al., 2005; Dugdill et al., 2005; Gidlow et al., 2008). Evaluation frameworks should therefore also be updated, to recommend that aspects such as social health are evaluated, and not just physical aspects of health. Moving forwards, based on the positive outcomes of studies 2 and 3, qualitative research is now necessary to understand the mechanisms that led this scheme to be successful. Similar work is also necessary in other ERSs identified as effective, to generate an understanding of the common principles behind successful ERSs, which can then be scaled across schemes.

Study 4 made an original contribution to knowledge by being one of the first known studies to assess change in habitual PA and sedentary behaviour of ERS participants, using device-based measurement. The findings showed that participants were highly sedentary upon entry to the scheme, and physically active below the recommended PA guidelines. By week 12, there were no significant changes in accelerometer-measured sedentary behaviour or PA. However, had the increases in

PA observed in this study been statistically significant, they would have been considered meaningful increases. It would therefore be beneficial for future research to confirm these changes in PA using a larger sample size.

Analyses also showed that participants under-reported sedentary behaviour, and over-reported PA, which supports previous research finding comparable results (Cerin et al., 2016; Chastin et al., 2018; Steene-Johannessen et al., 2016). However, despite participants over-reporting PA at weeks 1 and 12, the accelerometer-measured PA levels of participants at these time points can still be considered meaningful levels of health enhancing PA.

In light of several limitations, it is difficult to generalise these findings to those reported in study 2, or to draw strong implications for methods of PA measurement in ERSs. Several technical limitations of device-based methods, for example, may lead to outcomes which are not a true reflection of actual PA. However, despite the observation of over-reporting in this study, and previous studies, there were significant improvements in the majority of health-related outcomes in study 2, which indicates that levels of health enhancing PA did increase in this population. Therefore, it may be beneficial for future research to interpret changes in self-reported PA levels alongside changes in health-related outcomes.

This study concluded that both device-based measures and self-report measures have limitations, agreeing with the argument that there is no single best method for the measurement of PA behaviour without error (Bergman, 2018). There has been a noticeable lack of critical discussion about methods of assessing PA in ERSs, and this study has raised important questions about its measurement and interpretation, particularly as it is considered to be the primary outcome of importance. It was therefore argued that researchers and stakeholders of ERSs should place equal emphasis upon change in PA, as well as change in health, considering that the aim of an ERS is to increase someone's PA in order to positively benefit their health (NICE, 2014a).

In summary, these studies appear to demonstrate that the 'Live Active' ERS was successful at engaging participants, increasing PA behaviour and improving health holistically. It is hypothesised that this should translate in to a long-term public health benefit for the community of Tameside, and contribute towards reducing the burden of premature mortality in the area (PHE, 2018a; Steel et al., 2018).

Considering the recent drive to implement social prescribing in the UK (Department of Health & Social Care, 2018b), and updated global policy to recommend the referral of individuals with health conditions to community-based PA interventions (WHO, 2018), these findings support the value of referral to ERSs. Furthermore, the findings from studies 2 and 3 have indicated benefits such as improved mental health, social health, and medication utilisation; areas in which there is a current imperative in the UK to improve these outcomes, reduce associated expenditure, and therefore alleviate the overall burden of these issues to both individuals and wider society. It could be argued that further investment in ERSs, to allow a greater number of individuals to be referred, could result in a much larger national impact.

Aside from these implications, the findings of this thesis have also raised methodological considerations for how ERSs are evaluated in the future, and whether the content of ERS evaluation frameworks provide a framework for a balanced appraisal of schemes. These considerations will be further discussed below in section 9.5.

9.3. Strengths and limitations of the thesis

The strengths and limitations of each study have been discussed within the respective chapters, therefore the main methodological strengths and limitations of this thesis will be summarised in this section. As the evaluation of the 'Live Active' scheme was not a RCT, and did not have a comparison group, it is not possible to say with certainty that the results observed in this thesis were due to the ERS intervention. For this evaluation, neither research design was possible as the ERS had commenced prior to the evaluation. The role of RCTs of ERSs has been discussed throughout this thesis. It may be more feasible for future evaluations to use a non-randomised design which includes a comparison group of similar demographics. Despite this, using secondary data as part of an observational design has provided an opportunity for the evaluation of a real-life ERS, providing a higher level of ecological validity rather than internal validity (Gidlow et al., 2008). In the case of study 2, large sample sizes were achieved, leading to one of few longitudinal follow-up studies of ERS impact. The observational design also provided greater opportunities for learning, based on feedback of real-life practice (Gidlow et al., 2008). Feedback from the data analyses was provided to stakeholders during regular meetings and in the form of a quarterly written report, throughout the evaluation period. This provided an opportunity for regular, informed refinement of the scheme based on the feedback. The flexibility of the study design also meant that new questions from stakeholders could be answered with ease. Feedback loops are characteristic of complex interventions, and may impact their delivery and outcomes (Moore et al., 2018; Shiell et al., 2008). However, it is noted that substantial changes to practice as a result of this feedback were not introduced until after the conclusion of this evaluation. Any changes introduced during the evaluation period were not judged to significantly change practice or outcomes.

However, there are disadvantages of using secondary data, which mostly relate to the researcher's loss of influence over the data collection (McKnight & McKnight, 2011; Vartanian, 2011). One important limitation of studies 1 and 2 was the impact of missing data, which could not be controlled by the researcher. In study 1, this impacted the prediction model of adherence, as some demographic variables could not be included in the multivariate analysis. In study 2, this impacted the level of data available for each health outcome, potentially not indicating the real impact of participation upon each outcome. Furthermore, 30.5% of eligible participants did not have a follow-up at week 52, and it was not possible to determine if these participants were active or inactive at this time point. Therefore, the study could not determine the true impact of participation on long-term PA behaviour.

The mixed methods approach of this thesis has provided a greater understanding of the scheme's impact than could be gained from a purely quantitative or qualitative approach. The importance of this design was evident in study 3, which concluded that the impact of the scheme would have been substantially underestimated by the findings of study 2 alone. Study 3 was also one of few qualitative explorations of a scheme's impact, with no identified studies exploring the perceived benefits of participation to this extent using mixed methods.

Finally, the findings of this thesis may relate to the scheme's context within a specific geographical area, with particular strategic aims, referral pathways, methods of delivery and inclusion criteria, and may not be comparable to other ERSs, or have the same impact in another location (Arsenijevic & Groot, 2017; Campbell et al., 2015; Tobi et al., 2017). The findings may, however, be useful to researchers, practitioners, commissioners, and other stakeholders working in similar communities to Tameside.

9.4. Applied impact of this thesis for 'Live Active'

The findings from this thesis have provided the 'Live Active' ERS with new knowledge that has been used to direct service improvement, and evidence of impact that has been used to secure additional funding and continued commissioning of the scheme.

Evidence of impact, for example, has been provided to existing referrers, throughout the evaluation period, to sustain existing referral partnerships based on the positive impact observed for participants they have referred. This evidence has also been shared with potential new referrers, to inform of expected benefits for participants. This had led to the establishment of new referral partnerships, and additional funding to establish new referral pathways.

The findings of this thesis were summarised in a final evaluation report for the stakeholders of the 'Live Active' scheme, with recommendations for future practice. This report has since been shared with colleagues across the North-West, and researchers across the UK, in the interests of shared learning and recommendations to improve practice.

Regular data analysis and feedback was provided throughout the evaluation period, which enabled the scheme to make several improvements in delivery and practice. This included, for example, the introduction of new PA classes based on frequency of feedback, such as women's only classes and water-based classes. Regular data analysis and feedback also provided opportunities for more frequent learning, and identification of areas for additional staff training, in areas such as data collection and behaviour change techniques, in order to improve impact.

Working within the research setting, I have been able to communicate the importance of data collection, illustrate what can be learned from data analyses, and share examples of the positive impact that an evidence base can provide. This has helped to establish an evaluation culture within the scheme that can be continued, and has helped to educate the staff members on how they can evidence and communicate the scheme's impact.

Finally, the work of the ‘Live Active’ ERS has received frequent recognition, with staff invited to share evidence for the impact of the scheme at several local and national events. Using the evidence generated during this thesis in their application, ‘Live Active’ was a runner up for the ‘healthy communities’ award at the UK active awards 2018. This work contributed towards Active Tameside receiving the ‘organisation of the year’ award at this event.

9.5. Implications for research

As topic specific recommendations for future research have been made in each respective study chapter, this section will summarise what are considered to be the most important areas for future research, to advance the field of ERS evaluation. The majority of these recommendations relate to methods of evaluation, rather than changes to aspects of ERSs which could improve impact. This is due to the learning gained from this thesis highlighting the importance of establishing methods that provide a balanced evaluation of impact, before making recommendations for improvement in delivery.

Usually, research informs practice, when interventions demonstrated as successful in controlled research are implemented at scale across communities (Milat, King, Bauman, & Redman, 2011). However, in the case of ERSs, heterogeneous models of delivery were scaled across the nation without an evidence base in place for their effectiveness. With an estimated 600 plus ERSs in operation across the UK, it is now difficult to withdraw schemes, or implement radical change in practice (Sowden & Raine, 2008). The role of research today, in the field of ERSs, is potentially therefore to improve and sustain good practice (Pratt et al., 2015). However, estimation of the scale of ERSs was made in 2011 (Pavey et al., 2011), and therefore a renewed national audit of ERSs would be valuable to understand the current scale of implementation.

Without guidelines for the standardised delivery of ERSs, stakeholders have been free to develop schemes as they desire, tailoring the delivery according to professional experience and the local context. Such freedom to operate can lead to the development of schemes, such as ‘Live Active’, which are successful potentially due to this tailoring. However, this has led to great variation in scheme design across the UK, on key factors such as eligibility criteria, scheme duration, and PA prescription. There may also exist variation within schemes, as has been observed within the ‘Live Active’ scheme, in factors such as participant characteristics, and type of PA prescription. Therefore, the results of this thesis will be valuable to the stakeholders of the ‘Live Active’ scheme, and stakeholders of ERSs in communities with a similar local context. However, researchers tend to be concerned with generating a consensus on the impact of an intervention, by comparing and contrasting evaluations of ERSs across the country, using methods such as systematic reviewing. The learnings from this thesis have led to questions of whether it is appropriate, at this time, to systematically review and provide generalisations about the impact of highly heterogeneous interventions. It is considered that this is an important argument, as the results of systematic reviews provide recommendations on whether ERSs should continue to be commissioned and can ultimately influence policy and decision making.

However, the real world of policy and practice tends to be ahead of research (Pratt et al., 2015), and is also influenced by competing factors such as politics, finances and resources (Curtis, Fulton, & Brown, 2018). Local authorities that commission ERSs have a deep understanding of their community, and therefore tend to be interested in what works locally, or what worked elsewhere for a similar community (Curtis et al., 2018). Commissioners can therefore often place more value on local evaluations, than national, research-based evidence (Wye et al., 2015), supporting arguments for the relevance of systematic reviews of impact to influence ERS policy and decision making. There can also be a preference for informal, anecdotal evidence or ‘stories’, more so than formal published RCTs or systematic reviews (Curtis et al., 2018). Public health programmes can therefore be commissioned completely independent of research-based evidence (Pratt et al., 2015). Researchers must adopt a more flexible understanding of how research can influence ERS policy and practice (Pratt et al., 2015), in order to have such influence, and acknowledge the culture of decision making (Orton, Lloyd-Williams, Taylor-Robinson, O’Flaherty, & Capewell, 2011), and what can be considered to be useful evidence (Curtis et al., 2018; Orton et al., 2011). There is, therefore, an argument for more local evaluations of ERSs, with the aim to demonstrate local impact and to inform future practice, rather than to generate a consensus on the impact of ERSs nationally, whilst such heterogeneity exists in scheme design.

Considering the design and content of evaluations, vague evaluation frameworks, and no standardisation of the outcomes that ERSs assess, have led to substantial variation in the measures collected by schemes, and therefore those reported in evaluations. In order to improve ERS data collection practices, future external evaluations, and the content of evaluation frameworks, it is recommended that future research, in collaboration with stakeholders, establish which quantitative measures are most valuable to assess and report. The findings of this study suggest that measures of physical, mental and social health would be valuable, as well as adjunct measures such as medication and healthcare utilisation. This would reduce the variation in measures reported in evaluations, and potentially allow for the inclusion of a wider range of outcomes which could provide a more representative evaluation of impact. There may be a practical value to ERSs, if a standardised set of relevant process and outcome measures were published from this work, with a recommended definition or measurement tool for each measure.

However, public health policy and practice should be informed by a rich evidence base (Gidlow et al., 2008), and not quantitative indicators alone (Curtis et al., 2018). Therefore, even with the establishment of recommended measures, there is still a role for qualitative research in future evaluations of ERSs, and particularly for decision makers who prefer this type of evidence (Curtis et al., 2018). If the aim of an ERS is to increase PA to positively benefit health (NICE, 2014a), then the concept of health can have a myriad of interpretations, and purely quantitative evaluations may not fully answer the question of whether or not a scheme met this aim. As observed in study 3, qualitative inquiry provided a compelling narrative for the impact of this scheme, which would have been underestimated by quantitative evaluation alone. Even with the inclusion of additional quantitative outcomes such as social health, issues of underestimation may still persist, as researchers cannot fully predict the experiences of participants. Therefore, quantitative evaluations will always hold value to both researchers and stakeholders, but it is suggested that they are not the sole method for evaluation of an ERSs impact. This may dictate a philosophical change amongst some researchers, in what they believe to represent a successful ERS, moving beyond a set of traditional, physiological based measures.

Whilst it is noted that the SEF for PA (Cavill et al., 2012) is an evaluation framework generalised to PA interventions, and not just ERSs, it is the recommended framework for ERSs to use (NICE, 2014a). Therefore, it is recommended that this framework, and any subsequent frameworks relevant to ERSs, prompt ERSs and researchers to consider which outcomes and methods can demonstrate the achievement of an intervention’s aims, without underestimation of impact. This should therefore reflect the value of participants’ opinions when evaluating an intervention.

Finally, in order to progress from evaluations of ERSs, towards improvements in practice, it would be beneficial for researchers to identify ERSs with positive outcomes where qualitative research can then be focused to understand the mechanisms leading to a successful ERS. To assist this, there may be value in a shared database of ERSs, detailing their local context and demonstrated outcomes. This may overcome the barriers of affording external evaluation, and success in publishing findings, to incorporate the outcomes of internally evaluated schemes. Similar work is being undertaken to collate the outcomes of ERSs that use Referall as an online database to record participant outcomes (UKactive, 2018). In the short-term, a description of the local context of an ERS in any database or evaluation can help researchers, practitioners and decision makers to understand what works in similar communities to those they are working in. In the long term, following the identification of successful ERSs, and research to understand the mechanisms leading to this success, it is argued that a review of these mechanisms rather than a review of ERS impact, can provide more valuable information to inform the refinement and improvement of practice in ERSs, or changes to wider policy, which can ultimately contribute to increasing population PA levels.

9.6. Implications for practice

The recommendations summarised in this section are relevant to individuals working within ERSs, and aim to translate research in to practice, based on the findings of the literature review and research studies contained within this thesis. As recommendations for practice have been made in each respective study chapter, this section will summarise what are considered to be the primary recommendations. Specific recommendations made to the ‘Live Active’ scheme as a result of this research have been summarised in appendix 12.

The first recommendations relate to referral partnerships and practices. As observed in study 1, 55.2% of referrals were received from secondary care and community services, which was in contrast with previous studies that received the majority of referrals from primary care services (e.g. Hanson et al., 2013; Kelly et al., 2017; McGeechan et al., 2018). It was also noted that several ERSs accepted referrals from primary care services only (e.g. Duda et al., 2014; Edwards et al., 2013; Harrison et al., 2005a), potentially limiting the reach of ERSs in these areas. ERSs originated in primary care settings, targeting referrals predominantly from GPs, with ERSs therefore previously known better to some as GP ERSs (Dugdill et al., 2005). This ERS has demonstrated that referral partnerships can be successfully developed with not just GPs, but nurses and allied health care professionals in primary care, secondary care, and community settings, reaching a greater number of participants in need of intervention. Therefore, it is recommended that ERSs should not limit their referral partnerships to primary care settings, but should seek to establish partnerships with a range of allied health care professionals, in both primary and secondary care, in order to reach more participants in need of intervention. This is particularly relevant, given recent initiatives to drive the implementation of social prescribing in primary and secondary care (e.g. Department of Health & Social Care, 2018b).

With regards to referral practices, study 1 identified that the scheme received a substantial number of inappropriate referrals (e.g. >8% of non-uptake). Referrals were received, for example, for participants with contraindications to PA, which inflated non-uptake and dropout levels. It would therefore be beneficial for ERS staff to ensure that all referrers are aware of the inclusion and exclusion criteria, and to provide feedback when inappropriate referrals have been received. These steps could reduce the incidence of inappropriate referrals, and provide an opportunity for health care professionals to prescribe alternative interventions for those not eligible for their local ERS.

Recommendations can also be made in relation to improving participant engagement with ERSs. Study 1 reported the reasons for non-uptake and dropout from the scheme, which provided an opportunity for learning and informed service improvement, to prevent non-engagement. It was noted, for example, that 15.1% of participants who dropped out did so because they perceived they could not be active in respect of their health status. This has identified an area for additional staff training, in using behaviour change techniques to overcome health related barriers to PA. It is therefore recommended, that ERSs record and regularly review the reasons for non-engagement, to identify areas for improvements in practice, which could ultimately improve engagement and therefore the likelihood of positive PA behaviour change.

In study 3, the findings from the feedback questionnaire indicated that participants had often made new friends during their time on the scheme, and felt less isolated. This was also frequently reported in the semi-structured interviews, in which many participants were taking part in group activities. These findings are perceived to have two important implications for practice. The findings of this study, and previous research (e.g. Kelly et al., 2017, Moore et al., 2013), have identified that the majority of particpants dropout from an ERS within weeks of uptaking, and that socialisation and social support within schemes can facilitate adherence (Morgan et al., 2016). It may be pertinent therefore, for exercise professionals to prescribe group-based activities during this early period, to facilitate enjoyment and self-efficacy for PA in a supportive environment, before introducing independent activities.

The second implication relates to the maintenance of PA in those who prefer group-based activities. In this ERS, participants had access to classes provided by the ERS for one year. It is important to note that participants were also able to take part in other group activities in the community, to which there is no time limit. However, it is not known what happens to the PA levels of participants who attend these ERS classes, once they are withdrawn. To prevent a decline in PA in this group, it would seem pertinent to offer a ‘maintenance class’ to previous participants of the scheme. ERSs are often based in local leisure facilities, in which group-based activities are often targeted towards the general population. The provision of a maintenance class, or other group-based classes suitable for an ERS population, would allow participants to continue to be physically active in a way they find enjoyable and beneficial to their health.

Lastly, ERSs were developed in the early 1990s, when the landscape of public health services was substantially different. Over the last decade, the UK has faced a period of advancement in public health practice, but also austerity which has seen cuts to public services (Thomson et al., 2014). Today, ERSs that receive local funding to operate, are under more pressure than ever to demonstrate their impact and value for money. Finance, time, and skill are generally required to evaluate the impact of an ERS, usually in the form of an external evaluation. However, this is not a financially viable reality for most ERSs and is usually reliant on additional funding. Therefore, it is important for ERSs to develop an evaluation culture amongst staff members (Gidlow et al., 2008), which reinforces the importance of data collection to demonstrate impact. Managers would also benefit from education and guidance on how to evaluate data and disseminate impact effectively. This would allow ERSs to regularly review their impact to inform service refinement, and to provide evidence of impact to stakeholders when necessary. Based on the findings of study 3, it would be valuable for ERSs to collect not just a range of quantitative health-related outcomes, but to also include a feedback questionnaire such as the one used in study 3, in addition to several short qualitative case studies, to identify and evidence adjunct benefits to participation which may represent additional value for money. This would also be useful information for commissioners, who often prefer a variety of evidence types, and those that provide a compelling case for investment (Curtis et al., 2018; Wye et al., 2015).

9.7. Implications for policy

This section summarises the most important implications of this research for ERS related policy, which have arisen from the results of the studies within this thesis and the findings of the literature review. There is recognition amongst researchers that the current guidelines for ERSs are ambiguous (Campbell et al., 2015; Pavey et al., 2011), leading to a myriad of potential interpretations (Henderson et al., 2018; Oliver et al., 2016), as evidenced by the substantial diversity in delivery between schemes (Oliver et al., 2016). Even within ERSs, stakeholders such as exercise professionals, scheme managers, and commissioners can have conflicting interpretations about what the aims of an ERS are (Henderson et al., 2018). It would therefore be beneficial for aspects of the ERS guidelines in the UK to be updated, to reduce ambiguity and uncertainty.

From a practice perspective, updated guidelines could reduce the substantial heterogeneity in delivery between ERSs, whilst still allowing for tailoring according to the local context. Usually, the aim is to identify and scale a successful PA intervention, to impact population PA levels (Milat et al., 2011). However, given that ERSs are already widespread, implementing substantial change in delivery could be difficult and subject to resistance to change (Sowden & Raine, 2008). It may be that there is no single model of best practice that is successful when scaled. In this scenario, it may beneficial for any future ERS guidelines to include a set of quality standards for delivery, derived from future research which identifies the factors leading to success in ERSs shown to be effective. Similar work is already underway in the Greater Manchester region, to promote consistency across ERSs in the area, by scaling up the principles leading to successful outcomes, through a standards framework (PHE, 2017). The findings of this study suggest that factors such as referral practices (e.g. eligible participants and referrers), the duration of participation (≥6 months) (Rowley et al., 2018), and the facilitation of social inclusion and independent PA (Morgan et al., 2016) may have contributed to success in this ERS. However, further research is required to confirm the exact mechanisms, and whether these contribute to success in other ERSs.

Often, local strategic aims can influence the eligibility criteria for ERSs, creating inequality in access between local areas. A set of standards, which promotes one eligibility criteria, can provide an opportunity to reduce this inequality in access, providing opportunities for PA to those who need it. Furthermore, from an evaluation perspective, a clearer statement of the aims of an ERS, and what constitutes positive health benefits, can provide clearer guidance to researchers evaluating ERSs about what the expected outcomes are. This would also facilitate the update of evaluation frameworks, and aid the decision making of local commissioners who are evaluating the impact of an ERS in their area.

