University of Southampton

Faculty of Engineering and Physical sciences

Computational Engineering and Design Research Group

SMARTPHONES FOR CONTINUOUS ASSESSMENT OF PHYSICAL ACTIVITY IN HEALTHCARE

by

Jimmy Caroupapoullé

ORCID ID 0000-0002-4763-5957

PhD Thesis

November 2022

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences

Doctor of Philosophy

SMARTPHONES FOR CONTINUOUS ASSESSMENT OF PHYSICAL ACTIVITY IN HEALTHCARE

by

Jimmy Caroupapoullé

The benefits of exercise and physical activity are well-known and documented. However, most recommendations do not consider the stress caused on the joint and bone structure in the context of musculoskeletal conditions. A committee of selected experts recognise the importance of measuring the ground force reaction on the lower limbs. However, measuring load rate estimates is not standardised and requires specialised facilities.

The primary aim of this project was to develop a framework to measure load rate outside of the clinical setting. A methodology was developed to monitor load rate estimates using smartphones, as cheap and broadly available technology to achieve this. The method was used to investigate the reliability in ideal conditions, as part of a clinical trial and over an extended timeframe.

The novelty of this project was to develop a protocol using smartphones as a surrogate of the lower limbs to monitor patients affected by musculoskeletal conditions. The key finding was that significant challenges are associated with using smartphones, but passive monitoring can be achieved to record substantial amounts of data without patient input. Secondly, load rate provides more qualitative details on the physical activity than measures such as energy expenditure, step count and met-min. However, the interpretation and visualisation of load rate are more complex.

To conclude, this thesis tested and verified a methodology that can continuously monitor load rate estimates on the lower limbs. This methodology addresses the need to measure joint loading in activity and time, which can further be used in healthcare for musculoskeletal diseases.

Table of Contents

Table of Contents ..... **i**

Table of Tables ..... **vii**

Table of Figures ..... **ix**

Research Thesis: Declaration of Authorship ..... **xiii**

Acknowledgements ..... **xv**

Definitions and Abbreviations ..... **xvi**

Chapter 1 Introduction **19**

1.1 Overview 19

1.2 Thesis Structure 20

Chapter 2 Literature review **21**

2.1 Musculoskeletal conditions overview 21

2.2 Musculoskeletal diseases & physical activity 23

2.2.1 Personalised recommendations 23

2.2.2 Physical activity and bones 25

2.2.3 Physical activity and joints 26

2.3 Physical Activity & load forces 27

2.4 Smartphones Applications & remote monitoring 30

2.5 Project Hypothesis 