In the event of updating the ERS guidelines, however, it is important for a broad range of evidence to be included in the evidence review, including not only RCTs, but also, for example, high quality observational and qualitative studies. The value of different research study designs, in the evaluation of complex interventions, should also be considered by those aiming to identify and disseminate best practice in PA promotion. As previously discussed in section 2.5, PHE (2014b) have introduced ‘NESTA’ standards, which are based on academic standards of research (Puttick & Ludlow, 2013), to identify promising community-based PA interventions which could be scaled across the country to increase population PA levels. In the most recent submission round, only two of 127 ERS submissions could demonstrate ‘proven practice' by including a control or comparison group (PHE, 2018c), which highlights the potential disconnect in what types of evidence are possible for real-life ERSs to demonstrate. Consideration of what types of evaluation are feasible for community-based interventions to conduct (Gidlow et al., 2008), alongside mounting criticism of positivist approaches to evaluating public health interventions (Allmark, 2015; Pettman et al., 2012; Rutter et al., 2017), provides an argument for the reconsideration of how 'promising practice' is identified.

Furthermore, if a community PA intervention has been designed according to the local context, it cannot be assumed that it would have the same success in a different community. In the case of ERSs, it is also known that success can be at least in part attributed to aspects of the scheme such as the personality and approach of the exercise professionals (Morgan et al., 2016), which cannot necessarily be replicated. This raises a question as to whether policy should recommend the identification and replication of successful ERSs, or if it would be more valuable to identify the principles associated with successful ERSs, and to scale these principles, such as through a set of standards for ERSs.

ERSs are an example of implementation of key policy actions in both global and national policy. At a global level, the 2018 global action plan for PA, published by the WHO, recommends the incorporation of PA promotion in to health care settings, which includes the referral of individuals by health care professionals to community-based PA opportunities, and referral to supervised PA as part of the management of LTCs. Partnerships between health care services, and the sports and leisure sector are also advised, to support the provision of PA opportunities for clinical populations (WHO, 2018, action 3.2). Nationally, the sporting futures policy also recommends that the sports and leisure sector develop partnerships with health care providers, to integrate PA in to care pathways and maximise the potential of referral interventions

(Department for Culture, Media & Sport, 2015). Lastly, activating health care professionals to promote PA, and integrating PA in to care pathways is a key element of PHE's national framework for PA (PHE, 2014a, action 2).

Whilst systematic reviews have provided equivocal conclusions about the impact of ERSs, they have also provided an unfair assessment of impact (Beck et al., 2016). For a global and national policy action, the evidence base for ERSs is relatively small, and requires further local evaluations. However, the findings of this study, and other published evaluations demonstrating a positive long-term impact of ERSs (e.g. Martín-Borràs et al., 2018; Murphy et al., 2012), suggests that these policy actions are beneficial at increasing levels of health enhancing PA in previously inactive adults, and can have a long-term public health impact. In the UK, where it is imperative to address the current burden of LTCs, and other issues such as social isolation (Department of Health & Social Care, 2018a), there is an argument for further investment in ERSs to allow the referral of more individuals, to achieve a substantial contribution towards alleviating these burdens.

A large proportion of policy actions are directed at increasing PA levels at a general population level. However, the aforementioned policy actions are directed at promoting PA in populations with health conditions, who are often of older age, and potentially require more support than the general population to become active in a safe and sustainable manner. These populations are also more likely to be inactive than the general population (Sport England, 2018), strengthening the case for intervention. To-date, in the absence of ERSs, there are limited examples of these policy actions being implemented on a national scale for populations with health conditions, and therefore decommissioning of ERSs at this time could introduce an inequality in opportunity for PA amongst these groups. It is therefore recommended that PA referral-based policy actions continue to be included in PA policy, and implemented, until there is sufficient evidence that these actions are not justified.

To finish, ERSs are all but one solution to national levels of physical inactivity. They focus on a specific segment of the population and cannot be guaranteed to successfully address the problem of physical inactivity in this population alone. PA is argued to be a socio-ecological behaviour, influenced not just by individual capabilities, but environmental, organisational and policy determinants (Sallis et al., 2008). There is, therefore, no single policy solution to physical inactivity, it is a complex problem, with multi-sector policy action required (Pratt et al., 2015; Pate & Dowda, 2019). The efforts of ERSs and other PA initiatives can therefore be strengthened by not just policy recommendations, but policy implementation in a range of connected areas.

9.8. Conclusions

The aim of this thesis was to evaluate the effectiveness of the 'Live Active' ERS, at increasing PA levels and improving health outcomes in its participants. The research, conducted over a 3-year period, found that participation led to significant medium to long-term improvements in self-reported PA level and the majority of health-related outcomes. This demonstrates a potential long-term public health impact for the community of Tameside, which can alleviate the burden associated with LTCs. Further mixed methods research identified that participants benefited from taking part in a multitude of ways, which represented a holistic benefit of the scheme. Many of these benefits were not routinely evaluated by the scheme, or by previous evaluations, and nor were they guided to by existing evaluation frameworks. A myriad of limitations can be encountered when evaluating an ERS or seeking to generate a consensus on their impact nationally. The implications of this thesis are primarily concerned with, but not limited to, the content of evaluation frameworks used in this field, and how future evaluations are conducted, to take in to account the complex nature of ERSs, and to provide an evaluation that is representative of the scheme’s true impact. Moving forwards, successful ERSs can then be identified, and the mechanisms of their success explored, and translated in to methods of practice which ultimately contribute to substantially reducing national levels of physical inactivity, and the associated burden to health. Further investment in ERSs, to allow a greater number of individuals to be referred, could result in a much larger national impact.

Appendices

Appendix 1. Literature review search strategy

The following electronic databases were searched: Academic search premier (1900), AMED (1985), CINAHL (1900), Cochrane central, MEDLINE (1946), Psych extra (1908), Psych info (1806), Pubmed, SciELO Citation Index, SportDiscus (1900), Web of science core collection.

Keywords:
Databases were searched throughout the period of study using the keywords 'exercise referral' OR 'physical activity referral' OR 'GP referral' OR 'exercise on prescription' OR 'active lifestyle scheme' OR 'exercise intervention' OR 'physical activity intervention'. These terms were also coupled with key words including: evaluation, outcome, impact, benefits, uptake, adherence, barriers, facilitators, and physical activity measurement. Supplementary methods included checking the references of published articles, and citation tracking.

Inclusion criteria included:
- Published and peer-reviewed research studies or reviews, editorials, responses, and conference-proceedings available as abstracts or short papers.
- English language.
- Research studies of ERS interventions were included if they met the following criteria, based on the NICE (2014a) definition of an ERS: 1) referral of an individual who is sedentary and/or inactive with a LTC or significant health risk factor, by a healthcare professional, to a third-party PA service provider with the aim of achieving an increase in PA; 2) initial assessment of the individual by ERS staff and continued monitoring throughout the programme, and; 3) provision of a PA programme by the third party provider, tailored to individual needs.

Google and Google Scholar were also searched for physical activity related grey literature, practice guidelines and government reports.

Appendix 2. 'Live Active' ERS feedback questionnaire

1. How much have you enjoyed the programme so far?
    a. Not at all
    b. Ok
    c. Enjoyed
    d. Enjoyed thoroughly
2. Was your programme at an appropriate level?
    a. Yes
    b. No
3. How much do you feel your level of fitness has improved?
    a. Not at all
    b. A little
    c. Noticeable improvements
    d. Good improvements
4. Have you been less reliant on certain medications through being more active?
    a. Yes
        i. In what way have you been less reliant on certain medications through being more active? [open text]
    b. No
5. Have you experienced any other benefits so far?
    a. Improved confidence
    b. Made friends
    c. Reduction in pain
    d. Improved sleep
    e. Weight loss
    f. Improved self-esteem
    g. Reduction in stress
    h. Better sense of wellbeing
    i. Feel less isolated
    j. Ability to be more independent at home
6. What have you liked about the scheme so far?
    a. Support from instructor/s
    b. Ease of booking
    c. Programme/activities
    d. Meeting new people
    e. Price of activities
    f. Accessibility of venue
    g. Welcome within centres
    h. Other
7. Overall do you feel an improved quality of life from participating in the programme?
    a. Yes
    b. No
8. Is there anything you would like to improve about the scheme?
    a. Support from instructor/s
    b. Ease of booking
    c. Programme/activities
    d. Meeting new people
    e. Price of activities
    f. Accessibility of venue
    g. Welcome within centres
    h. Other

9. How helpful were the instructors?
    a. Not at all
    b. Ok
    c. Quite helpful
    d. Very helpful
10. Would you recommend the service to a friend?
    a. Yes
    b. No
11. Other comments [open text]

Appendix 3. Stage of change questionnaire

1. I am not currently very physically active and I don't intend to become more active in the next six months/I'm too busy right now (Pre-contemplation)
2. I am not currently very physically active, but I am thinking about increasing the amount of activity I take in the next six months (Contemplation)
3. The amount of activity I take varies: sometimes I am physically active, other times not (Preparation)
4. I am currently physically active on most days, but have only just begun to be so within the last six months (Action)
5. I am currently physically active on most days, and have been so for longer than six months (Maintenance)
6. A year ago I was physically active on most days, but in the last few months I have been less active (Relapse)

Appendix 4. IPAQ short form questionnaire

The next questions will ask you about the time you spent being physically active in the last 7 days. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the vigorous activities that you did in the last 7 days. Vigorous physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

1. During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling?

   _____ days per week

   ☐ No vigorous physical activities ➞ *Skip to question 3*

2. How much time did you usually spend doing vigorous physical activities on one of those days?

   _____ hours per day
   _____ minutes per day
   ☐ Don’t know/Not sure

Think about all the moderate activities that you did in the last 7 days. Moderate activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

3. During the last 7 days, on how many days did you do moderate physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.

   _____ days per week

   ☐ No moderate physical activities ➞ *Skip to question 5*

4. How much time did you usually spend doing moderate physical activities on one of those days?

_____ hours per day
_____ minutes per day
☐ Don't know/Not sure

Think about the time you spent walking in the last 7 days. This includes at work and at home, walking to travel from place to place, and any other walking that you might do solely for recreation, sport, exercise, or leisure.

5. During the last 7 days, on how many days did you walk for at least 10 minutes at a time?

_____ days per week

☐ No walking → *Skip to question 7*

6. How much time did you usually spend walking on one of those days?

_____ hours per day
_____ minutes per day
☐ Don't know/Not sure

The last question is about the time you spent sitting on weekdays during the last 7 days. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

7. During the last 7 days, how much time did you spend sitting on a weekday?

_____ hours per day
_____ minutes per day
☐ Don't know/Not sure

This is the end of the questionnaire, thank you for participating.

Appendix 5. EQ-5D-3L questionnaire

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

I have no problems in walking about ☐

I have some problems in walking about ☐

I am confined to bed ☐

Self-Care

I have no problems with self-care ☐

I have some problems washing or dressing myself ☐

I am unable to wash or dress myself ☐

Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities ☐

I have some problems with performing my usual activities ☐

I am unable to perform my usual activities ☐

Pain / Discomfort

I have no pain or discomfort ☐

I have moderate pain or discomfort ☐

I have extreme pain or discomfort ☐

Anxiety / Depression

I am not anxious or depressed ☐

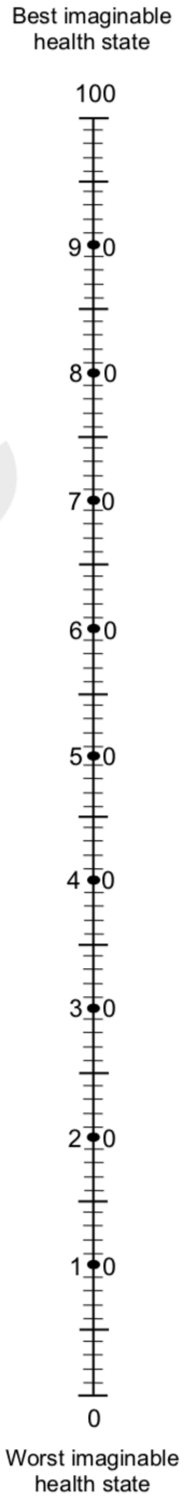
I am moderately anxious or depressed ☐

I am extremely anxious or depressed ☐

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

Your own health state today



Appendix 6. WEMWBS questionnaire

Below are some statements about feelings and thoughts.

Please tick (√) the box that best describes your experience of each over the **last 2 weeks**

STATEMENTS	None of the time	Rarely	Some of the time	Often	All of the time
I've been feeling optimistic about the future	1	2	3	4	5
I've been feeling useful	1	2	3	4	5
I've been feeling relaxed	1	2	3	4	5
I've been feeling interested in other people	1	2	3	4	5
I've had energy to spare	1	2	3	4	5
I've been dealing with problems well	1	2	3	4	5
I've been thinking clearly	1	2	3	4	5
I've been feeling good about myself	1	2	3	4	5
I've been feeling close to other people	1	2	3	4	5
I've been feeling confident	1	2	3	4	5
I've been able to make up my own mind about things	1	2	3	4	5
I've been feeling loved	1	2	3	4	5
I've been interested in new things	1	2	3	4	5
I've been feeling cheerful	1	2	3	4	5

Appendix 7. July 2016 – feedback questionnaire data

		Number	%
Programme at appropriate level	Yes	128	99.2
	No	1	0.8
Enjoyed the programme	Not at all	2	1.6
	Okay	14	10.9
	Enjoyed	43	33.3
	Enjoyed thoroughly	70	54.3
Fitness improvement	Not at all	4	3.1
	A little	34	26.4
	Noticeable improvement	50	38.8
	Good improvement	41	31.8
Quality of life has improved	Yes	112	86.8
	No	17	13.2
Less reliant on medication	Yes	18	14
	No	111	86
	Medications less reliant on: respiratory related (37.5%), pain related (31.3%), mental health related (12.5%), diabetes related (12.5%), CVD related (6.3%).		
Instructors were helpful	Not at all	-	-
	Okay	5	3.9
	Quite helpful	15	11.6
	Very helpful	109	84.5
Recommends the service	Yes	129	100
	No	-	-
Benefits	Improved confidence	34	26.4
	Made new friends	31	21.7
	Reduced pain	28	21.7
	Improved sleep	18	14
	Weight loss	52	40.3
	Improved self-esteem	26	20.2
	Reduced stress	18	14
	Better sense of wellbeing	49	38
	Reduced social isolation	9	7
	Able to be more independent	16	12.4
Liked	Support from instructors	89	69
	Ease of booking	13	10.1
	Programme/activities	43	33.3
	Meeting new people	33	25.6
	Price	33	25.6
	Accessibility	31	24
	Welcome in centres	30	23.3
	Other	4	3.1
Improve upon	Support from instructors	3	2.3
	Ease of booking	1	0.8
	Programme/activities	3	2.3
	Meeting new people	-	-
	Price	2	1.6
	Accessibility	-	-

		Number	%
	Welcome in centres	-	-
	Other	5	3.9

Appendix 8. PA and sedentary behaviour by quarter

		Week 1		Week 12	
Quarter	**Month**	**MET mins of PA per week**	**Sitting time per day (mins)**	**MET mins of PA per week**	**Sitting time per week (mins)**
4	April 2016	0	480	539	450
1	July 2016	0	480	636	420
2	October 2016	0	480	675	420
3	January 2017	0	480	612	480

Appendix 9. Uptake and adherence by quarter

Quarter	Month	Uptake %	Adherence %
4	April 2016	85	86
1	July 2016	82.5	83
2	October 2016	82.8	78.8
3	January 2017	83	76
4	April 2017	81	73
1	July 2017	80.9	70.7
2	October 2017	81	69.7
3	January 2018	80.7	69.2
4	April 2018	79.8	69.4
1	July 2018	78.9	67.6

Appendix 10. Interview guide

1. Could you tell me about how you came to be referred the Live Active scheme?
   [Probes]
   a) Who suggested that you should be referred here?
   b) What was is about your health and wellbeing that triggered the suggestion of referral?
   c) What was the intended outcome of you attending?

2. When you were initially referred to the scheme, before you came to your first consultation, what were your expectations of the scheme?
   [Probes]
   a) Did you understand what participation would involve?
   b) Did you think it would benefit you in a particular way?

3. Are you currently physically active?
   [Probes]
   a. If so, what are the activities?
   b. If not, why not?

4. Now that you've been with the Live Active service for over six months, could you tell me about how you think you've benefitted from attending?
   [Probes]
   a) Physical activity
   b) Physical, mental, social, condition specific
   c) Lifestyle
   d) Are these benefits important to them?

5. Are you satisfied with what you've got out of participation so far?
   [Probes]
   a) Are there any goals you haven't achieved?
   b) Are there some benefits you expected to see, which you haven't managed to achieve yet?
   c) Do you think that in the future you might achieve some of these (goals/benefits)?

6. Could you image where you would be today if you were not referred to this scheme?
   [Probes]
   a. Physically active or not?
   b. Improved health status or not?

7. We've come to the end of my questions, so at this point, I'd like to ask if there's anything else further you want to add or consider further, which you haven't mentioned so far?

Activity Monitor Instructions & Daily Log

Return Appointment: _____ on
_____/_____/_____, at _________________

If you have any questions or concerns, please contact

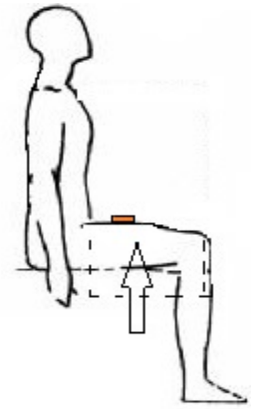
Faye Prior on ########### or email: #######@edu.salford.ac.uk

or

Live Active Office ###########

Activity Monitor Instructions

How do I attach the monitor?



1. Sit down on a chair when attaching the monitor so that your thigh is in a horizontal position. This will also make it easier to find the top of your thigh (the crease between your leg and your upper body).
2. The monitor is to be attached directly on to your skin, one third of the way down between the top of your **right** thigh and top of your knee. Position the monitor in the midline of your **right** thigh as shown in the picture.
3. Swab the area where the monitor is to be attached with the provided alcohol pad and let the area dry for a few seconds, so that the area is free from any oils that may cause the monitor to move.
4. Place the monitor in the correct position on the thigh (ensuring that the man on the front of the monitor is **standing up** when you stand up, so that the curved part of the monitor is at the top).
5. Peel the backing off the dressing and place it over the monitor. Press the dressing onto your skin, starting with the bottom, sides, then top, to smooth out the air bubbles and wrinkles as much as possible to ensure that the monitor is firmly secured to your thigh.

Wearing the monitor

- On the day that you receive the activity monitor, it will begin recording your activity at midnight.
- Please wear the activity monitor continuously for **7 days** (not including the day that you receive it), and remove it on the morning of day 8.
- The activity monitor can be worn during sleep and whilst showering, but **please remove it if you wish to take a bath or go swimming**, and re-attach with a fresh dressing after this activity.
- The dressings that stick the activity monitor to your skin may last up to 7 days, but to avoid skin irritation or water leakage, you may change the dressing regularly using the additional dressings provided.
- The activity monitor will emit a green flash every 6 seconds. This is an indication that it is working and recording data. Although it may not be visible through the waterproof dressing.

If you need to change the dressing

During your wear time, you may need to change the dressing which attaches the monitor to your thigh. To do this:

1. Peel off the dressing and remove the monitor from your thigh.
2. With an alcohol wipe provided in your pack, wipe down the area of your leg where the monitor was attached and follow the same procedure for applying the activity monitor as described in the above section.

What else do I need to do?

It is important that you fill in the **Daily Log** on the following pages, every day for the **7 days** while you are wearing the monitor.

Returning your activity monitor and daily log

After you have worn your activity monitor for 7 days, please return it to the research team at the time and date agreed (written on the front page of this booklet), along with your activity diary and any unused adhesive patches and alcohol wipes. If you become unable to return the activity monitor at this time, please contact the researcher, or call the Live Active office to arrange an alternative time (details on the front page).

Other notes

If you experience any skin irritations due to dressings, try to reattach the activity monitor as described above, or reattach the activity monitor to your left leg. If you still continue to experience irritation then take the monitor off and contact the researcher, or the Live Active office for further instructions (details on the front page).

How to fill in the daily log

- The log is divided into 7 days. Please complete each question for all of the seven days. Please try and be as accurate as possible—record the exact times if you can, or at least to the nearest 5 minutes of your estimated times.
- Start by writing the **date** in the top row.
- Record the time that you **woke up** and the time that you actually **got out of bed** (these times may be the same for some days). We ask for these two times because people sometimes spend time in bed before getting up, and we are interested in distinguishing between actual sleeping time and time in bed once awake, for example waking up and then watching TV for an hour before getting out of bed.
- At the end of the day, record what time you **got into bed** to go to sleep and then the time that you actually **went to sleep** time. (i.e., the estimated time that you fell asleep).
- Please record the time that you fell asleep first thing in the morning when you wake up, along with your wake time.
- If you **remove the activity monitor** for longer than 10minutes during the day please note down the time that you removed the device, the time that you re-attached it and the reason why you removed the device.
- There is also a space for you to make **comments**. It is useful for us to know if you have had any skin irritations, accidentally worn the monitor upside down or any other information that you think we should know.

Day and Date	**Time woke up**	**Time got out of bed**	**Did you remove your monitor for >10 mins today?**	**If removed, record time of removal and reason why**	**Time got into bed**	**Time went to sleep**	**Other comments**
Day 1 *17/12/2016*	*07:00am*	*07:15am*	[✓ *Yes*] *No*	*Time off: 18:00pm* *Time on:18:45pm* *Reason: Swimming in the sea*	*21:45pm*	*22:10pm*	*Slight irritation on right leg so put monitor on left leg*
Day 1 Date:			Yes No				
Day 2 Date:			Yes No				
Day 3 Date:			Yes No				
Day 4 Date:			Yes No				
Day 5 Date:			Yes No				
Day 6 Date:			Yes No				
Day 7 Date:			Yes No				

Appendix 12. Recommendations made to the 'Live Active' ERS

Excerpt from the 'Live Active' ERS final evaluation report.

Based on observations made over the 3-year evaluation period, and the results to emerge from this study, the following recommendations are offered to the scheme:

Referral processes

- At present, a proportion of non-uptake is attributable to inappropriate referrals. Ensure that referrers are aware of the eligibility criteria for referral, including the exclusion criteria. Provide feedback when inappropriate referrals have been received, to reduce the incidence of these referrals.
- Participants do not appear to be well informed of the scheme at the point of referral, and misinterpretations at this point may contribute to non-uptake and dropout. Ensure that referrers provide a sufficient explanation of the scheme to potential participants, so that they are informed of what to expect.
- At present, most referrers do not receive feedback about the effectiveness of the scheme for individuals they have referred. Consider offering feedback reports (e.g. every six months) to referrer organisations, detailing the engagement levels and outcomes of participants they have referred (at the population level). This can provide insight, and may encourage sustained partnerships based on the positive outcomes reported. Alternatively, a short annual report of the scheme may be sufficiently informative.
- The results of this study suggest that participation is particularly beneficial for those with musculoskeletal conditions, mental health issues, or those who are socially isolated and/or lonely. The targeted referral of these groups is therefore recommended, given the benefits of participation for these groups, and the current national imperative to reduce the burden of these issues.

Additional staff training

- Consider implementing additional staff training in areas such as behaviour change, to facilitate the retention of participants at risk of dropout. This includes, for example: those who are from more deprived areas, and may experience barriers related to inequalities; those who are referred or enter the scheme during the winter months, when the weather or ill health may present a barrier, and; those who perceive they cannot be active in respect of health problems, such as anticipated aggravation of musculoskeletal pain.
- At present, goals appear to be mostly descriptive and non-specific. Ensure that goals agreed between the exercise professional and participant are specific, measurable, attainable, realistic and timely, to facilitate attainment and to provide participants with an outcome they can measure progress towards.

Data management

- Consider reducing the number of outcomes collected, to reduce overlap of themes, increase compliance with data collection, and increase time within the consultation for physical activity counselling.
    - Retention of the IPAQ-SF and EQ-5D are recommended as a minimum, or replacement with similar questionnaires.
    - Replacement of the EQ-5D-3L with the EQ-5D-5L is recommended to

  reduce the ceiling effect and obtain more insightful outcomes.
  - Measurement of body mass index and blood pressure have low compliance, and are influenced by external factors. Measurement is recommended for participants in which these outcomes are relevant, or as part of a risk assessment screening.
  - Collection of the WEMWBS has low compliance. Observations suggest that this is, in-part, due to time restraints. Therefore, replacement with the WEMWBS short form is recommended, if this outcome is to be continued to be measured.
- The results of this study indicate that there are a wide range of perceived benefits to participation, many of which are not currently captured by the existing data collection. In the feedback questionnaire, consider expanding the tick box list of benefits of participation (e.g. physical activity knowledge, mobility, diet, return to work), or re-configure to allow an open text box response for each additional benefit reported.

Programme and activities

- Participants in this scheme appear to enjoy the social elements of participation, which can facilitate adherence to physical activity. Therefore, it is recommended that the formation of social networks and social support within 'Live Active' classes and gym settings is facilitated, and that participants are encouraged to attend group-based activities in the first weeks of participation. This may facilitate enjoyment and self-efficacy for physical activity in the first weeks of participation when the majority of dropout occurs, promote positive social health, and ultimately facilitate adherence to the scheme, and physical activity beyond completion.
- Group-based physical activity classes in leisure facilities are often targeted towards the general population. Therefore, at the point of scheme completion, suitable opportunities to continue group-based physical activities may be limited. Therefore, it is recommended to introduce maintenance classes, or other group-based classes suitable for an ERS population, to the general class timetable at Active Tameside sites, to provide opportunities for group-based physical activity when a participant has completed the Live Active scheme.
- Results from the feedback questionnaire, and semi-structured interviews, indicate that some participants find the 'Live Active' class and gym timetable restrictive. This is a particular concern for participants who are employed full-time, or do not wish to attend classes in the evening. Therefore, it is recommended that the 'Live Active' class timetable is expanded, to provide the core classes more than once per week, ideally during both off-peak and peak-time hours, to increase opportunities for attendance.

Appendix 13. Dissemination of research findings

Peer reviewed papers

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2019). Long term health outcomes associated with an exercise referral scheme: an observational longitudinal follow-up study. *Journal of Physical Activity and Health*. Awaiting publication.

Papers to be submitted

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2019). Levels and predictors of engagement in an exercise referral scheme and reasons for non-uptake and dropout: an evaluation of the Live Active exercise referral scheme.

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2019). A qualitative exploration of the perceived benefits of participation in an exercise referral scheme.

Unpublished reports

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2018). Live Active exercise referral scheme evaluation report.

Invited talks

Prior, F. (2018). *The benefits of physical activity and approaches to promotion: a case study of the 'Live Active' exercise referral scheme*. Presented at the University of Salford Public Health Research Seminar Series. Salford, UK.

Prior, F. (2018). *A case study of the 'Live Active' exercise referral scheme*. Presented at the GM Active January board meeting. Greater Manchester, UK.

Prior, F. (2017). *Long-term outcomes of the 'Live Active' exercise referral scheme*. Presented at the Active Tameside annual celebration event. Tameside, UK.

Prior, F. (2017). *Evaluation of the 'Live Active' exercise referral scheme*. Presented at the University of Salford Equity, Health & Wellbeing research group meeting. Salford, UK.

Prior, F. (2016). *Exercise referral schemes: what does the evidence say? A case study of the 'Live Active' exercise referral scheme*. Presented at the GM Active June exercise referral workshop. Greater Manchester, UK.

Conference presentations

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2018). *Comparison of self-reported and device-measured physical activity in a sample of exercise referral participants*. Presented at the British Association of Sport and Exercise Sciences Physical Activity Division Day. Coventry, UK.

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2017). *Evaluation of an exercise referral scheme in the United Kingdom: medium term outcomes*. Presented at 8th Conference of the HEPA Europe. Zagreb, Croatia.

Conference posters

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2018). *Evaluation of the 'Live Active' exercise referral scheme: long term outcomes*. Poster presented at the 7th International Society for Physical Activity and Health Congress. London, UK.

Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2018). *Perceived benefits of participation in the 'Live Active' exercise referral scheme*. Poster presented at the 7th

International Society for Physical Activity and Health Congress. London, UK.

Roberts, K., Prior, F., Coffey, M., Robins, A., & Cook, P. A. (2017). *Live Active exercise referral scheme*. Poster presented at the UKactive promising practice roadshow. Greater Manchester, UK.

References

Adams, J., & White, M. (2003). Are activity promotion interventions based on the transtheoretical model effective? A critical review. *British Journal of Sports Medicine*, *37*(2), 106–114. doi: 10.1136/bjsm.37.2.106

Adams, J., & White, M. (2004). Why don’t stage-based activity promotion interventions work? *Health Education Research*, *20*(2), 237–243. doi: 10.1093/her/cyg105

Adams, S., Matthews, C., Ebbeling, C., Moore, C., Cunningham, J., Fulton, J., & Hebert, J. (2005). The effect of social desirability and social approval on self-reports of physical activity. *American Journal of Epidemiology*, *161*(4), 389–398. doi: 10.1093/aje/kwi054

Age UK. (2017). ‘Grannannying’ is the word as five million grandparents take on childcare responsibilities. Retrieved from https://www.ageuk.org.uk/bp-assets/globalassets/surrey/original-blocks/about-us/latest-news/age-uk--age-uk-surrey-press-releases/grannannying-is-the-word.pdf

Ainsworth, B. (2009). How do I measure physical activity in my patients? Questionnaires and objective methods. *British Journal of Sports Medicine*, *43*(1), 6-9. doi: 10.1136/bjsm.2008.052449

Ainsworth, B., Cahalin, L., Buman, M., & Ross, R. (2015). The current state of physical activity assessment tools. *Progress in Cardiovascular Diseases*, *57*(4), 387–395. doi: 10.1016/j.pcad.2014.10.005

Ainsworth, B., Caspersen, C., Matthews, C., Mâsse, L., Baranowski, T., & Zhu, W. (2012). Recommendations to improve the accuracy of estimates of physical activity derived from self report. *Journal of Physical Activity & Health*, *9*(S1), 76–S84.

Ajzen, I. (1991). The theory of planned behavior. *Theories of Cognitive Self-Regulation*, *50*(2), 179–211. doi: 10.1016/0749-5978(91)90020-T

Akobeng, A. (2005). Understanding randomised controlled trials. *Archives of Disease in Childhood*, *90*, 840-844. doi: 10.1136/adc.2004.058222

Ali, S., & Rouse, A. (2002). Practice audits: reliability of sphygmomanometers and blood pressure recording bias. *Journal of Human Hypertension*, *16*(5), 359–361. doi: 10.1038/sj.jhh.1001384

Allmark, P. (2015). Health promotion in public health: philosophical analysis. In T. Schramme & S. Edwards (Eds.), *Handbook of the philosophy of medicine* (pp. 1–20). Dordrecht: Springer.

Altheide, D., & Johnson, J. (1994). Criteria for assessing interpretive validity in qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 485-499). Thousand Oaks: SAGE Publications.

Altschuler, A., Picchi, T., Nelson, M., Rogers, J., Hart, J., & Sternfeld, B. (2009). Physical activity questionnaire comprehension-lessons from cognitive interviews. *Medicine and Science in Sports and Exercise*, *41*(2), 336–343. doi: 10.1249/MSS.0b013e318186b1b1

American College of Sports Medicine. (2017). *ACSM’s guidelines for exercise testing and prescription* (10th ed.). Philadelphia: Lippincott Williams and Wilkins.

Andreoli, A., Garaci, F., Cafarelli, F., & Guglielmi, G. (2016). Body composition in clinical practice. *European Journal of Radiology*, *85*(8), 1461–1468. doi: 10.1016/j.ejrad.2016.02.005

Anokye, N., Trueman, P., Green, C., Pavey, T., Hillsdon, M., & Taylor, R. (2011). The cost-effectiveness of exercise referral schemes. *BMC Public Health*, *11*, 954. doi:10.1186/1471-2458-11-954

Arem, H., Moore, S., Patel, A., Hartge, P., de Gonzalez, A., Visvanathan, K., …

Matthews, C. (2015). *Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship*. *JAMA Internal Medicine*, *175*(6), 959-967. doi: 10.1001/jamainternmed.2015.0533

Arsenijevic, J., & Groot, W. (2017). Physical activity on prescription schemes (PARS): do programme characteristics influence effectiveness? Results of a systematic review and meta-analyses. *BMJ Open*, *7*, e012156. doi: 10.1136/bmjopen-2016-012156

Arthritis Research UK. (2018). *State of musculoskeletal health 2018*. Retrieved from https://www.arthritisresearchuk.org/arthritis-information/data-and-statistics/state-of-musculoskeletal-health.aspx

Ashe, M., Miller, W., Eng, J., & Noreau, L. (2009). Older adults, chronic disease and leisure-time physical activity. *Gerontology*, *55*(1), 64–72. doi: 10.1159/000141518

Ashford, S., Edmunds, J., & French, D. (2010). What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. *British Journal of Health Psychology*, *15*(2), 265–288. doi: 10.1348/135910709X461752

Aune, D., Norat, T., Leitzmann, M., Tonstad, S., & Vatten, L. (2015). Physical activity and the risk of type 2 diabetes: a systematic review and dose–response meta-analysis. *European Journal of Epidemiology*, *30*(7), 529–542. doi: 10.1007/s10654-015-0056-z

Aune, D., Sen, A., Prasad, M., Norat, T., Janszky, I., Tonstad, S., … Vatten, L. (2016). BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ*, *353*, i2156. doi: 10.1136/bmj.i2156

Bahrke, M., & Morgan, W. (1978). Anxiety reduction following exercise and meditation. *Cognitive Therapy and Research*, *2*(4), 323–333. doi: 10.1007/BF01172650

Baker, C. (2018). *Briefing paper: obesity statistics*. Retrieved from https://researchbriefings.files.parliament.uk/documents/SN03336/SN03336.pdf

Bamberger, M., Rao, V., & Woolcock, M. (2010). *Using mixed methods in monitoring and evaluation: experiences from international development*. Retrieved from https://openknowledge.worldbank.org/handle/10986/3732

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.

Baranowski, T. (1988). Validity and reliability of self report measures of physical activity: an information-processing perspective. *Research Quarterly for Exercise and Sport*, *59*(4), 314–327. doi: 10.1080/02701367.1988.10609379

Baranowski, T., & Stables, G. (2000). Process evaluations of the 5-a-day projects. *Health Education & Behavior*, *27*(2), 157–166. doi: 10.1177/109019810002700202

Barbour, R. (2000). The role of qualitative research in broadening the ‘evidence base’ for clinical practice. *Journal of Evaluation in Clinical Practice*, *6*(2), 155–163. doi: 10.1046/j.1365-2753.2000.00213.x

Barry, G., Galna, B., Lord, S., Rochester, L., & Godfrey, A. (2015). Defining ambulatory bouts in free-living activity: Impact of brief stationary periods on bout metrics. *Gait & Posture*, *42*(4), 594–597. doi: 10.1016/j.gaitpost.2015.07.062

Bartram, D., Yadegarfar, G., Sinclair, J., & Baldwin, D. (2011). Validation of the Warwick–Edinburgh Mental Well-being Scale (WEMWBS) as an overall indicator of population mental health and well-being in the UK veterinary profession. *The Veterinary Journal*, *187*(3), 397–398. doi: 10.1016/j.tvjl.2010.02.010

Baum, F. (1995). Researching public health: behind the qualitative-quantitative methodological debate. *Social Science & Medicine*, *40*(4), 459–468. doi: 10.1016/0277-9536(94)E0103-Y

Bauman, A., Reis, R., Sallis, J., Wells, J., Loos, R., & Martin, B. (2012). Correlates of physical activity: why are some people physically active and others not? *The Lancet*, *380*(9838), 258–271. doi: 10.1016/S0140-6736(12)60735-1

Beck, F., Gillison, F., Koseva, M., Standage, M., Brodrick, J., Graham, C., & Young, H. (2016). The systematic identification of content and delivery style of an exercise intervention. *Psychology & Health*, *31*(5), 605–621. doi: 10.1080/08870446.2015.1132718

Berger, P., & Luckmann, T. (1966). *The social construction of reality: a treatise in the sociology of knowledge*. London: Penguin Books Ltd.

Bergman, P. (2018). The number of repeated observations needed to estimate the habitual physical activity of an individual to a given level of precision. *PLOS ONE*, *13*(2), e0192117. doi: 10.1371/journal.pone.0192117

Bertaux, D. (1981). From the life-history approach to the transformation of sociological practice. In D. Bertaux (Eds.), *Biography and society: the life history approach in the social sciences* (pp. 29-46). London: SAGE Publications.

Bharmal, M., & Thomas, J. (2006). Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. *Value in Health*, *9*(4), 262–271. doi: 10.1111/j.1524-4733.2006.00108.x

Blackmore, C., Medina, L., Ravenel, J., Silvestri, G., & Applegate, K. (2011). Critically assessing the literature: understanding error and bias. In L. Medina, C. Blackmore, & K. Applegate (Eds.), *Evidence-based imaging: improving the quality of imaging in patient care* (pp. 19–26). New York: Springer.

Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310. doi: 10.1016/S0140-6736(86)90837-8

Blondell, S., Hammersley-Mather, R., & Veerman, J. (2014). Does physical activity prevent cognitive decline and dementia? A systematic review and meta-analysis of longitudinal studies. *BMC Public Health*, *14*(1), 1–12. doi: 10.1186/1471-2458-14-510

Boon, R., Hamlin, M., Steel, G., & Ross, J. (2010). Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with accelerometry. *British Journal of Sports Medicine*, *44*(10), 741–746. doi: 10.1136/bjsm.2008.052167

Booth, F., Roberts, C., & Laye, M. (2012). Lack of exercise is a major cause of chronic diseases. *Comprehensive Physiology*, *2*(2), 1143–1211. doi: 10.1002/cphy.c110025

Bordieu, P. (1977). *Outline of a theory of practice*. Cambridge: Cambridge University Press.

Bordieu, P. (1986). The forms of capital. In J. Richardson (Eds.), *Handbook of theory and research for the sociology of education* (pp.241-258). New York: Greenwood Press.

Börjesson, M., Onerup, A., Lundqvist, S., & Dahlöf, B. (2016). Physical activity and exercise lower blood pressure in individuals with hypertension: narrative review of 27 RCTs. *British Journal of Sports Medicine*, *50*(6), 356-361. doi: 10.1136/bjsports-2015-095786

Bornstein, M., Jager, J., & Putnick, D. (2013). Sampling in developmental science: situations, shortcomings, solutions, and standards. *Developmental Review,* *33*(4), 357–370. doi: 10.1016/j.dr.2013.08.003

Borrell, L. (2014). The effects of smoking and physical inactivity on advancing mortality in U.S. adults. *Annals of Epidemiology*, *24*(6), 484–487. doi: 10.1016/j.annepidem.2014.02.016

Boslaugh, S. (2007). *Secondary data sources for public health: a practical guide*. Cambridge: Cambridge University Press.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*(3), 281–291. doi: 10.1093/pubmed/fdi031

Bowling, A. (2014). *Research methods in health: investigating health and health services* (4th ed.). Maidenhead: Open University Press.

Boyle, P., Barnes, L., Buchman, A., & Bennett, D. (2009). Purpose in life is associated with mortality among community-dwelling older persons. *Psychosomatic Medicine*, *71*(5), 574–579. doi: 10.1097/PSY.0b013e3181a5a7c0

Brannen, J. (1992). *Mixing methods: qualitative and quantitative research*. Aldershot: Ashgate Publishing Limited.

Brawner, C., Churilla, J., & Keteyian, S. (2016). Prevalence of physical activity is lower among individuals with chronic disease. *Medicine and Science in Sports and Exercise*, *48*(6), 1062–7. doi: 10.1249/MSS.0000000000000861

Brewer, J., & Hunter, A. (1989). *Multimethod research: a synthesis of styles*. Newbury Park: SAGE Publications.

Bridle, C., Riemsma, R., Pattenden, J., Sowden, A., Mather, L., Watt, I., & Walker, A. (2005). Systematic review of the effectiveness of health behavior interventions based on the transtheoretical model. *Psychology & Health*, *20*(3), 283–301. doi: 10.1080/08870440512331333997

British Heart Foundation National Centre. (2010). *A toolkit for the design, implementation & evaluation of exercise referral schemes. A Snapshot of ER schemes operating in England, Scotland & Northern Ireland - 2006-2008.* Retrieved from http://www.bhfactive.org.uk/sites/Exercise-Referral-Toolkit/downloads/s2-current-practice.pdf

Brook, R., Weder, A., & Rajagopalan, S. (2011). “Environmental hypertensionology” the effects of environmental factors on blood pressure in clinical practice and research. *The Journal of Clinical Hypertension*, *13*(11), 836–842. doi: 10.1111/j.1751-7176.2011.00543.x

Brown, W., Bauman, A., Bull, F., & Burton, N. (2012). Development of evidence-based physical activity recommendations for adults (18-64 years). Retrieved from http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines/$File/DEB-PAR-Adults-18-64years.pdf

Brownson, R., Boehmer, T., & Luke, D. (2004). Declining rates of physical activity in the United States: what are the contributors? *Annual Review of Public Health*, *26*(1), 421–443. doi: 10.1146/annurev.publhealth.26.021304.144437

Bryman, A. (1984). The debate about quantitative and qualitative research: a question of method or epistemology? *The British Journal of Sociology*, *35*(1), 75–92.

Bryman, A. (2001). *Social research methods*. Oxford: Oxford University Press.

Buchan, D., Ollis, S., Thomas, N., & Baker, J. (2012). Physical activity behaviour: an overview of current and emergent theoretical practices. *Journal of Obesity*, *2012*, 1–11. doi: 10.1155/2012/546459

Busschaert, C., De Bourdeaudhuij, I., Van Holle, V., Chastin, S., Cardon, G., & De Cocker, K. (2015). Reliability and validity of three questionnaires measuring context-specific sedentary behaviour and associated correlates in adolescents, adults and older adults. *International Journal of Behavioral*

Nutrition and Physical Activity, *12*(1), 117. doi: 10.1186/s12966-015-0277-2

Butte, N., Ekelund, U., & Westerterp, K. (2012). Assessing physical activity using wearable monitors: measures of physical activity. *Medicine & Science in Sports & Exercise*, *44*(S1), 5-12. doi: 10.1249/MSS.0b013e3182399c0e

Campbell, F., Holmes, M., Everson-Hock, E., Davis, S., Buckley Woods, H., Anokye, N., … Kaltenthaler, E. (2015). A systematic review and economic evaluation of exercise referral schemes in primary care: a short report. *Health Technology Assessment, 19*(60), 1–110. doi: 10.3310/hta19600

Canadian Society for Exercise Physiology. (2011). *Canadian physical activity guidelines clinical practice guideline development report*. Retrieved from http://csepguidelines.ca/wp-content/uploads/2018/04/CPAGuideline_Report_JAN2011.pdf

Caracelli, V., & Greene, J. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *15*(2), 195–207.

Carey, K., Purnine, D., Maisto, S., & Carey, M. (1999). Assessing readiness to change substance abuse: a critical review of instruments. *Clinical Psychology: Science and Practice*, *6*(3), 245–266. doi: 10.1093/clipsy.6.3.245

Casebeer, A., & Verhoef, M. (1997). Combining qualitative and quantitative research methods: considering the possibilities for enhancing the study of chronic diseases. *Chronic Diseases in Canada*, *18*(3), 130–135.

Caspersen, C., Powell, K., & Christenson, G. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126–131.

Cavill, N., Roberts, K., & Rutter, H. (2012). *Standard evaluation framework for physical activity interventions*. Retrieved from http://www.noo.org.uk/uploads/doc/vid_16722_SEF_PA.pdf

Centre for Economics and Business Research. (2015). *The economic cost of physical inactivity in Europe: an ISCA/CEBR report*. Retrieved from http://inactivity-time-bomb.nowwemove.com/download-report/The Economic Costs of Physical Inactivity in Europe (June 2015).pdf

Cerin, E., Cain, K., Oyeyemi, A., Owen, N., Conway, T., Cochrane, T., … Sallis, J. (2016). Correlates of agreement between accelerometry and self-reported physical activity. *Medicine & Science in Sports & Exercise*, *48*(6), 1075–1084. doi: 10.1249/MSS.0000000000000870

Chastin, S., Culhane, B., & Dall, P. (2014). Comparison of self-reported measure of sitting time (IPAQ) with objective measurement (activPAL). *Physiological Measurement*, *35*(11), 2319-2328. Doi: 10.1088/0967-3334/35/11/2319

Chastin, S., Dontje, M., Skelton, D., Čukić, I., Shaw, R., Gill, J., … Dall, P. (2018). Systematic comparative validation of self-report measures of sedentary time against an objective measure of postural sitting (activPAL). *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 21. doi: 10.1186/s12966-018-0652-x

Chatterjee, R., Chapman, T., Brannan, M., & Varney, J. (2017). GPs' knowledge, use, and confidence in national physical activity and health guidelines and tools: a questionnaire-based survey of general practice in England. *British Journal of General Practice*, *67*(663), e668-e675. doi: 10.3399/bjgp17X692513

Chen, K., Janz, K., Zhu, W., & Brychta, R. (2012). Redefining the roles of sensors in objective physical activity monitoring. *Medicine and Science in Sports and Exercise*, *44*(S1), 13–23. doi: 10.1249/MSS.0b013e3182399bc8

Chodzko-Zajko, W., Proctor, D., Fiatarone Singh, M., Minson, C., Nigg, C., Salem, G., & Skinner, J. (2009). Exercise and physical activity for older adults. *Medicine & Science in Sports & Exercise*, *41*(7), 1510–30. doi: 10.1249/MSS.0b013e3181a0c95c

Choi, B., & Pak, A. (2005). A catalog of biases in questionnaires. *Preventing Chronic Disease*, *2*(1), 1–13.

Clarke, P., & Yaros, P. (1988). Research blenders: commentary and response. *Nursing Science Quarterly*, *1*(4), 147–149. doi: 10.1177/089431848800100406

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). Abingdon: Routledge.

Colberg, S., Sigal, R., Yardley, J., Riddell, M., Dunstan, D., Dempsey, P., … Tate, D. (2016). Physical activity/exercise and diabetes: a position statement of the American Diabetes Association. *Diabetes Care*, *39*(11), 2065-2079. doi: 10.2337/dc16-1728

Coleman, A., Steel, S., Ashworth, M., Vowler, S., & Shennan, A. (2005). Accuracy of the pressure scale of sphygmomanometers in clinical use within primary care. *Blood Pressure Monitoring*, *10*(4), 181–188.

Collins, K., Onwuegbuzie, A., & Jiao, Q. (2006). Prevalence of mixed-methods sampling designs in social science research. *Evaluation & Research in Education*, *19*(2), 83–101. doi: 10.2167/eri421.0

Connor Gorber, S., Tremblay, M., Moher, D., & Gorber, B. (2007). A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Reviews*, *8*(4), 307–326. doi: 10.1111/j.1467-789X.2007.00347.x

Cooper, D., & Schindler, P. (2013). *Business research methods* (12th ed.). New York: McGraw-Hill Irwin.

Corbetta, P. (2003). *Social research: theory, methods and techniques*. London: SAGE Publications.

Cornelissen, V., & Smart, N. (2013). Exercise training for blood pressure: a systematic review and meta-analysis. *Journal of the American Heart Association*, *2*(1), e004473. doi: 10.1161/JAHA.112.004473

Cornier, M., Després, J., Davis, N., Grossniklaus, D., Klein, S., Lamarche, B., … Poirier, P. (2011). Assessing adiposity: a scientific statement from the American Heart Association. *Circulation*, *124*(18), 1996–2019. doi: 10.1161/CIR.0b013e318233bc6a

Cotterell, N., Buffel, T., & Phillipson, C. (2018). Preventing social isolation in older people. *Maturitas*, *113*, 80–84. doi: 10.1016/j.maturitas.2018.04.014

Courneya, K., Forbes, C., Trinh, L., Sellar, C., Friedenreich, C., & Reiman, T. (2013). Patient satisfaction with participation in a randomized exercise trial: effects of randomization and a usual care posttrial exercise program. *Clinical Trials*, *10*(6), 959–966. doi: 10.1177/1740774513495985

Craft, L., & Perna, F. (2004). The benefits of exercise for the clinically depressed. *Primary Care Companion to The Journal of Clinical Psychiatry*, *6*(3), 104–111.

Craig, B., Pickard, A., & Lubetkin, E. (2014). Health problems are more common, but less severe when measured using newer EQ-5D versions. *Journal of Clinical Epidemiology*, *67*(1), 93–99. doi: 10.1016/j.jclinepi.2013.07.011

Craig, C., Marshall, A., Sjöström, M., Bauman, A., Booth, M., Ainsworth, B., … Oja, P. (2003). International Physical Activity Questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381-1395.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research

Council guidance. *The BMJ*, *337*, a1655. https://doi.org/10.1136/bmj.a1655

Crespo, C., Keteyian, S., Snelling, A., Smit, E., & Anderson, R. (1999). Prevalence of no leisure-time physical activity in persons with chronic disease: results from the third national health and nutrition examination survey. *Clinical Exercise Physiology*, *1*(2), 68-73.

Creswell, J. (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks: SAGE Publications.

Creswell, J. (2003). *Research design: qualitative, quantitative and mixed methods approaches* (2nd ed.). Thousand Oaks: SAGE Publications.

Creswell, J., & Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks: SAGE Publications.

Crone, D., Smith, A., & Gough, B. (2005). 'I feel totally at one, totally alive and totally happy': a psycho-social explanation of the physical activity and mental health relationship. *Health Education Research*, *20*(5), 600–11. doi: 10.1093/her/cyh007

Crone, D., Johnston, L., Gidlow, C., Henley, C., & James, D. (2008). Uptake and participation in physical activity referral schemes in the UK: an investigation of patients referred with mental health problems. *Issues in Mental Health Nursing*, *29*(10), 1088–1097. doi: 10.1080/01612840802319837

Curtis, K., Fulton, E., & Brown, K. (2018). Factors influencing application of behavioural science evidence by public health decision-makers and practitioners, and implications for practice. *Preventive Medicine Reports*, *12*, 106–115. doi: 10.1016/j.pmedr.2018.08.012

Dahlgren, G., Carlsson, D., Moorhead, A., Häger-Ross, C., & McDonough, S. (2010). Test–retest reliability of step counts with the ActivPAL^{TM} device in common daily activities. *Gait & Posture*, *32*(3), 386–390. doi: 10.1016/j.gaitpost.2010.06.022

Dale, A. (2004). *The SAGE encyclopedia of social science research methods*. Thousand Oaks: SAGE Publications.

Darden, D., Richardson, C., & Jackson, E. (2013). Physical activity and exercise for secondary prevention among patients with cardiovascular disease. *Current Cardiovascular Risk Reports*, *7*(6), 411-416. doi: 10.1007/s12170–013–0354–5

Davidson, S., & Rossall, P. (2015). Evidence review: loneliness in later life. Retrieved from https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/reports-and-briefings/health--wellbeing/rb_june15_lonelines_in_later_life_evidence_review.pdf

Davis, R., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2015). Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychology Review*, *9*(3), 323–344. doi: 10.1080/17437199.2014.941722

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, *210*, 2–21. doi: 10.1016/j.socscimed.2017.12.005

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Denscombe, M. (2007). *The good research guide: for small-scale social research projects* (3rd ed.). Buckingham: Open University Press.

Denscombe, M. (2010). *The good research guide: for small-scale social research projects* (4th ed.). Berkshire: Open University Press.

Denzin, N., & Lincoln, Y. (1994). *Handbook of qualitative research*. Thousand Oaks: SAGE Publications.

Denzin, N., & Lincoln, Y. (2017). *The SAGE handbook of qualitative research* (5th ed.). Thousand Oaks: SAGE Publications.

Department for Culture, Media & Sport. (2015). *Sporting future: a new strategy for an active nation*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/486622/Sporting_Future_ACCESSIBLE.pdf

Department for Digital, Culture, Media & Sport. (2018). *A connected society: a strategy for tackling loneliness - laying the foundations for change*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/750909/6.4882_DCMS_Loneliness_Strategy_web_Update.pdf

Department of Health. (1999). *Saving lives: our healthier nation*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/265576/4386.pdf

Department of Health. (2001). *Exercise referral systems: a national quality assurance framework*. Retrieved from http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4079009.pdf

Department of Health. (2004). *At least five a week. Evidence on the impact of physical activity and its relationship to health*. Retrieved from http://webarchive.nationalarchives.gov.uk/20130105001829/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4080981.pdf

Department of Health. (2006). *Our health, our care, our say: a new direction for community services*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/272238/6737.pdf

Department of Health. (2007). *DH statement on exercise referral*. Retrieved from http://webarchive.nationalarchives.gov.uk/20080817183313/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_072689

Department of Health. (2011). *start active, stay active: a report on physical activity from the four home countries' Chief Medical Officers*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216370/dh_128210.pdf

Department of Health. (2012). *Long term conditions compendium of information: third edition*. Retrieved from https://www.gov.uk/government/publications/long-term-conditions-compendium-of-information-third-edition

Department of Health & Social Care. (2018a). *The government's mandate to NHS England for 2018-19*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/691998/nhse-mandate-2018-19.pdf

Department of Health & Social Care. (2018b). *Social prescribing schemes to be funded by the health and wellbeing fund: 2018*. Retrieved from https://www.gov.uk/government/publications/social-prescribing-schemes-to-be-funded-by-the-health-and-wellbeing-fund-2018

Devlin, N., Parkin, D., & Browne, J. (2010). Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Economics*, *19*(8), 886–905. doi: 10.1002/hec.1608

Diaz, K., & Shimbo, D. (2013). Physical activity and the prevention of hypertension.

Current Hypertension Reports, *15*(6), 659–668. doi: 10.1007/s11906-013-0386-8

DiClemente, C., & Prochaska, J. (1982). Self-change and therapy change of smoking behavior: a comparison of processes of change in cessation and maintenance. *Addictive Behaviors*, *7*(2), 133–142. doi: 10.1016/0306-4603(82)90038-7

Din, N., Moore, G., Murphy, S., Wilkinson, C., & Williams, N. (2015). Health professionals' perspectives on exercise referral and physical activity promotion in primary care: findings from a process evaluation of the national exercise referral scheme in Wales. *Health Education Journal*, *74*(6), 743–757. doi: 10.1177/0017896914559785

Dobson, F., Bennell, K., French, S., Nicolson, P., Klaasman, R., Holden, M., … Hinman, R. (2016). Barriers and facilitators to exercise participation in people with hip and/or knee osteoarthritis: synthesis of the literature using behavior change theory. *American Journal of Physical Medicine & Rehabilitation*, *95*(5), 372-389. Doi: 10.1097/PHM.0000000000000448

Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care*, *35*(11), 1095–1108.

Donaldson, G., & Wedzicha, J. (2014). The causes and consequences of seasonal variation in COPD exacerbations. *International Journal of Chronic Obstructive Pulmonary Disease*, *9*, 1101–1110. doi: 10.2147/COPD.S54475

Donders, A., van der Heijden, G., Stijnen, T., & Moons, K. (2006). Review: a gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, *59*(10), 1087–1091. doi: 10.1016/j.jclinepi.2006.01.014

Doolan, D., & Froelicher, E. (2009). Using an existing data set to answer new research questions: a methodological review. *Research and Theory for Nursing Practice*, *23*(3), 203-215.

Dopico, X., Evangelou, M., Ferreira, R., Guo, H., Pekalski, M., Smyth, D., … Todd, J. (2015). Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nature Communications*, *6*, 7000. doi: 10.1038/ncomms8000

Dowd, K., Harrington, D., & Donnelly, A. (2012). Criterion and concurrent validity of the activPAL^{TM} professional physical activity monitor in adolescent females. *PLoS ONE*, *7*(10), e47633. doi: 10.1371/journal.pone.0047633

Dowd, K., Szeklicki, R., Minetto, M., Murphy, M., Polito, A., Ghigo, E., … Donnelly, A. (2018). A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. *The International Journal of Behavioral Nutrition and Physical Activity*, *15*, 15. doi: 10.1186/s12966-017-0636-2

Duda, J., Williams, G., Ntoumanis, N., Daley, A., Eves, F., Mutrie, N., … Jolly, K. (2014). Effects of a standard provision versus an autonomy supportive exercise referral programme on physical activity, quality of life and well-being indicators: a cluster randomised controlled trial. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 10. doi: 10.1186/1479-5868-11-10

Dugdill, L., Graham, R., & McNair, F. (2005). Exercise referral: the public health panacea for physical activity promotion? A critical perspective of exercise referral schemes; their development and evaluation. *Ergonomics*, *48*(11–14), 1390–1410. doi: 10.1080/00140130500101544

Dumith, S., Hallal, P., Reis, R., & Kohl III, H. (2011). Worldwide prevalence of physical inactivity and its association with human development index in 76 countries. *Preventive Medicine*, *53*(1–2), 24–28. doi:

10.1016/j.ypmed.2011.02.017

Dunlop, D., Song, J., Semanik, P., Chang, R., Sharma, L., Bathon, J., … Hootman, J. (2011). Objective physical activity measurement in the osteoarthritis initiative: are guidelines being met? *Arthritis and Rheumatism*, *63*(11), 3372–3382. doi: 10.1002/art.30562

Durante, R., & Ainsworth, B. (1996). The recall of physical activity: using a cognitive model of the question-answering process. *Medicine & Science in Sports & Exercise*, *28*(10), 1282-1291.

Dyer, M., Goldsmith, K., Sharples, L., & Buxton, M. (2010). A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health and Quality of Life Outcomes*, *8*(1), 13–13. doi: 10.1186/1477-7525-8-13

Edmunds, J., Ntoumanis, N., & Duda, J. (2007). Adherence and well-being in overweight and obese patients referred to an exercise on prescription scheme: a self-determination theory perspective. *Advances in Self-Determination Theory Research in Sport and Exercise*, *8*(5), 722–740. doi: 10.1016/j.psychsport.2006.07.006

Edwards, R., & Holland, J. (2013). *What is qualitative interviewing?* London: Bloomsbury Academic.

Edwards, R., Linck, P., Hounsome, N., Raisanen, L., Williams, N., Moore, L., & Murphy, S. (2013). Cost-effectiveness of a national exercise referral programme for primary care patients in Wales: results of a randomised controlled trial. *BMC Public Health*, *13*(1), 1021. doi: 10.1186/1471-2458-13-1021

Edwardson, C., Winkler, E., Bodicoat, D., Yates, T., Davies, M., Dunstan, D., & Healy, G. (2017). Considerations when using the activPAL monitor in field-based research with adult populations. *Journal of Sport and Health Science*, *6*(2), 162–178. doi: 10.1016/j.jshs.2016.02.002

Egger, M., Dickersin, K., & Smith, G. (2008). Problems and limitations in conducting systematic reviews. In M. Egger, G. Davey Smith & D. Altman (Eds.), *Systematic reviews in health care* (pp. 43–68). doi: 10.1002/9780470693926.ch3

Ekelund, U., Ward, H., Norat, T., Luan, J., May, A., Weiderpass, E., … Riboli, E. (2015). Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European prospective investigation into cancer and nutrition study (EPIC). *The American Journal of Clinical Nutrition*, *101*(3), 613-621. doi: 10.3945/ajcn.114.100065

Elkjaer, B., & Simpson, B. (2011). Pragmatism: a lived and living philosophy. What can it offer to contemporary organization theory? In H. Tsoukas & R. Chia (Eds.), *Philosophy and organization theory* (pp. 55–84). Bingley: Emerald Group Publishing Limited.

EuroQol Group. (1990). EuroQol–a new facility for the measurement of health-related quality of life. *Health Policy, 16*(3), 199–208.

EuroQol Group. (n.d.). EQ-5D index calculator. Retrieved from https://eurqol.org

Evenson, K., Butler, E., & Rosamond, W. (2014). United States prevalence of physical activity and sedentary behavior among adults with cardiovascular disease. *Journal of Cardiopulmonary Rehabilitation and Prevention*, *34*(6), 406–419. doi: 10.1097/HCR.0000000000000064

Ewbank, L., Omojomolo, D., Sullivan, K., & McKenna, H. (2018). *The rising cost of medicines to the NHS. What's the story?* Retrieved from https://www.kingsfund.org.uk/sites/default/files/2018-04/Rising-cost-of-medicines.pdf

Eynon, M., O'Donnell, C., & Williams, L. (2016). Gaining qualitative insight into the subjective experiences of adherers to an exercise referral scheme: a thematic analysis. *Journal of Health Psychology*, *23*(11), 1476–1487. doi: 10.1177/1359105316656233

Eynon, M., O'Donnell, C., & Williams, L. (2017). Assessing the impact of autonomous motivation and psychological need satisfaction in explaining adherence to an exercise referral scheme. *Psychology, Health and Medicine*, *22*(9), 1056-1062. doi: 10.1080/13548506.2016.1274041

Family & Childcare Trust. (2018). *Childcare survey 2018*. Retrieved from https://www.familyandchildcaretrust.org/childcare-survey-2018

Farquhar, M., Ewing, G., & Booth, S. (2011). Using mixed methods to develop and evaluate complex interventions in palliative care research. *Palliative Medicine*, *25*(8), 748–757. doi: 10.1177/0269216311417919

Farrell, L., Hollingsworth, B., Propper, C., & Shields, M. (2014). The socioeconomic gradient in physical inactivity: evidence from one million adults in England. *Social Science & Medicine*, *123*, 55–63. doi: 10.1016/j.socscimed.2014.10.039

Feng, Y., Devlin, N., & Herdman, M. (2015). Assessing the health of the general population in England: how do the three- and five-level versions of EQ-5D compare? *Health and Quality of Life Outcomes*, *13*(1), 171–171. doi: 10.1186/s12955-015-0356-8

Fenton, G., Hill, K., Stocker, R., & House, A. (2015). Older adults at risk of a cardiovascular event: a preliminary investigation of their experiences of an active lifestyle scheme in England. *Ageing and Society*, *35*(10), 2141–2155. doi: 10.1017/S0144686X1400083X

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. doi: 10.1177/001872675400700202

Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th ed.). London: SAGE Publications.

Finger, J., Gisle, L., Mimilidis, H., Santos-Hoevener, C., Kruusmaa, E., Matsi, A., … Lange, C. (2015). How well do physical activity questions perform? A European cognitive testing study. *Archives of Public Health*, *73*, 57. doi: 10.1186/s13690-015-0109-5

Fisman, D. (2012). Seasonality of viral infections: mechanisms and unknowns. *Clinical Microbiology and Infection*, *18*(10), 946–954. doi: 10.1111/j.1469-0691.2012.03968.x

Fitzsimons, C., Kirk, A., Murphy, M., & Mutrie, N. (2012). Agreement between the IPAQ-long weekday sitting item and the activPAL activity monitor in Scottish adults. *Journal of Science and Medicine in Sport*, *15*(S1), 295–296. doi: 10.1016/j.jsams.2012.11.719

Fogelman, N., & Canli, T. (2015). 'Purpose in life' as a psychosocial resource in healthy aging: an examination of cortisol baseline levels and response to the trier social stress test. *npj Aging And Mechanisms Of Disease*, *1*, 15006.

Forouzanfar, M., Alexander, L., Anderson, H., Bachman, V., Biryukov, S., Brauer, M., … Murray, C. (2015). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, *386*(10010), 2287–2323. doi: 10.1016/S0140-6736(15)00128-2

Fox, K. (2000). Self-esteem, self-perceptions and exercise. *International Journal of Sport Psychology*, *31*(2), 228–240.

Fox, K, Biddle, S., Edmunds, L., Bowler, I., & Killoran, A. (1997). Physical activity promotion through primary health care in England. *The British Journal of General Practice*, *47*(419), 367–369.

Fransen, M., & Edmonds, J. (1999). Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. *Rheumatology*, *38*(9), 807–813.

Freedson, P., & Miller, K. (2000). Objective monitoring of physical activity using motion sensors and heart rate. *Research Quarterly for Exercise and Sport*, *71*(S2), 21–29. doi: 10.1080/02701367.2000.11082782

Furber, C. (2010). Framework analysis: a method for analysing qualitative data. *African Journal of Midwifery and Women's Health*, *4*(2), 97–100. doi: 10.12968/ajmw.2010.4.2.47612

Gabriel, Z., & Bowling, A. (2004). Perspectives on quality of life in older age: older people talking. *Ageing & Society*, *24*, 675–691. doi: 10.1017/S0144686X03001582

Gabrys, L., Michallik, L., Thiel, C., Vogt, L., & Banzer, W. (2013). Effects of a structured physical-activity counseling and referral scheme in long-term unemployed individuals: a pilot accelerometer study. *Behavioral Medicine, 39*(2), 44–50. doi: 10.1080/08964289.2013.782852

Gale, N., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, *13*(1), 117–117. doi: 10.1186/1471-2288-13-117

Gallegos-Carrillo, K., García-Peña, C., Salmerón, J., Salgado-de-Snyder, N., & Lobelo, F. (2017). Brief counseling and exercise referral scheme: a pragmatic trial in Mexico. *American Journal of Preventive Medicine*, *52*(2), 249–259. doi: 10.1016/j.amepre.2016.10.021

Garber, C., Blissmer, B., Deschenes, M., Franklin, B., Lamonte, M., Lee, I., … Swain, D. (2011). Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine & Science in Sports & Exercise*, *43*(7), 1334–59. doi: 10.1249/MSS.0b013e318213fefb

Garson, D. (2015). *Missing values analysis and data imputation*. Asheboro: Statistical Associates Publishers.

Gentles, S., Charles, C., Ploeg, J., & McKibbon, K. (2015). Sampling in qualitative research: insights from an overview of the methods literature. *The Qualitative Report*, *20*(11), 1772–1789.

Gentles, S., & Vilches, S. (2017). Calling for a shared understanding of sampling terminology in qualitative research: proposed clarifications derived from critical analysis of a methods overview by McCrae and Purssell. *International Journal of Qualitative Methods*, *16*(1). doi: 10.1177/1609406917725678

Gerst-Emerson, K., & Jayawardhana, J. (2015). Loneliness as a public health issue: the impact of loneliness on health care utilization among older adults. *American Journal of Public Health*, *105*(5), 1013–1019. doi: 10.2105/AJPH.2014.302427

Gidlow, C., Johnston, L., Crone, D., & James, D. (2005). Attendance of exercise referral schemes in the UK: a systematic review. *Health Education Journal*, *64*(2), 168–186. doi: 10.1177/001789690506400208

Gidlow, C., Johnston, L., Crone, D., Morris, C., Smith, A., Foster, C., & James, D. (2007). Socio-demographic patterning of referral, uptake and attendance in physical activity referral schemes. *Journal of Public Health*, *29*(2), 107-113. doi: 10.1093/pubmed/fdm002

Gidlow, C., Johnston, L., Crone, D., & James, D. (2008). State of the art reviews: methods of evaluation: issues and implications for physical activity referral schemes. *American Journal of Lifestyle Medicine*, *2*(1), 46–50. doi: 10.1177/1559827607308733.

Gill, J., & Cooper, A. (2008). Physical activity and prevention of type 2 diabetes mellitus. *Sports Medicine*, *38*(10), 807–824. doi: 10.2165/00007256-200838100-00002

Giné-Garriga, M., Coll-Planas, L., Guerra, M., Domingo, À., Roqué, M., Caserotti, P., … Salvà, A. (2017). The SITLESS project: exercise referral schemes enhanced by self-management strategies to battle sedentary behaviour in older adults: study protocol for a randomised controlled trial. *Trials*, *18*, 221. doi: 10.1186/s13063-017-1956-x

Global Advocacy Council for Physical Activity, & International Society for Physical Activity and Health. (2010). *The Toronto charter for physical activity: a global call to action*. Retrieved from https://ispah-org-draft.squarespace.com/resources

Global Advocacy Council for Physical Activity, & International Society for Physical Activity and Health. (2011). *NCD prevention: investments that work for physical activity*. Retrieved from https://ispah-org-draft.squarespace.com/resources

Goff, D., Lloyd-Jones, D., Bennett, G., Coady, S., D'Agostino, R., Gibbons, R., … Wilson, P. (2013). 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *Circulation*, *129*(25 S2), 49-73. doi: 10.1161/01.cir.0000437741.48606.98

Graham, R., Dugdill, L., & Cable, N. (2005). Health professionals' perspectives in exercise referral: implications for the referral process. *Ergonomics*, *48*(11–14), 1411–22. doi: 10.1080/00140130500101064

Grant, P., Dall, P., Mitchell, S., & Granat, M. (2008). Activity-monitor accuracy in measuring step number and cadence in community-dwelling older adults. *Journal of Aging and Physical Activity*, *16*(2), 201–214. https://doi.org/10.1123/japa.16.2.201

Grant, P., Ryan, C., Tigbe, W., & Granat, M. (2006). The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British Journal of Sports Medicine*, *40*(12), 992–997. doi: 10.1136/bjsm.2006.030262

Gravetter, F., & Forzano, L. (2011). *Research methods for the behavioral sciences* (4th ed.). Belmont: CENGAGE Learning Custom Publishing.

Gray, D. (2004). *Doing research in the real world.* London: SAGE Publications.

Greene, J., Caracelli, V., & Graham, W. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274. doi: 10.3102/01623737011003255

Greene, J., & Caracelli, V. (1997). *Advances in mixed-method evaluation: the challenges and benefits of integrating diverse paradigms*. San Francisco: Jossey-Bass.

Gremigni, P., & Bianco, D. (2012). *Performance of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) as a screening tool for depression in UK and Italy*. Retrieved from http://www2.warwick.ac.uk/fac/med/research/platform/wemwbs/development/papers/donatella_bianco-thesis.pdf

Gusi, N., Reyes, M., Gonzalez-Guerrero, J., Herrera, E., & Garcia, J. (2008). Cost-utility of a walking programme for moderately depressed, obese, or overweight elderly women in primary care: a randomised controlled trial. *BMC Public Health*, *8*, 231. doi: 10.1186/1471-2458-8-231

Guthold, R., Stevens, G., Riley, L., & Bull, F. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health*, *6*(10), e1077–

e1086. doi: 10.1016/S2214-109X(18)30357-7

Haase, J., & Myers, S. (1988). Reconciling paradigm assumptions of qualitative and quantitative research. *Western Journal of Nursing Research*, *10*(2), 128–137. doi: 10.1177/019394598801000202

Hagger, M., Chatzisarantis, N., & Biddle, S. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: predictive validity and the contribution of additional variables. *Journal of Sport and Exercise Psychology*, *24*(1), 3–32. doi: 10.1123/jsep.24.1.3

Hallal, P., Andersen, L., Bull, F., Guthold, R., Haskell, W., & Ekelund, U. (2011). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, *380*(9838), 247–257. doi: 10.1016/S0140-6736(12)60646-1

Hallal, P., Simoes, E., Reichert, F., Azevedo, M., Ramos, L., Pratt, M., & Brownson, R. (2010). Validity and reliability of the telephone-administered international physical activity questionnaire in Brazil. *Journal of Physical Activity & Health*, *7*(3), 402–409.

Hallgren, M., Vancampfort, D., Giesen, E., Lundin, A., & Stubbs, B. (2017). Exercise as treatment for alcohol use disorders: systematic review and meta-analysis. *British Journal of Sports Medicine*, *51*(14), 1058. doi: 10.1136/bjsports-2016-096814

Hansen, E. (2006). Successful qualitative health research: a practical introduction. Sydney: Allen & Unwin.

Hanson, C., Allin, L., Ellis, J., & Dodd-Reynolds, C. (2013). An evaluation of the efficacy of the exercise on referral scheme in Northumberland, UK: association with physical activity and predictors of engagement. A naturalistic observation study. *BMJ Open*, *3*, e002849. doi: 10.1136/bmjopen-2013-002849

Hanson, S., & Jones, A. (2017). Missed opportunities in the evaluation of public health interventions: a case study of physical activity programmes. *BMC Public Health*, *17*(1), 674. doi: 10.1186/s12889-017-4683-z

Hardcastle, S., & Taylor, A. (2001). Looking for more than weight loss and fitness gain: psychosocial dimensions among older women in a primary-care exercise-referral program. *Journal of Aging and Physical Activity*, *9*(3), 313–328. doi: 10.1123/japa.9.3.313

Harrington, D., Welk, G., & Donnelly, A. (2011). Validation of MET estimates and step measurement using the ActivPAL physical activity logger. *Journal of Sports Sciences*, *29*(6), 627–633. doi: 10.1080/02640414.2010.549499

Harrison, R., McNair, F., & Dugdill, L. (2005a). Access to exercise referral schemes – a population based analysis. *Journal of Public Health*, *27*(4), 326–30. doi: 10.1093/pubmed/fdi048

Harrison, R., Roberts, C., & Elton, P. (2005b). Does primary care referral to an exercise programme increase physical activity one year later? A randomized controlled trial. *Journal of Public Health*, *27*(1), 25–32. doi: 10.1093/pubmed/fdh197

Hart, T., Swartz, A., Cashin, S., & Strath, S. (2011). How many days of monitoring predict physical activity and sedentary behaviour in older adults? *International Journal of Behavioral Nutrition and Physical Activity*, *8*(1), 62. doi: 10.1186/1479-5868-8-62

Haskell, W. (2012). Physical activity by self-report: a brief history and future issues. *Journal of Physical Activity & Health*, *9*(S1), 5–10.

Hausenblas, H., Carron, A., & Mack, D. (1997). Application of the theories of reasoned action and planned behavior to exercise behavior: a meta-analysis. *Journal of Sport and Exercise Psychology*, *19*(1), 36–51. doi:

10.1123/jsep.19.1.36

Hawkins, J., Oliver, E., Wyatt-Williams, J., Scale, E., & van Woerden, H. (2014). Assessing the feasibility of using uniaxial accelerometers with an online support platform in the delivery of a community-based exercise referral scheme. *Journal of Primary Care & Community Health*, *5*(4), 288–90. doi: 10.1177/2150131914536053

Hays, R., Kim, S., Spritzer, K., Kaplan, R., Tally, S., Feeny, D., … Fryback, D. (2009). Effects of mode and order of administration on generic health-related quality of life scores. *Value in Health*, *12*(6), 1035–1039. doi: 10.1111/j.1524-4733.2009.00566.x

Health and Social Care Information Centre. (2009). *Health survey for England - 2008, physical activity and fitness: summary of key findings*. Retrieved from http://digital.nhs.uk/catalogue/PUB00430/heal-surv-phys-acti-fitn-eng-2008-rep-v1.pdf

Health and Social Care Information Centre. (2013). *Health survey for England - 2012: chapter 2, physical activity in adults*. Retrieved from http://digital.nhs.uk/catalogue/PUB13218/HSE2012-Ch2-Phys-act-adults.pdf

Health and Social Care Information Centre. (2017). *Health survey for England, 2016*. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2016

Heesch, K., van Uffelen, J., Hill, R., & Brown, W. (2010). What do IPAQ questions mean to older adults? Lessons from cognitive interviews. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 35. doi: 10.1186/1479-5868-7-35

Helmerhorst, H., Brage, S., Warren, J., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), 1–55. doi: 10.1186/1479-5868-9-103

Henderson, H., Evans, A., Allen-Collinson, J., & Siriwardena, N. (2018). The ‘wild and woolly’ world of exercise referral schemes: contested interpretations of an exercise as medicine programme. *Qualitative Research in Sport, Exercise and Health*, *10*(4), 505-523. https://doi.org/10.1080/2159676X.2017.1352018

Henry, G. (1990). *Practical sampling*. London: SAGE Publications.

Hertogh, E., Schuit, A., Peeters, P., & Monninkhof, E. (2010). Noncompliance in lifestyle intervention studies: the instrumental variable method provides insight into the bias. *Journal of Clinical Epidemiology*, *63*(8), 900–906. doi: 10.1016/j.jclinepi.2009.10.007

Hills, A., Mokhtar, N., & Byrne, N. (2014). Assessment of physical activity and energy expenditure: an overview of objective measures. *Frontiers in Nutrition*, *1*, 5. doi: 10.3389/fnut.2014.00005

Hoggarth, L., & Comfort, H. (2010). *A practical guide to outcome evaluation*. London: Jessica Kingsley Publishers.

Holt-Lunstad, J., Smith, T., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspectives on Psychological Science*, *10*(2), 227–237. doi: 10.1177/1745691614568352

Holt-Lunstad, J., Smith, T., & Layton, J. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine*, *7*(7), e1000316. doi: 10.1371/journal.pmed.1000316

Huber, M. (2010). *Invitational conference ‘is health a state or an ability? Towards a dynamic concept of health’: report of the meeting December 10-11, 2009*.

Retrieved from https://www.gezondheidsraad.nl/sites/default/files/bijlage A1004_1.pdf

Huber, M., Knottnerus, J., Green, L., Horst, H., Jadad, A., Kromhout, D., … Smid, H. (2011). How should we define health? *BMJ*, *343*, d4163. doi: 10.1136/bmj.d4163

Hupin, D., Roche, F., Gremeaux, V., Chatard, J., Oriol, M., Gaspoz, J., … Edouard, P. (2015). Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged ≥60 years: a systematic review and meta-analysis. *British Journal of Sports Medicine*, *49*(19), 1262-1267. Doi: 10.1136/bjsports-2014-094306

Hutchison, A., Breckon, J., & Johnston, L. (2009). Physical activity behavior change interventions based on the transtheoretical model: a systematic review. *Health Education & Behavior*, *36*(5), 829–845. doi: 10.1177/1090198108318491

Iliffe, S., Tai, S., Gould, M., Thorogood, M., & Hillsdon, M. (1994). Prescribing exercise in general practice. *BMJ*, *309*(6953), 494–495.

International Society for Physical Activity and Health. (2016). The Bangkok declaration on physical activity for global health and sustainable development. Retrieved from https://ispah-org-draft.squarespace.com/resources

IPAQ. (2005). *Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)*. Retrieved from https://docs.google.com/viewer?a=v & pid=sites & srcid=ZGVmYXVsdGRvbWFpbnx0aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWRlZTM

Irving, G., Holden, J., Stevens, R., & McManus, R. (2016). Which cuff should I use? Indirect blood pressure measurement for the diagnosis of hypertension in patients with obesity: a diagnostic accuracy review. *BMJ Open*, *6*(11). doi: 10.1136/bmjopen-2016-012429

Isaacs, A., Critchley, J., Tai, S., Buckingham, K., Westley, D., Harridge, S., … Gottlieb, J. (2007). Exercise evaluation randomised trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technology Assessment*, *11*(10), 1–165. doi: 10.3310/hta11100

Jadad, A., & O'Grady, L. (2008). How should health be defined? *BMJ*, *337*, a2900. doi: 10.1136/bmj.a2900

James, D., Johnston, L., Crone, D., Sidford, A., Gidlow, C., Morris, C., & Foster, C. (2008). Factors associated with physical activity referral uptake and participation. *Journal of Sports Sciences*, *26*(2), 217–24. doi: 10.1080/02640410701468863

James, D., Mills, H., Crone, D., Johnston, L., Morris, C., & Gidlow, C. (2009). Factors associated with physical activity referral completion and health outcomes. *Journal of Sports Sciences*, *27*(10), 1007–1017. doi: 10.1080/02640410903214248

James, S., Abate, D., Abate, K., Abay, S., Abbafati, C., Abbasi, N., … Murray, C. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, *392*(10159), 1789–1858. doi: 10.1016/S0140-6736(18)32279-7

Janssen, M., Lubetkin, E., Sekhobo, J., & Pickard, A. (2011). The use of the EQ-5D preference-based health status measure in adults with type 2 diabetes mellitus. *Diabetic Medicine*, *28*(4), 395–413. doi: 10.1111/j.1464-5491.2010.03136.x

Janssen, M., Pickard, A., Golicki, D., Gudex, C., Niewada, M., Scalone, L., …

Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research*, *22*(7), 1717–1727. doi: 10.1007/s11136-012-0322-4

Johnson, J., & Coons, S. (1998). Comparison of the EQ-5D and SF-12 in an adult US sample. *Quality of Life Research*, *7*(2), 155–166.

Johnson, R., & Onwuegbuzie, A. (2004). Mixed methods research: a research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. doi: 10.3102/0013189X033007014

Johnson, R., Onwuegbuzie, A., & Turner, L. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112–133. doi: 10.1177/1558689806298224

Johnston, L., Warwick, J., De Ste Croix, M., Crone, D., & Sldford, A. (2005). The nature of all 'inappropriate referrals' made to a countywide physical activity referral scheme: Implications for practice. *Health Education Journal*, *64*(1), 58–69. doi: 10.1177/001789690506400107

Jones, F., Harris, P., Waller, H., & Coggins, A. (2005). Adherence to an exercise prescription scheme: the role of expectations, self-efficacy, stage of change and psychological well-being. *British Journal of Health Psychology*, *10*(3), 359–378. doi: 10.1348/135910704X24798

Jootun, D., McGhee, G., & Marland, G. (2009). Reflexivity: promoting rigour in qualitative research. *Nursing Standard*, *23*(23), 42–46.

Kaczynski, A., Manske, S., Mannell, R., & Grewal, K. (2008). Smoking and physical activity: a systematic review. *American Journal of Health Behavior*, *32*(1), 93–110. doi: 10.5993/AJHB.32.1.9

Kaptein, S., & Badley, E. (2012). Sex differences, age, arthritis, and chronic disease: influence on physical activity behaviors. *Journal of Physical Activity and Health*, *9*(4), 540–548. doi: 10.1123/jpah.9.4.540

Kelly, M., & Barker, M. (2016). Why is changing health-related behaviour so difficult? *Public Health*, *136*, 109–116. doi: 10.1016/j.puhe.2016.03.030

Kelly, M., Rae, G., Walker, D., Partington, S., Dodd-Reynolds, C., & Caplan, N. (2017). Retrospective cohort study of the South Tyneside exercise referral scheme 2009–14: predictors of dropout and barriers to adherence. *Journal of Public Health*, *39*(4), 257–264. doi: 10.1093/pubmed/fdw122

Kelly, P., Fitzsimons, C., & Baker, G. (2016). Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *International Journal of Behavioral Nutrition and Physical Activity*, *13*(1), 32. doi: 10.1186/s12966-016-0351-4

Keys, A., Fidanza, F., Karvonen, M., Kimura, N., & Taylor, H. (1972). Indices of relative weight and obesity. *Journal of Chronic Diseases*, *25*(6), 329–343.

Kim, E., Sun, J., Park, N., Kubzansky, L., & Peterson, C. (2013a). Purpose in life and reduced risk of myocardial infarction among older U.S. adults with coronary heart disease: a two-year follow-up. *Journal of Behavioral Medicine*, *36*(2), 124–133. doi: 10.1007/s10865-012-9406-4

Kim, E., Sun, J., Park, N., & Peterson, C. (2013b). Purpose in life and reduced incidence of stroke in older adults: 'the health and retirement study'. *Journal of Psychosomatic Research*, *74*(5), 427–432. doi: 10.1016/j.jpsychores.2013.01.013

Kim, Y., Barry, V., & Kang, M. (2015). Validation of the ActiGraph GT3X and activPAL accelerometers for the assessment of sedentary behavior. *Measurement in Physical Education and Exercise Science*, *19*(3), 125–137. doi: 10.1080/1091367X.2015.1054390

Kim, Y., White, T., Wijndaele, K., Westgate, K., Sharp, S., Helge, J., … Brage, S.

(2018). The combination of cardiorespiratory fitness and muscle strength, and mortality risk. *European Journal of Epidemiology, 33*(10), 953-964. doi: 10.1007/s10654-018-0384-x

King, G., Keohane, R., & Verba, S. (1994). *Designing social inquiry: scientific inference in qualitative research*. Princeton: Princeton University Press.

Knight, T., Leech, F., Jones, A., Walker, L., Wickramasinghe, R., Angris, S., & Rolfe, P. (2001). Sphygmomanometers in use in general practice: an overlooked aspect of quality in patient care. *Journal of Human Hypertension*, *15*(10), 681–684. doi.org/10.1038/sj.jhh.1001251

Korkiakangas, E., Alahuhta, M., & Laitinen, J. (2009). Barriers to regular exercise among adults at high risk or diagnosed with type 2 diabetes: a systematic review. *Health Promotion International*, *24*(4), 416–427. doi: 10.1093/heapro/dap031

Kowalski, K., Rhodes, R., Naylor, P., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), 148. doi: 10.1186/1479-5868-9-148

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. (2011). Validation of wearable monitors for assessing sedentary behavior. *Medicine & Science in Sports & Exercise*, *43*(8), 1561-1567. doi: 10. 1249/MSS.0b01e31820ce174

Kruger, J., Ham, S., & Sanker, S. (2008). Physical inactivity during leisure time among older adults–behavioral risk factor surveillance system, 2005. *Journal of Aging and Physical Activity*, *16*(3), 280–91.

Kushman, J. (1992). The organizational dynamics of teacher workplace commitment: a study of urban elementary and middle schools. *Educational Administration Quarterly*, *28*(1), 5–42. doi: 10.1177/0013161X92028001002

Kuzel, A., & Like, R. (1991). Standards of trustworthiness for qualitative studies in primary care. In P. Norton, M. Steward, F. Tudiver, M. Bass & E. Dunn (Eds.), *Primary Care Research* (pp. 138-158). Thousand Oaks: SAGE Publications.

Kyu, H., Abate, D., Abate, K., Abay, S., Abbafati, C., Abbasi, N., … Murray, C. (2018). Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, *392*(10159), 1859–1922. doi: 10.1016/S0140-6736(18)32335-3

Lacey, A., & Luff, D. (2009). *Qualitative research analysis. the NIHR RDS for the East Midlands/Yorkshire & the Humber, 2007*. Retrieved from https://www.rds-yh.nihr.ac.uk/wp-content/uploads/2013/05/9_Qualitative_Data_Analysis_Revision_2009.pdf

Lazarus, R. (1976). *Patterns of adjustment*. New York: McGraw Hill.

Lee, A., Griffin, S., & Simmons, R. (2009). An evaluation of the effectiveness of 'active for life'': an exercise referral scheme in West Suffolk'. *Public Health*, *123*(10), 670–672. doi: 10.1016/j.puhe.2009.09.005

Lee, I., & Shiroma, E. (2014). Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. *British Journal of Sports Medicine*, *48*(3), 197-201.

Lee, I., Shiroma, E., Lobelo, F., Puska, P., Blair, S., & Katzmarzyk, P. (2012). Impact of physical inactivity on the world's major non-communicable diseases. *Lancet*, *380*(9838), 219–229. doi: 10.1016/S0140-6736(12)61031-9

Lee, P., Macfarlane, D., Lam, T., & Stewart, S. (2011). Validity of the International Physical Activity Questionnaire short form (IPAQ-SF): a systematic review.

International Journal of Behavioral Nutrition and Physical Activity, *8*, 115. doi: 10.1186/1479-5868-8-115

Lee, S., Kuk, J., Davidson, L., Hudson, R., Kilpatrick, K., Graham, T., & Ross, R. (2005). Exercise without weight loss is an effective strategy for obesity reduction in obese individuals with and without type 2 diabetes. *Journal of Applied Physiology*, *99*(3), 1220–1225. doi: 10.1152/japplphysiol.00053.2005

Levin, S., Jacobs, D., Ainsworth, B., Richardson, M., & Leon, A. (1999). Intra-individual variation and estimates of usual physical activity. *Annals of Epidemiology*, *9*(8), 481–488. doi: 10.1016/S1047-2797(99)00022-8

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: SAGE Publications.

Littlecott, H., Moore, G., Moore, L., & Murphy, S. (2014). Psychosocial mediators of change in physical activity in the Welsh national exercise referral scheme: secondary analysis of a randomised controlled trial. *International Journal of Behavioral Nutrition and Physical Activity*, *11*, 109. doi: 10.1186/s12966-014-0109-9

Liu, K., O'Brien, E., Guralnik, J., Criqui, M., Martin, G., Greenland, P., & McGrae McDermott, M. (2000). Measuring physical activity in peripheral arterial disease: a comparison of two physical activity questionnaires with an accelerometer. *Angiology*, *51*(2), 91–100. doi: 10.1177/000331970005100201

Lord, J., & Green, F. (1995). Exercise on prescription: does it work? *Health Education Journal*, *54*(4), 453–464.

Lozano, F., Lobos, J., March, J., Carrasco, E., Barros, M., & González-Porras, J. (2016). Self-administered versus interview-based questionnaires among patients with intermittent claudication: do they give different results? A cross-sectional study. *Sao Paulo Medical Journal*, *134*(1), 63–69. doi: 10.1590/1516-3180.2015.01733009

MacNealy, M. (1999). *Strategies for empirical research in writing*. New York: Addison Wesley Longman.

Macran, S., Weatherly, H., & Kind, P. (2003). Measuring population health. *Medical Care*, *41*(2), 218–231. doi: 10.1097/01.MLR.0000044901.57067.19

Madden, A., & Smith, S. (2016). Body composition and morphological assessment of nutritional status in adults: a review of anthropometric variables. *Journal of Human Nutrition and Dietetics*, *29*(1), 7–25. doi: 10.1111/jhn.12278

Maddison, R., Ni Mhurchu, C., Jiang, Y., Vander Hoorn, S., Rodgers, A., Lawes, C., & Rush, E. (2007). International Physical Activity Questionnaire (IPAQ) and New Zealand Physical Activity Questionnaire (NZPAQ): a doubly labelled water validation. *International Journal of Behavioral Nutrition and Physical Activity*, *4*, 62. doi: 10.1186/1479-5868-4-62

Maheswaran, H., Weich, S., Powell, J., & Stewart-Brown, S. (2012). Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): group and individual level analysis. *Health and Quality of Life Outcomes*, *10*, 156. doi: 10.1186/1477-7525-10-156

Mancia, G. (2012). Short- and long-term blood pressure variability. *Hypertension*, *60*(2), 512-517.

Mancia, G., Bertinieri, G., Grassi, G., Parati, G., Pomidossi, G., Ferrari, A., … Zanchetti, A. (1983a). Effects of blood-pressure measurement by the doctor on patient's blood pressure and heart rate. *Lancet*, *2*(8352), 695–698.

Mancia, G., Fagard, R., Narkiewicz, K., Redon, J., Zanchetti, A., Böhm, M., … Wood, D. (2013). 2013 ESH/ESC guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of

Cardiology (ESC). *European Heart Journal*, *34*(28), 2159–2219. doi: 10.1093/eurheartj/eht151

Mancia, G., Ferrari, A., Gregorini, L., Parati, G., Pomidossi, G., Bertinieri, G., … Zanchetti, A. (1983b). Blood pressure and heart rate variabilities in normotensive and hypertensive human beings. *Circulation Research*, *53*(1), 96-104.

Marchman Andersen, M., Dalton, S., Lynch, J., Johansen, C., & Holtug, N. (2013). Social inequality in health, responsibility and egalitarian justice. *Journal of Public Health*, *35*(1), 4–8. doi: 10.1093/pubmed/fdt012

Markandu, N., Whitcher, F., Arnold, A., & Carney, C. (2000). The mercury sphygmomanometer should be abandoned before it is proscribed. *Journal of Human Hypertension*, *14*(1), 31–36.

Marmot, M., Allen, J., Goldblatt, P., Boyce, T., McNeish, D., Grady, M., & Geddes, I. (2010). *Fair society healthy lives (the Marmot review)*. Retrieved from http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review/fair-society-healthy-lives-full-report

Martin, C., & Woolf-May, K. (1999). The retrospective evaluation of a general practitioner exercise prescription programme. *Journal of Human Nutrition and Dietetics*, *12*(S1), 32–42. doi: 10.1046/j.1365-277X.1999.00005.x

Martín-Borràs, C., Giné-Garriga, M., Puig-Ribera, A., Martín, C., Solà, M., & Cuesta-Vargas, A. I. (2018). A new model of exercise referral scheme in primary care: is the effect on adherence to physical activity sustainable in the long term? A 15-month randomised controlled trial. *BMJ Open*, *8*, e017211. doi: 10.1136/bmjopen-2017-017211

Marttila, J., & Nupponen, R. (2003). Assessing stage of change for physical activity: how congruent are parallel methods? *Health Education Research*, *18*(4), 419–428. doi: 10.1093/her/cyf034

Mason, J. (2002). *Qualitative researching* (2nd ed.). London: SAGE Publications.

Mastellos, N., Gunn, L., Felix, L., Car, J., & Majeed, A. (2014). Transtheoretical model stages of change for dietary and physical exercise modification in weight loss management for overweight and obese adults. *Cochrane Database of Systematic Reviews*, *2*, CD008066. doi: 10.1002/14651858.CD008066.pub3

Matthews, C., Ainsworth, B., Thompson, R., & Bassett, D. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine & Science in Sports & Exercise*, *34*(8), 1376-1381.

Matthews, C., Hagströmer, M., Pober, D., & Bowles, H. (2012). Best practices for using physical activity monitors in population-based research. *Medicine and Science in Sports and Exercise*, *44*(S1), 68–76. doi: 10.1249/MSS.0b013e3182399e5b

Maukonen, M., Männistö, S., & Tolonen, H. (2018). A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review. *Scandinavian Journal of Public Health*, *46*(5), 565–579. doi: 10.1177/1403494818761971

Mays, N., & Pope, C. (2000). Assessing quality in qualitative research. *BMJ*, *320*(7226), 50. https://doi.org/10.1136/bmj.320.7226.50

McAuley, E., Szabo, A., Gothe, N., & Olson, E. (2011). Self-efficacy: implications for physical activity, function, and functional limitations in older adults. *American Journal of Lifestyle Medicine*, *5*(4), 361-369. doi: 10.1177/1559827610392704

McGeechan, G., Phillips, D., Wilson, L., Whittaker, V., O'Neill, G., & Newbury-Birch, D. (2018). Service evaluation of an exercise on referral scheme for adults with existing health conditions in the United Kingdom. *International Journal of*

Behavioral Medicine, 25(3), 304-311. doi: 10.1007/s12529-017-9699-3

McKnight, P., & McKnight, K. (2011). Missing data in secondary data analysis. In K. Trzesniewski, M. Donnellan & R. Lucas (Eds.), *Secondary data analysis: an introduction for psychologists* (pp. 83-101). Washington: American Psychological Association Press.

McLaren, L., & Hawe, P. (2005). Ecological perspectives in health research. *Journal of Epidemiology and Community Health*, *59*(1), 6–14. doi: 10.1136/jech.2003.018044

McLeroy, K., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly*, *15*(4), 351–377.

McManus, S., Bebbington, P., Jenkins, R., & Brugha, T. (2016). *Adult psychiatric morbidity survey: survey of mental health and wellbeing, England, 2014*. Retrieved from https://webarchive.nationalarchives.gov.uk/20180328140249/http://digital.nhs.uk/media/35660/APMS-2014-Full-Report/pdf/Mental_health_and_wellbeing_in_England_full_report

McNair, F., Graham, R., Dugdill, L., Cable, T., Harrison, R., & Eubank, M. (2005). Determining the indicators of success in exercise referral schemes. *Journal of Sports Sciences*, *23*, 220–222.

McPhail, S., Lane, P., Russell, T., Brauer, S., Urry, S., Jasiewicz, J., … Haines, T. (2009). Telephone reliability of the Frenchay Activity Index and EQ-5D amongst older adults. *Health and Quality of Life Outcomes*, *7*(1), 48–48. doi: 10.1186/1477-7525-7-48

Medical Research Council. (2000). *A framework for the development and evaluation of RCTs for complex interventions to improve health*. Retrieved from https://mrc.ukri.org/documents/pdf/rcts-for-complex-interventions-to-improve-health/

Medical Research Council. (2008). *Developing and evaluating complex interventions*. Retrieved from https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/

Medical Research Council. (2015). *Process evaluation of complex interventions UK Medical Research Council (MRC) guidance*. Retrieved from https://mrc.ukri.org/documents/pdf/mrc-phsrn-process-evaluation-guidance-final/

Meek, P., Bergeron, D., Towne, D., Ahn, S., Ory, G., & Smith, L. (2018). Restricted social engagement among adults living with chronic conditions. *International Journal of Environmental Research and Public Health*, *15*(1), e158. doi: 10.3390/ijerph15010158

Melanson, E., Freedson, P., & Blair, S. (1996). Physical activity assessment: a review of methods. *Critical Reviews in Food Science and Nutrition*, *36*(5), 385–396. doi: 10.1080/10408399609527732

Michie, S., West, R., Campbell, R., Brown, J., & Gainforth, H. (2014). *ABC of behaviour change theories*. Great Britain: Silverback Publishing.

Mikkelsen, K., Stojanovska, L., Polenakovic, M., Bosevski, M., & Apostolopoulos, V. (2017). Exercise and mental health. *Maturitas*, *106*, 48–56. doi: 10.1016/j.maturitas.2017.09.003

Milanović, Z., Sporiš, G., & Weston, M. (2015). Effectiveness of high-intensity interval training (hit) and continuous endurance training for vo2max improvements: a systematic review and meta-analysis of controlled trials. *Sports Medicine*, *45*(10), 1469–1481. doi: 10.1007/s40279-015-0365-0

Milat, A., King, L., Bauman, A., & Redman, S. (2011). Scaling up health promotion interventions: an emerging concept in implementation science. *Health*

Promotion Journal of Australia, *22*(3), 238–238. doi: 10.1071/HE11238

Miles, M., & Huberman, A. (1984). *Qualitative data analysis: a sourcebook of new methods*. Beverly Hills: SAGE Publications.

Miller, S. (2006). Mixed methods as methodological innovations: problems and prospects. *Methodological Innovations Online*, *1*(1), 29–33. doi: 10.4256/mio.2006.0005

Miller, W. (1983). Motivational interviewing with problem drinkers. *Behavioural Psychotherapy*, *11*(2), 147–172. doi: 10.1017/S0141347300006583

Mills, H., Crone, D., James, D., & Johnston, L. (2012). Exploring the perceptions of success in an exercise referral scheme: a mixed method investigation. *Evaluation Review*, *36*(6), 407–29. doi: 10.1177/0193841X12474452

Misselbrook, D. (2014). W is for wellbeing and the WHO definition of health. *British Journal of General Practice*, *64*(628), 582–582. doi: 10.3399/bjgp14X682381

Molloy, G., McGee, H., O'Neill, D., & Conroy, R. (2010). Loneliness and emergency and planned hospitalizations in a community sample of older adults. *Journal of the American Geriatrics Society*, *58*(8), 1538–1541. doi: 10.1111/j.1532-5415.2010.02960.x

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., … Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, *350*, h1258. doi: 10.1136/bmj.h1258

Moore, G., Evans, R., Hawkins, J., Littlecott, H., Melendez-Torres, G., Bonell, C., & Murphy, S. (2018). From complex social interventions to interventions in complex social systems: future directions and unresolved questions for intervention development and evaluation. *25*(1), 23-45. *Evaluation*. doi: 10.1177/1356389018803219

Moore, G., Moore, L., & Murphy, S. (2011). Facilitating adherence to physical activity: exercise professionals' experiences of the national exercise referral scheme in Wales: a qualitative study. *BMC Public Health*, *11*, 935. doi: 10.1186/1471-2458-11-935

Moore, G., Moore, L., & Murphy, S. (2012). Integration of motivational interviewing into practice in the national exercise referral scheme in Wales: a mixed methods study. *Behavioural and Cognitive Psychotherapy*, *40*(3), 313–330. doi: 10.1017/S1352465811000531

Moore, G., Raisanen, L., Moore, L., Ud Din, N., & Murphy, S. (2013). Mixed-method process evaluation of the Welsh national exercise referral scheme. *Health Education*, *113*(6), 476–501. doi: 10.1108/HE-08-2012-0046

Moore, S., Lee, I., Weiderpass, E., Campbell, P., Sampson, J., Kitahara, C., … Patel, A. (2016). Leisure-time physical activity and risk of 26 types of cancer in 1.44 million adults. *JAMA Internal Medicine*, *176*(6), 816–825. doi: 10.1001/jamainternmed.2016.1548

Morgan, F., Battersby, A., Weightman, A., Searchfield, L., Ruth, R., Morgan, H., … Ellis, S. (2016). Adherence to exercise referral schemes by participants – what do providers and commissioners need to know? A systematic review of barriers and facilitators. *BMC Public Health*, *16*(1), 1–11. doi: 10.1186/s12889-016-2882-7

Morgan, M., Madden, A., Jennings, G., Elia, M., & Fuller, N. (2006). Two-component models are of limited value for the assessment of body composition in patients with cirrhosis. *The American Journal of Clinical Nutrition*, *84*(5), 1151–1162.

Morgan, O. (2005). Approaches to increase physical activity: reviewing the evidence for exercise-referral schemes. *Public Health*, *119*(5), 361–70. doi: 10.1016/j.puhe.2004.06.008

Morris, J. (1994). Exercise in the prevention of coronary heart disease: today's best buy in public health. *Medicine & Science in Sports & Exercise*, *26*(7), 807–814.

Morse, J. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, *40*(2), 120-123.

Morse, J. (1994). Designing qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research*. Thousand Oaks: SAGE Publications.

Morton, K., Biddle, S., & Beauchamp, M. (2008). Changes in self-determination during an exercise referral scheme. *Public Health*, *122*(11), 1257–1260. doi: 10.1016/j.puhe.2007.11.006

Motl, R., McAuley, E., & DiStefano, C. (2005). Is social desirability associated with self-reported physical activity? *Preventive Medicine*, *40*(6), 735–739. doi: 10.1016/j.ypmed.2004.09.016

Murphy, S., Edwards, R., Williams, N., Raisanen, L., Moore, G., Linck, P., … Moore, L. (2012). An evaluation of the effectiveness and cost effectiveness of the national exercise referral scheme in Wales, UK: a randomised controlled trial of a public health policy initiative. *Journal of Epidemiology and Community Health*, *66*(8), 745–753. doi: 10.1136/jech-2011-200689

Murray, C., Barber, R., Foreman, K., Ozgoren, A., Abd-Allah, F., Abera, S., … Vos, T. (2015). Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition. *The Lancet*, *386*(10009), 2145–2191. doi: 10.1016/S0140-6736(15)61340-X

Mushtaq, R., Shoib, S., Shah, T., & Mushtaq, S. (2014). Relationship between loneliness, psychiatric disorders and physical health? A review on the psychological aspects of loneliness. *Journal of Clinical and Diagnostic Research, 8*(9), WE01-04. doi: 10.7860/JCDR/2014/10077.4828

Myers, C., & Wilks, D. (1999). Comparison of Euroqol EQ-5D and SF-36 in patients with chronic fatigue syndrome. *Quality of Life Research*, *8*(1–2), 9–16.

National Health Service. (2016). *The five year forward view for mental health*. Retrieved from https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf

National Health Service. (2018). *Making the case for the personalised approach*. Retrieved from https://www.england.nhs.uk/blog/making-the-case-for-the-personalised-approach/

National Institute for Clinical Excellence. (2006). *Four commonly used methods to increase physical activity*. Retrieved from https://www.nice.org.uk/guidance/ph2

National Institute for Clinical Excellence. (2008). *Physical activity and the environment (PH8)*. Retrieved from https://www.nice.org.uk/guidance/ph8/resources/physical-activity-and-the-environment-pdf-55460874949

National Institute for Clinical Excellence. (2013). *Physical activity: brief advice for adults in primary care (PH44)*. Retrieved from https://www.nice.org.uk/guidance/ph44/resources/physical-activity-brief-advice-for-adults-in-primary-care-pdf-1996357939909

National Institute for Clinical Excellence. (2014a). *Physical activity: exercise referral schemes (PH54)*. Retrieved from https://www.nice.org.uk/guidance/ph54

National Institute for Clinical Excellence. (2014b). *Behaviour change: individual approaches (PH49)*. Retrieved from https://www.nice.org.uk/guidance/ph49

National Institute for Clinical Excellence. (2018). *NICE charter 2017*. Retrieved from https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE_Charter.pdf

Naylor, C., Parsonage, M., McDaid, D., Knapp, M., Fossey, M., & Galea, A. (2012).

Long-term conditions and mental health: the cost of co-morbidities. Retrieved from https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/long-term-conditions-mental-health-cost-comorbidities-naylor-feb12.pdf

Neeland, I., & de Lemos, J. (2016). Time to retire the BMI? *Journal of the American College of Cardiology*, *68*(14), 1522-1524.

Netea, R., Lenders, J., Smits, P., & Thien, T. (1999). Arm position is important for blood pressure measurement. *Journal of Human Hypertension*, *13*(2), 105–109.

Netea, R., Lenders, J., Smits, P., & Thien, T. (2003). Influence of body and arm position on blood pressure readings: an overview. *Journal of Hypertension*, *21*(2), 237–241. doi: 10.1097/01.hjh.0000052439.12292.c2

Newton, J., Briggs, A., Murray, C., Dicker, D., Foreman, K., Wang, H., … Davis, A. (2015). Changes in health in England, with analysis by English regions and areas of deprivation, 1990–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, *386*(10010), 2257–2274. doi: 10.1016/S0140-6736(15)00195-6

Ng, E., & de Colombani, P. (2015). Framework for selecting best practices in public health: a systematic literature review. *Journal of Public Health Research*, *4*(3), 577–577. doi: 10.4081/jphr.2015.577

Ng, S., Howard, A., Wang, H., Su, C., & Zhang, B. (2014). The physical activity transition among adults in China: 1991–2011. *Obesity Reviews*, *15*(S1), 27-36. doi: 10.1111/obr.12127

Ng, S., & Popkin, B. (2012). Time use and physical activity: a shift away from movement across the globe. *Obesity Reviews*, *13*(8), 659–680. doi: 10.1111/j.1467-789X.2011.00982.x

Nigg, C., Geller, K., Motl, R., Horwath, C., Wertin, K., & Dishman, R. (2011). A research agenda to examine the efficacy and relevance of the transtheoretical model for physical activity behavior. *Psychology of Sport and Exercise*, *12*(1), 7–12. doi: 10.1016/j.psychsport.2010.04.004

Norcross, J., Krebs, P., & Prochaska, J. (2011). Stages of change. *Journal of Clinical Psychology*, *67*(2), 143–154. doi: 10.1002/jclp.20758

Norris, S., Grothaus, L., Buchner, D., & Pratt, M. (2000). Effectiveness of physician-based assessment and counseling for exercise in a staff model HMO. *Preventive Medicine*, *30*(6), 513–523. doi: 10.1006/pmed.2000.0673

O’Brien, E. (1996). Review: a century of confusion; which bladder for accurate blood pressure measurement? *Journal of Human Hypertension*, *10*(9), 565–572.

O’Brien, E, Asmar, R., Beilin, L., Imai, Y., Mallion, J., Mancia, G., … Verdecchia, P. (2003). European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. *Journal of Hypertension*, *21*(5), 821-848.

O’Donovan, G., Blazevich, A., Boreham, C., Cooper, A., Crank, H., Ekelund, U., … Stamatakis, E. (2010). The ABC of physical activity for health: a consensus statement from the British Association of Sport and Exercise Sciences. *Journal of Sports Sciences*, *28*(6), 573–591. doi: 10.1080/02640411003671212

Office for National Statistics. (2011). *Census: aggregate data (England and Wales): UK data service census support*. Retrieved from http://infuse.ukdataservice.ac.uk

Office for National Statistics. (2015). *Life expectancy at birth and at age 65 for the UK and local areas in England and Wales*. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarri ages/lifeexpectancies/datasets/lifeexpectancyatbirthandatage65bylocalareasintheunitedkingdomtable1ukandlocalareasinenglandandwales

Office for National Statistics. (2016a). *Labour market profile - Tameside*. Retrieved from https://www.nomisweb.co.uk/reports/lmp/la/1946157088/report.pdf?

Office for National Statistics. (2016b). *Subnational population projections for England: 2014-based*. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/subnationalpopulationprojectionsforengland/2014basedprojections

Office for National Statistics. (2017a). *Health state life expectancies, UK: 2014 to 2016*. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/healthstatelifeexpectanciesuk/2014to2016

Office for National Statistics. (2017b). *Population estimates - local authority based by single year of age*. Retrieved from https://www.nomisweb.co.uk/datasets/pestsyoala

Office for National Statistics. (2018a). *Changing trends in mortality: a cross-UK comparison, 1981 to 2016.* Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/articles/changingtrendsinmortality/acrossukcomparison1981to2016

Office for National Statistics. (2018b). *Sickness absence in the UK labour market*. Retrieved from https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/sicknessabsenceinthelabourmarket

Okorodudu, D., Jumean, M., Montori, V., Romero-Corral, A., Somers, V., Erwin, P., & Lopez-Jimenez, F. (2010). Diagnostic performance of body mass index to identify obesity as defined by body adiposity: a systematic review and meta-analysis. *International Journal of Obesity*, *34*(5), 791–799.

Oliver, E., Hanson, C., Lindsey, I., & Dodd-Reynolds, C. (2016). Exercise on referral: evidence and complexity at the nexus of public health and sport policy. *International Journal of Sport Policy and Politics*, *8*(4), 731-736. doi: 10.1080/19406940.2016.1182048

Olshansky, S., Passaro, D., Hershow, R., Layden, J., Carnes, B., Brody, J., … Ludwig, D. (2005). A potential decline in life expectancy in the United States in the 21st century. *New England Journal of Medicine*, *352*(11), 1138–1145. doi: 10.1056/NEJMsr043743

Ong, A., Uchino, B., & Wethington, E. (2016). Loneliness and health in older adults: a mini-review and synthesis. *Gerontology*, *62*(4), 443–449. doi: 10.1159/000441651

Organisation for Economic Co-operation & Development. (2018). *Health at a glance: Europe 2018*. Retrieved from http://www.oecd.org/health/health-at-a-glance-europe-23056088.htm

Orton, L., Lloyd-Williams, F., Taylor-Robinson, D., O'Flaherty, M., & Capewell, S. (2011). The use of research evidence in public health decision making processes: systematic review. *PloS One*, *6*(7), e21704. doi: 10.1371/journal.pone.0021704

Palatini, P., & Frick, G. (2012). Techniques for self-measurement of blood pressure: limitations and needs for future research. *The Journal of Clinical Hypertension*, *14*(3), 139–143. doi: 10.1111/j.1751-7176.2011.00586.x

Palatini, P., & Parati, G. (2011). Blood pressure measurement in very obese patients: a challenging problem. *Journal of Hypertension*, *29*(3), 425–429. doi:

10.1097/HJH.0b013e3283435b65

Pan, C., Sun, H., Wang, X., Ma, Q., Xu, Y., Luo, N., & Wang, P. (2015). The EQ-5D-5L index score is more discriminative than the EQ-5D-3L index score in diabetes patients. *Quality of Life Research*, *24*(7), 1767–1774. doi: 10.1007/s11136-014-0902-6

Parati, G., Castiglioni, P., Di Rienzo, M., Omboni, S., Pedotti, A., & Mancia, G. (1990). Sequential spectral analysis of 24-hour blood pressure and pulse interval in humans. *Hypertension*, *16*(4), 414-421.

Parati, G., Stergiou, G., Asmar, R., Bilo, G., de Leeuw, P., Imai, Y., … Mancia, G. (2008). European Society of Hypertension guidelines for blood pressure monitoring at home: a summary report of the second international consensus conference on home blood pressure monitoring. *Journal of Hypertension*, *26*(8), 1505-1526. doi: 10.1097/HJH.0b013e328308da66

Pardo, A., Violán, M., Cabezas, C., García, J., Miñarro, C., Rubinat, M., … Romàn-Viñas, B. (2014). Effectiveness of a supervised physical activity programme on physical activity adherence in patients with cardiovascular risk factors. *Apunts Medicina de l'Esport*, *49*(182), 37–44. doi: 10.1016/j.apunts.2014.02.001

Parkin, D., Devlin, N., & Feng, Y. (2014). *What determines the shape of an EQ-5D index distribution?* Retrieved from https://www.ohe.org/publications/what-determines-shape-eq-5d-index-distribution

Parkin, D., Rice, N., & Devlin, N. (2010). Statistical analysis of EQ-5D profiles: does the use of value sets bias inference? *Medical Decision Making*, *30*(5), 556–565. doi: 10.1177/0272989X09357473

Parkinson, S., Eatough, V., Holmes, J., Stapley, E., & Midgley, N. (2016). Framework analysis: a worked example of a study exploring young people's experiences of depression. *Qualitative Research in Psychology*, *13*(2), 109–129. doi: 10.1080/14780887.2015.1119228

Parretti, H., Bartington, S., Badcock, T., Hughes, L., Duda, J., & Jolly, K. (2017). Impact of primary care exercise referral schemes on the health of patients with obesity. *Pragmatic and Observational Research*, *8*, 189–201. doi: 10.2147/POR.S118648

Pate, R., & Dowda, M. (2019). Raising an active and healthy generation: a comprehensive public health initiative. *Exercise and Sport Sciences Reviews*, *47*(1), 3-14. doi: 10.1249/JES.0000000000000171

Patton, M. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks: SAGE Publications.

Patton, M. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks: SAGE Publications.

Patton, M. (2015). *Qualitative research & evaluation methods: integrating theory and practice* (4th ed.). Thousand Oaks: SAGE Publications.

Pavey, T., Anokye, N., Taylor, A., Trueman, P., Moxham, T., Fox, K., … Taylor, R. (2011). The clinical effectiveness and cost-effectiveness of exercise referral schemes: a systematic review and economic evaluation. *Health Technology Assessment*, *15*(44), 1–254. doi: 10.3310/hta15440

Pavey, T., Taylor, A., Hillsdon, M., Fox, K., Campbell, J., Foster, C., … Taylor, R. (2012). Levels and predictors of exercise referral scheme uptake and adherence: a systematic review. *Journal of Epidemiology and Community Health*, *66*(8), 737–44. doi: 10.1136/jech-2011-200354

Pedersen, H., Helmer-Nielsen, M., Dieperink, K., & Østergaard, B. (2016a). Exercise on prescription: a cross-sectional study with self-reported outcome. *Journal of Physical Activity and Health*, *13*(4), 377–384. doi: 10.1123/jpah.2015-0015

Pedersen, E., Danquah, I., Petersen, C., & Tolstrup, J. (2016b). Intra-individual variability in day-to-day and month-to-month measurements of physical activity and sedentary behaviour at work and in leisure-time among Danish adults. *BMC Public Health*, *16*, 1222. doi: 10.1186/s12889-016-3890-3

Pedišić, Ž., & Bauman, A. (2015). Accelerometer-based measures in physical activity surveillance: current practices and issues. *British Journal of Sports Medicine*, *49*(4), 219–223. doi: 10.1136/bjsports-2013-093407

Pettee-Gabriel, K., Morrow, J., & Woolsey, A. (2012). Framework for physical activity as a complex and multidimensional behavior. *Journal of Physical Activity and Health*, *9*(S1), 11–18. doi: 10.1123/jpah.9.s1.s11

Pettman, T., Armstrong, R., Doyle, J., Burford, B., Anderson, L., Hillgrove, T., … Waters, E. (2012). Strengthening evaluation to capture the breadth of public health practice: ideal vs. real. *Journal of Public Health*, *34*(1), 151–155. doi: 10.1093/pubmed/fds014

Pickard, A., Wilke, C., Jung, E., Patel, S., Stavem, K., & Lee, T. (2008). Use of a preference-based measure of health (EQ-5D) in COPD and asthma. *Respiratory Medicine*, *102*(4), 519–536. doi: 10.1016/j.rmed.2007.11.016

Pickering, T., Gerin, W., Schwartz, J., Spruill, T., & Davidson, K. (2008). Franz Volhard lecture: should doctors still measure blood pressure? The missing patients with masked hypertension. *Journal of Hypertension*, *26*(12), 2259–2267. doi: 10.1097/HJH.0b013e32831313c4

Pickering, T., Hall, J., Appel, L., Falkner, B., Graves, J., Hill, M., … Roccella, E. (2005). Recommendations for blood pressure measurement in humans and experimental animals. *Hypertension*, *45*(1), 142-161.

Piepoli, M., Hoes, A., Agewall, S., Albus, C., Brotons, C., Catapano, A., … Gale, C. (2016). 2016 European guidelines on cardiovascular disease prevention in clinical practice. The sixth joint task force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice. *European Heart Journal*, *37*(29), 2315–2315. doi: 10.1093/eurheartj/ehw106

Powell, T. (2001). Competitive advantage: logical and philosophical considerations. *Strategic Management Journal*, *22*(9), 875–888. doi: 10.1002/smj.173

Pratt, M., Perez, L., Goenka, S., Brownson, R., Bauman, A., Sarmiento, O., & Hallal, P. (2015). Can population levels of physical activity be increased? Global evidence and experience. *Progress in Cardiovascular Diseases*, *57*(4), 356–367. doi: 10.1016/j.pcad.2014.09.002

Prince, S., Saunders, T., Gresty, K., & Reid, R. (2014). A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. *Obesity Reviews*, *15*(11), 905–919. doi: 10.1111/obr.12215

Prochaska, J., & DiClemente, C. (1982). Transtheoretical therapy: toward a more integrative model of change. *Psychotherapy: Theory, Research & Practice*, *19*(3), 276–288.

Prochaska, J., & DiClemente, C. (1983). Stages and processes of self-change of smoking: toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*(3), 390–395.

Public Health England. (2014a). *Everybody active, every day: an evidence-based approach to physical activity*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/374914/Framework_13.pdf

Public Health England. (2014b). *Identifying what works for local physical inactivity interventions*. Retrieved from http://researchinstitute.ukactive.com/downloads/managed/Whatworks.pdf

Public Health England. (2015a). *Workplace health: applying all our health*. Retrieved from https://www.gov.uk/government/publications/workplace-health-applying-all-our-health/workplace-health-applying-all-our-health

Public Health England. (2015b). *Tameside local authority health profile 2015*. Retrieved from www.apho.org.uk/resource/view.aspx?RID=171635

Public Health England. (2016). *Active travel: a briefing for local authorities*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/523460/Working_Together_to_Promote_Active_Travel_A_briefing_for_local_authorities.pdf

Public Health England. (2017). *Everybody active, every day: two years on - an update on the national physical activity framework*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/601298/Everybody_active_every_day_2_years_on.pdf

Public Health England. (2018a). *Health profile for England: 2018. Chapter 5: inequalities in health.* Retrieved from https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-5-inequalities-in-health

Public Health England. (2018b). *Tameside local authority health profile 2018*. Retrieved from https://fingertips.phe.org.uk/profile/health-profiles

Public Health England. (2018c). *Moving at scale – promising practice and practical guidance on evaluation of physical activity programmes in the UK*. Retrieved from http://research.ukactive.com/001-moving-scale/Moving-at-scale---Pr.pdf

Puttick, R., & Ludlow, J. (2013). *Standards of evidence: an approach that balances the need for evidence with innovation*. Retrieved from https://media.nesta.org.uk/documents/standards_of_evidence_IMNHDLQ.pdf

Queen, M., Crone, D., & Parker, A. (2015). Long-term engagement with a practice-based exercise referral scheme: patients' perceptions of effectiveness. *European Journal for Person Centered Healthcare, 3*(3), 369-376. doi: 10.5750/ejpch.v3i3.1006

Rahman, R., Thogersen-Ntoumani, C., Thatcher, J., & Doust, J. (2011). Changes in need satisfaction and motivation orientation as predictors of psychological and behavioural outcomes in exercise referral. *Psychology & Health*, *26*(11), 1521–39. doi: 10.1080/08870446.2010.538849

Rebar, A., Stanton, R., Geard, D., Short, C., Duncan, M., & Vandelanotte, C. (2015). A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. *Health Psychology Review*, *9*(3), 366–378. doi: 10.1080/17437199.2015.1022901

Reenen, M., & Oppe, M. (2015). *EQ-5D-3L user guide version 5.1*. Retrieved from https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-3L_UserGuide_2015.pdf

Reichardt, C., & Cook, T. (1979). Beyond qualitative versus quantitative methods. In D. Cook & C. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research*. Beverly Hills: SAGE Publications.

Reichardt, C., & Rallis, S. (1994). Qualitative and quantitative inquiries are not incompatible: a call for a new partnership. *New Directions for Program Evaluation*, *1994*(61), 85–91.

Reilly, J., Penpraze, V., Hislop, J., Davies, G., Grant, S., & Paton, J. (2008). Objective measurement of physical activity and sedentary behaviour: review with new data. *Archives of Disease in Childhood*, *93*(7), 614–619. doi: 10.1136/adc.2007.133272

Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2013). Long-term health benefits of physical activity – a systematic review of longitudinal studies. *BMC Public Health*, *13*, 813. doi: 10.1186/1471-2458-13-813

Resnik, D. (2007). Responsibility for health: personal, social, and environmental. *Journal of Medical Ethics*, *33*(8), 444–445. doi: 10.1136/jme.2006.017574

Resnik, D. (2014). Genetics and personal responsibility for health. *New Genetics and Society*, *33*(2), 113–125. doi: 10.1080/14636778.2014.905195

Ribeiro, S., & Kehayias, J. (2014). Sarcopenia and the analysis of body composition. *Advances in Nutrition, 5*(3), 260–267. doi: 10.3945/an.113.005256

Riddoch, C., Puig-Ribera, A., & Cooper, A. (1998). *Effectiveness of physical activity promotion schemes in primary care: a review.* London: Health Education Authority.

Riemsma, R., Pattenden, J., Bridle, C., Sowden, A., Mather, L., Watt, I., & Walker, A. (2002). A systematic review of the effectiveness of interventions based on a stages-of-change approach to promote individual behaviour change. *Health Technology Assessment*, *6*(24), 1-231. doi: 10.3310/hta6240

Ritchie, J., Lewis, J., McNaughton-Nicholls, C., & Ormston, R. (2013). *Qualitative research practice: a guide for social science students and researchers* (2nd ed.). Thousand Oaks: SAGE Publications.

Ritchie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. Burgess (Eds.), *Analysing qualitative data* (pp. 173–194). London: Routledge.

Ritchie, J., Spencer, L., & O'Connor, W. (2003). Carrying out qualitative analysis. In J. Ritchie & J. Lewis (Eds.), *Qualitative research practice* (pp. 219–262). London: SAGE Publications.

Robinson, O. (2014). Sampling in interview-based qualitative research: a theoretical and practical guide. *Qualitative Research in Psychology*, *11*(1), 25–41. doi: 10.1080/14780887.2013.801543

Rosa, C., Gracia-Marco, L., Barker, A., Freitas, I., & Monteiro, H. (2015). Assessment of physical activity by accelerometer and IPAQ-short version in patients with chronic kidney disease undergoing hemodialysis. *Blood Purification*, *40*(3), 250–255. doi: 10.1159/000437040

Rosenberg, D., Bull, F., Marshall, A., Sallis, J., & Bauman, A. (2008). Assessment of sedentary behavior with the International Physical Activity Questionnaire. *Journal of Physical Activity & Health*, *5*(S1), 30–44.

Rossman, G., & Wilson, B. (1985). Numbers and words: combing quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review*, *9*(5), 627–643. doi: 10.1177/0193841X8500900505

Rouse, P., Ntoumanis, N., Duda, J., Jolly, K., & Williams, G. (2011). In the beginning: role of autonomy support on the motivation, mental health and intentions of participants entering an exercise referral scheme. *Psychology and Health*, *26*(6), 729–749. doi: 10.1080/08870446.2010.492454

Rowley, N., Mann, S., Steele, J., Horton, E., & Jimenez, A. (2018). The effects of exercise referral schemes in the United Kingdom in those with cardiovascular, mental health, and musculoskeletal disorders: a preliminary systematic review. *BMC Public Health*, *18*(1), 949. doi: 10.1186/s12889-018-5868-9

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Russel Bernard, H. (2000). *Social research methods: qualitative and quantitative approaches*. Thousand Oaks: SAGE Publications.

Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D., … White, M. (2017). The need for a complex systems model of evidence for public health. *The Lancet*, *390*(10112), 2602–2604. doi: 10.1016/S0140-6736(17)31267-9

Ryan, R., & Deci, E. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166. doi: 10.1146/annurev.psych.52.1.141

Saint Onge, J., & Krueger, P. (2017). Health lifestyle behaviors among U.S. adults. *SSM - Population Health*, *3*, 89–98. doi: 10.1016/j.ssmph.2016.12.009

Sale, J., Lohfeld, L., & Brazil, K. (2002). Revisiting the quantitative-qualitative debate: implications for mixed-methods research. *Quality and Quantity*, *36*(1), 43–53. doi: 10.1023/A:1014301607592

Salkind, N. (2010). *Encyclopedia of research design*. Thouand Oaks: Sage Publications.

Sallis, J., Haskell, W., Wood, P., Fortmann, S., Rogers, T., Blair, S., & Paffenbarger, R. (1985). Physical activity assessment methodology in the five-city project. *American Journal of Epidemiology*, *121*(1), 91–106. doi: 10.1093/oxfordjournals.aje.a113987

Sallis, J., Owen, N., & Fisher, E. (2008). Ecological models of health behavior. In K. Glanz, B. Rimer & K. Viswanath (Eds.), *Health behavior and health education: theory, research and practice* (4th ed., pp. 465–485). San Francisco: Jossey-Bass.

Sallis, J., & Saelens, B. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise and Sport*, *71*(S2), 1–14. doi: 10.1080/02701367.2000.11082780

Samitz, G., Egger, M., & Zwahlen, M. (2011). Domains of physical activity and all-cause mortality: systematic review and dose–response meta-analysis of cohort studies. *International Journal of Epidemiology*, *40*(5), 1382–1400. doi: 10.1093/ije/dyr112

Sasaki, J., Júnior, J., Meneguci, J., Tribess, S., Marocolo Júnior, M., Stabelini Neto, A., & Virtuoso Júnior, J. (2018). Number of days required for reliably estimating physical activity and sedentary behaviour from accelerometer data in older adults. *Journal of Sports Sciences*, *36*(14), 1572–1577. doi: 10.1080/02640414.2017.1403527

Saunders, M., Lewis, P., Thornhill, A., & Bristow, A. (2015). *Research methods for business students* (7th ed.). Harlow: Pearson Education.

Saunders, T., Campbell, N., Jason, T., Dechman, G., Hernandez, P., Thompson, K., & Blanchard, C. (2016). Objectively measured steps/day in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Journal of Physical Activity & Health*, *13*(11), 1275-1283. doi: 10.1123/jpah.2016-0087

Scarborough, P., Bhatnagar, P., Wickramasinghe, K., Allender, S., Foster, C., & Rayner, M. (2011). The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006–07 NHS costs. *Journal of Public Health*, *33*(4), 527–535. doi: 10.1093/pubmed/fdr033

Scholes, S., Bridges, S., Ng Fat, L., & Mindell, J. (2016). Comparison of the Physical Activity and Sedentary Behaviour Assessment Questionnaire and the Short-Form International Physical Activity Questionnaire: an analysis of health survey for England data. *PLoS ONE*, *11*(3), e0151647. doi: 10.1371/journal.pone.0151647

Schuch, F., Vancampfort, D., Richards, J., Rosenbaum, S., Ward, P., & Stubbs, B. (2016). Exercise as a treatment for depression: a meta-analysis adjusting for publication bias. *Journal of Psychiatric Research*, *77*, 42–51. doi: 10.1016/j.jpsychires.2016.02.023

Scott, S., Breckon, J., Copeland, R., & Hutchison, A. (2015). Determinants and strategies for physical activity maintenance in chronic health conditions: a qualitative study. *Journal of Physical Activity & Health*, *12*(5), 733–40. doi:

10.1123/jpah.2013-0286

Secker, J., Wimbush, E., Watson, J., & Milburn, K. (1995). Qualitative methods in health promotion research: some criteria for quality. *Health Education Journal*, *54*(1), 74–87. doi: 10.1177/001789699505400108

Sedentary Behaviour Research Network. (2012). Letter to the editor: standardized use of the terms “sedentary” and “sedentary behaviours”. *Applied Physiology, Nutrition, and Metabolism*, *37*(3), 540–542. doi: 10.1139/h2012-024

Sedgwick, P. (2013). Convenience sampling. *BMJ*, *347*, f6304. doi: 10.1136/bmj.f6304

Shankar, A., McMunn, A., Banks, J., & Steptoe, A. (2011). Loneliness, social isolation, and behavioral and biological health indicators in older adults. *Health Psychology*, *30*(4), 377–385.

Shenton, A. (2004). Strategies for Ensuring Trustworthiness in Qualitative Research Projects. *Education for Information*, *22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? Implications for health economic evaluation. *BMJ*, *336*(7656), 1281–1283. doi: 10.1136/bmj.39569.510521.AD

Sieber, S. (1973). The integration of fieldwork and survey methods. *American Journal of Sociology*, *78*(6), 1335–1359.

Silfee, V., Haughton, C., Jake-Schoffman, D., Lopez-Cepero, A., May, C., Sreedhara, M., … Lemon, S. (2018). Objective measurement of physical activity outcomes in lifestyle interventions among adults: a systematic review. *Preventive Medicine Reports*, *11*, 74–80. doi: 10.1016/j.pmedr.2018.05.003

Silsbury, Z., Goldsmith, R., & Rushton, A. (2015). Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. *BMJ Open*, *5*, e008430. doi: 10.1136/bmjopen-2015-008430

Silverman, D. (2001). *Interpreting qualitative data: methods for analysing talk, text and interaction* (2nd ed.). London: SAGE Publications.

Sim, J., & Wright, C. (2000). *Research in health care: concepts, designs and methods* (1st ed.). Hampshire: Cengage Learning EMEA.

Smith, A., Ayanian, J., Covinsky, K., Landon, B., McCarthy, E., Wee, C., & Steinman, M. (2011). Conducting high-value secondary dataset analysis: an introductory guide and resources. *Journal of General Internal Medicine, 26*(8), 920-929. doi: 10.1007/s11606-010-1621-5

Smith, E. (2008). *Using secondary data in educational and social research* (1st ed.). Berkshire: Open University Press.

Smith, J., & Firth, J. (2011). Qualitative data analysis: the framework approach. *Nurse Researcher*, *18*(2), 52–62. doi: 10.7748/nr2011.01.18.2.52.c8284

Smith, K., Stewart, S., Riddell, N., & Victor, C. (2018). Investigating the relationship between loneliness and social isolation with inflammation: a systematic review. *Innovation in Aging*, *2*(S1), 839–840. doi: 10.1093/geroni/igy023.3129

Smith, R., & Petticrew, M. (2010). Public health evaluation in the twenty-first century: time to see the wood as well as the trees. *Journal of Public Health*, *32*(1), 2–7. doi: 10.1093/pubmed/fdp122

Social Care Institute for Excellence. (2018). *Prevention for adults with long term conditions*. Retrieved from https://www.scie.org.uk/prevention/people/long-term-conditions

Sofi, F., Capalbo, A., Cesari, F., Abbate, R., & Gensini, G. (2008). Physical activity during leisure time and primary prevention of coronary heart disease: an updated meta-analysis of cohort studies. *European Journal of Cardiovascular*

Prevention & Rehabilitation, *15*(3), 247–257. doi: 10.1097/HJR.0b013e3282f232ac

Sørensen, J., Kragstrup, J., Skovgaard, T., & Puggaard, L. (2008). Exercise on prescription: a randomized study on the effect of counseling vs counseling and supervised exercise. *Scandinavian Journal of Medicine & Science in Sports*, *18*(3), 288–297.

Sørensen, J., Skovgaard, T., & Puggaard, L. (2006). Exercise on prescription in general practice: a systematic review. *Scandinavian Journal of Primary Health Care*, *24*(2), 69–74. doi: 10.1080/02813430600700027

Sørensen, J., Sørensen, J., Skovgaard, T., Bredahl, T., & Puggaard, L. (2011). Exercise on prescription: changes in physical activity and health-related quality of life in five Danish programmes. *European Journal of Public Health*, *21*(1), 56–62. doi: 10.1093/eurpub/ckq003

Sowden, S., & Raine, R. (2008). Running along parallel lines: how political reality impedes the evaluation of public health interventions. A case study of exercise referral schemes in England. *Journal of Epidemiology and Community Health*, *62*(9), 835–841. doi: 10.1136/jech.2007.069781

Sowden, S., Breeze, E., Barber, J., & Raine, R. (2008). Do general practices provide equitable access to physical activity interventions? *British Journal of General Practice*, *58*(555), e1–e8.

Sparkes, A., & Smith, B. (2014). *Qualitative research methods in sport, exercise and health: from process to product*. London: Routledge.

Sport England. (2015). *Active people survey, 2012-2013. 2nd edition.* Retrieved from http://doc.ukdataservice.ac.uk/doc/7493/mrdoc/pdf/7493_aps_7_technical_report.pdf

Sport England. (2016). *Towards an active nation: strategy 2016-21*. Retrieved from https://www.sportengland.org/media/10629/sport-england-towards-an-active-nation.pdf

Sport England. (2017). *Active lives survey 2015-16 year 1 report*. Retrieved from https://www.sportengland.org/media/11498/active-lives-survey-yr-1-report.pdf

Sport England. (2018). *Active lives adult survey: November 16/17 report*. Retrieved from https://www.sportengland.org/media/13217/v-mass-markets-digital-content-editorial-team-active-lives-march-2018-active-lives-adult-survey-nov-16-17-final.pdf

Spring, B., Moller, A., & Coons, M. (2012). Multiple health behaviours: overview and implications. *Journal of Public Health*, *34*(S1), 3–10. doi: 10.1093/pubmed/fdr111

Stanaway, J., Afshin, A., Gakidou, E., Lim, S., Abate, D., Abate, K., … Murray, C. (2018). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, *392*(10159), 1923–1994. doi: 10.1016/S0140-6736(18)32225-6

Stathi, A., McKenna, J., & Fox, K. (2004). The experiences of older people participating in exercise referral schemes. *The Journal of the Royal Society for the Promotion of Health*, *124*(1), 18–23. doi: 10.1177/146642400312400108

Steckler, A., McLeroy, K., Goodman, R., Bird, S., & McCormick, L. (1992). Toward integrating qualitative and quantitative methods: an introduction. *Health Education Quarterly*, *19*(1), 1–8. doi: 10.1177/109019819201900101

Steel, N., Ford, J., Newton, J., Davis, A., Vos, T., Naghavi, M., … Murray, C. (2018). Changes in health in the countries of the UK and 150 English local authority areas 1990–2016: a systematic analysis for the global burden of disease study

2016. *The Lancet, 392*(10158), 1647-1661. doi: 10.1016/S0140-6736(18)32207-4

Steene-Johannessen, J., Anderssen, S., Van der Ploeg, H., Hendriksen, I., Donnelly, A., Brage, S., & Ekelund, U. (2016). Are self-report measures able to define individuals as physically active or inactive? *Medicine & Science in Sports & Exercise*, *48*(2), 235–244. doi: 10.1249/mss.0000000000000760

Sterne, J., White, I., Carlin, J., Spratt, M., Royston, P., Kenward, M., … Carpenter, J. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, *338*, b2393. doi: 10.1136/bmj.b2393

Stevens, W., Hillsdon, M., Thorogood, M., & McArdle, D. (1998). Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old men and women: a randomised controlled trial. *British Journal of Sports Medicine*, *32*(3), 236–241.

Stewart, L., Dolan, E., Carver, P., & Swinton, P. (2017). Per-protocol investigation of a best practice exercise referral scheme. *Public Health*, *150*, 26–33. doi: 10.1016/j.puhe.2017.04.023

Stewart-Brown, S., Platt, S., Tennant, A., Maheswaran, H., Parkinson, J., Weich, S., … Clarke, A. (2011). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a valid and reliable tool for measuring mental well-being in diverse populations and projects. *Journal of Epidemiology and Community Health*, *65*(S2), A38–39. doi: 10.1136/jech.2011.143586.86

Stokols, D. (2000). Social ecology and behavioral medicine: implications for training, practice, and policy. *Behavioral Medicine*, *26*(3), 129–138.

Stranges, S., Samaraweera, P., Taggart, F., Kandala, N., & Stewart-Brown, S. (2014). Major health-related behaviours and mental well-being in the general population: the health survey for England. *BMJ Open*, *4*, e005878. doi: 10.1136/bmjopen-2014-005878

Strath, S., Kaminsky, L., Ainsworth, B., Ekelund, U., Freedson, P., Gary, R., … Swartz, A. (2013). Guide to the assessment of physical activity: clinical and research applications. *Circulation*, 128(20), 2259-2279. doi: 10.1161/01.cir.0000435708.67487.da

Sullivan, P., Lawrence, W., & Ghushchyan, V. (2005). A national catalog of preference-based scores for chronic conditions in the United States. *Medical Care*, *43*(7), 736–749.

Taggart, F., Friede, T., Weich, S., Clarke, A., Johnson, M., & Stewart-Brown, S. (2013). Cross cultural evaluation of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) - a mixed methods study. *Health and Quality of Life Outcomes*, *11*, 27–27. doi: 10.1186/1477-7525-11-27

Taggart, F, Stewart-Brown, S., & Parkinson, J. (2015). *Warwick-Edinburgh Mental Well-being Scale (WEMWBS) user guide – version 2*. Retrieved from http://www2.warwick.ac.uk/fac/med/research/platform/wemwbs/researchers/userguide/wemwbs_user_guide_jp_02.02.16.pdf

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology – combining qualitative and quantitative approaches*. London: SAGE Publications.

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioural research*. London: SAGE Publications.

Taylor, A., Doust, J., & Webborn, N. (1998). Randomised controlled trial to examine the effects of a GP exercise referral programme in Hailsham, East Sussex, on modifiable coronary heart disease risk factors. *Journal of Epidemiology & Community Health*, *52*(9), 595–601.

Taylor, A., & Fox, K. (2005). Effectiveness of a primary care exercise referral intervention for changing physical self-perceptions over 9 months. *Health*

Psychology, *24*(1), 11–21.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling. *Journal of Mixed Methods Research*, *1*(1), 77–100. doi: 10.1177/2345678906292430

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Stephen, S., Weich, S., … Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health and Quality of Life Outcomes*, *5*(1), 63. doi: 10.1186/1477-7525-5-63

Thirsk, L., & Clark, A. (2017). Using qualitative research for complex interventions: the contributions of hermeneutics. *International Journal of Qualitative Methods*, *16*(1). doi: 10.1177/1609406917721068

Thomson, L., Camic, P., & Chatterjee, H. (2015). *Social prescribing: a review of community referral schemes*. Retrieved from https://create.canterbury.ac.uk/15655/1/Social_Prescribing_Review_2015.pdf

Thomson, S., Figueras, J., Evetovits, T., Jowett, M., Mladovsky, P., Maresso, A., … Kluge, H. (2014). *Economic crisis, health systems and health in Europe: impact and implications for policy*. Retrieved from http://www.euro.who.int/__data/assets/pdf_file/0008/257579/Economic-crisis-health-systems-Europe-impact-implications-policy.pdf

Tierney, M., Fraser, A., & Kennedy, N. (2012). Physical activity in rheumatoid arthritis: a systematic review. *Journal of Physical Activity and Health*, *9*(7), 1036–1048. doi: 10.1123/jpah.9.7.1036

Timmermans, E., van der Pas, S., Schaap, L., Sánchez-Martínez, M., Zambon, S., Peter, R., … Deeg, D. (2014). Self-perceived weather sensitivity and joint pain in older people with osteoarthritis in six European countries: results from the European project on osteoarthritis (EPOSA). *BMC Musculoskeletal Disorders*, *15*, 66–66. doi: 10.1186/1471-2474-15-66

Tobi, P., Estacio, E., Yu, G., Renton, A., & Foster, N. (2012). Who stays, who drops out? Biosocial predictors of longer-term adherence in participants attending an exercise referral scheme in the UK. *BMC Public Health*, *12*, 347. doi: 10.1186/1471-2458-12-347

Tobi, P., Kemp, P., & Schmidt, E. (2017). Cohort differences in exercise adherence among primary care patients referred for mental health versus physical health conditions. *Primary Health Care Research & Development*, *18*(5), 463–471. doi: 10.1017/S1463423617000214

Toomey, C., Cremona, A., Hughes, K., Norton, C., & Jakeman, P. (2015). A review of body composition measurement in the assessment of health. *Topics in Clinical Nutrition*, *30*(1), 16-32.

Torrance, G., Thomas, W., & Sackett, D. (1972). A utility maximization model for evaluation of health care programs. *Health Services Research*, *7*(2), 118–133.

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, J. Tanur & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: building a bridge between disciplines* (pp. 73-100). Washington: National Academy Press.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–883.

Tracy, S. (2010). Qualitative Quality: Eight “Big-Tent” Criteria for Excellent Qualitative Research. *Qualitative Inquiry*, *16*(10), 837–851. https://doi.org/10.1177/1077800410383121

Tremblay, M., Aubert, S., Barnes, J., Saunders, T., Carson, V., Latimer-Cheung, A., … Wondergem, R. (2017). Sedentary Behavior Research Network (SBRN) –

terminology consensus project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 75. doi: 10.1186/s12966-017-0525-8

Troiano, R. (2009). Can there be a single best measure of reported physical activity? *The American Journal of Clinical Nutrition*, *89*(3), 736–737.

Troiano, R., McClain, J., Brychta, R., & Chen, K. (2014). Evolution of accelerometer methods for physical activity research. *British Journal of Sports Medicine*, *48*(13), 1019–1023. doi: 10.1136/bjsports-2014-093546

Trost, J. (1986). Statistically nonrepresentative stratified sampling: a sampling technique for qualitative studies. *Qualitative Sociology*, *9*(1), 54–57.

Trost, S. (2007). State of the art reviews: measurement of physical activity in children and adolescents. *American Journal of Lifestyle Medicine*, *1*(4), 299–314.

Trost, S., McIver, K., & Pate, R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine & Science in Sports & Exercise*, *37*(S11), 531-543.

Trost, S., & O'Neil, M. (2014). Clinical use of objective measures of physical activity. *British Journal of Sports Medicine*, *48*(3), 178–181. doi: 10.1136/bjsports-2013-093173

Trueman, P., & Anokye, N. (2013). Applying economic evaluation to public health interventions: the case of interventions to promote physical activity. *Journal of Public Health*, *35*(1), 32–39. doi: 10.1093/pubmed/fds050

Tucker, P., & Gilliland, J. (2007). The effect of season and weather on physical activity: a systematic review. *Public Health*, *121*(12), 909–922. doi: 10.1016/j.puhe.2007.04.009

Tudor-Locke, C., Han, H., Aguiar, E., Barreira, T., Schuna Jr, J., Kang, M., & Rowe, D. (2018). How fast is fast enough? Walking cadence (steps/min) as a practical estimate of intensity in adults: a narrative review. *British Journal of Sports Medicine*, *52*(12), 776. doi: 10.1136/bjsports-2017-097628

Tuttle, M., Montoye, A., & Kaminsky, L. (2016). The benefits of body mass index and waist circumference in the assessment of health risk. *ACSM's Health & Fitness Journal*, *20*(4), 15-20.

UKactive. (2014). *Turning the tide of inactivity*. Retrieved from http://ukactive.com/downloads/managed/Turning_the_tide_of_inactivity.pdf

UKactive. (2018). *Exercise referral*. Retrieved from https://www.ukactive.com/projects/exercise-referral/

US Department of Health and Human Services. (2008). *2008 physical activity guidelines for Americans.* Retrieved from https://health.gov/paguidelines/pdf/paguide.pdf

Ussher, M., Taylor, A., & Faulkner, G. (2014). Exercise interventions for smoking cessation. *The Cochrane Database of Systematic Reviews*, *8*, CD002295. doi: 10.1002/14651858.CD002295.pub5

Valtorta, N., & Hanratty, B. (2012). Loneliness, isolation and the health of older adults: do we need a new research agenda? *Journal of the Royal Society of Medicine*, *105*(12), 518–522. doi: 10.1258/jrsm.2012.120128

van de Willige, G., Wiersma, D., Nienhuis, F., & Jenner, J. (2005). Changes in quality of life in chronic psychiatric patients: a comparison between EuroQol (EQ-5D) and WHOQoL. *Quality of Life Research*, *14*(2), 441–451.

van der Ploeg, H., Tudor-Locke, C., Marshall, A., Craig, C., Hagströmer, M., Sjöström, M., & Bauman, A. (2010). Reliability and validity of the International Physical Activity Questionnaire for assessing walking. *Research Quarterly for Exercise and Sport*, *81*(1), 97–101.

Van Dyck, D., Cardon, G., Deforche, B., & De Bourdeaudhuij, I. (2015). IPAQ interview version: convergent validity with accelerometers and comparison of physical activity and sedentary time levels with the self-administered version. *The Journal of Sports Medicine and Physical Fitness*, *55*(7–8), 776–86.

van Poppel, M., Chinapaw, M., Mokkink, L., van Mechelen, W., & Terwee, C. (2010). Physical activity questionnaires for adults. *Sports Medicine*, *40*(7), 565–600. doi: 10.2165/11531930-000000000-00000

Vartanian, T. (2011). *Secondary data analysis (pocket guide to social work research methods)* (1st ed.). Oxford: Oxford University Press.

Vasudevan, S., Maddocks, M., Chamberlain, S., Spinou, A., Wood, C., Jolley, C., & Birring, S. (2013). P197 physical inactivity in sarcoidosis. *Thorax*, *68*(S3), A165. doi: 10.1136/thoraxjnl-2013-204457.349

Victora, C., Habicht, J., & Bryce, J. (2004). Evidence-Based Public Health: Moving Beyond Randomized Trials. *American Journal of Public Health*, *94*(3), 400–405.

Vingård, E., Alexanderson, K., & Norlund, A. (2004). Chapter 9: consequences of being on sick leave. *Scandinavian Journal of Public Health*, *32*(S63), 207–215. doi: 10.1080/14034950410021899

Vinson, D., & Parker, A. (2012). Exercise, service and support: client experiences of physical activity referral schemes (PARS). *Qualitative Research in Sport, Exercise and Health*, *4*(1), 15–31. doi: 10.1080/2159676X.2011.653501

Vorrink, S., Kort, H., Troosters, T., & Lammers, J. (2011). Level of daily physical activity in individuals with COPD compared with healthy controls. *Respiratory Research*, *12*(1), 33–33. doi: 10.1186/1465-9921-12-33

Vos, T., Barber, R., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., … Murray, C. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, *386*(9995), 743–800. doi: 10.1016/S0140-6736(15)60692-4

Vos, T., Flaxman, A., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., … Murray, C. (2012). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010. *The Lancet*, *380*(9859), 2163–2196. doi: 10.1016/S0140-6736(12)61729-2

Walker, E., & Druss, B. (2016). A public health perspective on mental and medical comorbidity. *JAMA*, *316*(10), 1104–1105. doi: 10.1001/jama.2016.10486

Walker, E., McGee, R., & Druss, B. (2015). Mortality in mental disorders and global disease burden implications: A systematic review and meta-analysis. *JAMA Psychiatry*, *72*(4), 334–341. doi: 10.1001/jamapsychiatry.2014.2502

Walker, S., Read, S., & Priest, H. (2013). Use of reflexivity in a mixed-methods study. *Nurse Researcher*, *20*(3), 38–43.

Warburton, D., & Bredin, S. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Current Opinion in Cardiology*, *32*(5), 541–556. doi: 10.1097/hco.0000000000000437

Ward, D., Furber, C., Tierney, S., & Swallow, V. (2013). Using framework analysis in nursing research: a worked example. *Journal of Advanced Nursing*, *69*(11), 2423–2431. doi: 10.1111/jan.12127

Ward, M., Phillips, C., Farr, A., & Harries, D. (2010). Heartlinks—a real world approach to effective exercise referral. *International Journal of Health Promotion and Education*, *48*(1), 20–27. doi: 10.1080/14635240.2010.10708176

Warren, J., Ekelund, U., Besson, H., Mezzani, A., Geladas, N., & Vanhees, L. (2010). Assessment of physical activity – a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *European Journal of Cardiovascular Prevention & Rehabilitation*, *17*(2), 127–139. doi: 10.1097/HJR.0b013e32832ed875

Waters, L., Reeves, M., Fjeldsoe, B., & Eakin, E. (2012). Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. *Journal of Physical Activity and Health*, *9*(6), 884–895. doi: 10.1123/jpah.9.6.884

Weathington, B., Cunningham, C., & Pittenger, D. (2010). *Research methods for the behavioral and social sciences* (1st ed.). Hoboken: John Wiley & Sons.

Webb, R., Thompson, J., Ruffino, J., Davies, N., Watkeys, L., Hooper, S., … McDonnell, B. (2016). Evaluation of cardiovascular risk-lowering health benefits accruing from laboratory-based, community-based and exercise-referral exercise programmes. *BMJ Open Sport & Exercise Medicine*, *2*, 000089. doi: 10.1136/bmjsem-2015-000089

Weed, M. (2016). Evidence for physical activity guidelines as a public health intervention: efficacy, effectiveness, and harm – a critical policy sciences approach. *Health Psychology and Behavioral Medicine*, *4*(1), 56–69. doi: 10.1080/21642850.2016.1159517

Wen, C., Wai, J., Tsai, M., Yang, Y., Cheng, T., Lee, M., … Wu, X. (2011). Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *The Lancet*, *378*(9798), 1244–1253. doi: 10.1016/S0140-6736(11)60749-6

Whitehead, P., Worthington, E., Parry, R., Walker, M., & Drummond, A. (2015). Interventions to reduce dependency in personal activities of daily living in community dwelling adults who use homecare services: a systematic review. *Clinical Rehabilitation*, *29*(11), 1064–1076. doi: 10.1177/0269215514564894

Wilke, C., Pickard, A., Walton, S., Moock, J., Kohlmann, T., & Lee, T. (2010). Statistical implications of utility weighted and equally weighted HRQL measures: an empirical study. *Health Economics*, *19*(1), 101–110. doi: 10.1002/hec.1467

Williams, B., Mancia, G., Spiering, W., Agabiti Rosei, E., Azizi, M., Burnier, M., … Desormais, I. (2018). 2018 ESC/ESH guidelines for the management of arterial hypertension. *European Heart Journal*, *39*(33), 3021–3104. https://doi.org/10.1093/eurheartj/ehy339

Williams, N., Hendry, M., France, B., Lewis, R., & Wilkinson, C. (2007). Effectiveness of exercise-referral schemes to promote physical activity in adults: systematic review. *The British Journal of General Practice*, *57*(545), 979–986. doi: 10.3399/096016407782604866

Williams, O., & Gibson, K. (2018). Exercise as a poisoned elixir: inactivity, inequality and intervention. *Qualitative Research in Sport, Exercise and Health*, *10*(4), 412–428. doi: 10.1080/2159676X.2017.1346698

Williams, S., & French, D. (2011). What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour—and are they the same? *Health Education Research*, *26*(2), 308–322. doi: 10.1093/her/cyr005

Wills, T. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, *90*(2), 245–271.

Wilson, D. (2015). Behavior matters: the relevance, impact, and reach of behavioral medicine. *Annals of Behavioral Medicine*, *49*(1), 40–48. doi: 10.1007/s12160-014-9672-1

Wilson, M., Kellock, C., Adams, D., & Landsberg, J. (2015). *The Scottish health* *survey topic report: mental health and wellbeing*. Retrieved from http://www.gov.scot/Resource/0046/00469088.pdf

Wofford, M., Harkins, K., King, D., Habeeb, G., Wyatt, S., & Jones, D. (2002). P-175: Accurate cuff size in blood pressure measurement. *American Journal of Hypertension*, *15*(S3), 92A.

Wolin, K., Fagin, C., Ufere, N., Tuchman, H., & Bennett, G. (2010). Physical activity in US Blacks: a systematic review and critical examination of self-report instruments. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 73. doi: 10.1186/1479-5868-7-73

Woods, C., & Mutrie, N. (2012). Putting physical activity on the policy agenda. *Quest*, *64*(2), 92–104. doi: 10.1080/00336297.2012.669318

World Health Organization. (2000). *Obesity: preventing and managing the global epidemic*. Retrieved from http://apps.who.int/iris/bitstream/10665/42330/1/WHO_TRS_894.pdf?ua=1&ua=1

World Health Organization. (2006). *Basic documents, forty-fifth edition, supplement, october 2006: constitution of the World Health Organization*. Retrieved from http://www.who.int/governance/eb/who_constitution_en.pdf

World Health Organization. (2009). *Global health risks: mortality and burden of disease attributable to selected major risks.* Retrieved from www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf

World Health Organization. (2010). *Global recommendations on physical activity for health*. Retrieved from http://apps.who.int/iris/bitstream/10665/44399/1/9789241599979_eng.pdf

World Health Organization. (2011). *Global status report on noncommunicable diseases 2010*. Retrieved from www.who.int/nmh/publications/ncd_report_full_en.pdf

World Health Organization. (2018). *Global action plan on physical activity 2018–2030: more active people for a healthier world*. Retrieved from http://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf

Wormald, H., & Ingle, L. (2004). GP exercise referral schemes: improving the patient's experience. *Health Education Journal*, *63*(4), 362–373. doi: 10.1177/001789690406300407

Wormald, H., Waters, H., Sleap, M., & Ingle, L. (2006). Participants' perceptions of a lifestyle approach to promoting physical activity: targeting deprived communities in Kingston-Upon-Hull. *BMC Public Health*, *6*(1), 1–12. doi: 10.1186/1471-2458-6-202

Wye, L., Brangan, E., Cameron, A., Gabbay, J., Klein, J., & Pope, C. (2015). Evidence based policy making and the 'art' of commissioning – how English healthcare commissioners access and use information and academic research in 'real life' decision-making: an empirical qualitative study. *BMC Health Services Research*, *15*(1), 430. doi: 10.1186/s12913-015-1091-x

Yates, T., Henson, J., Edwardson, C., Bodicoat, D., Davies, M., & Khunti, K. (2015). Differences in levels of physical activity between White and South Asian populations within a healthcare setting: impact of measurement type in a cross-sectional study. *BMJ Open*, *5*, e006181. doi: 10.1136/bmjopen-2014-006181

Zilioli, S., Slatcher, R., Ong, A., & Gruenewald, T. (2015). Purpose in life predicts allostatic load ten years later. *Journal of Psychosomatic Research*, *79*(5), 451–457. doi:10.1016/j.jpsychores.2015.09.013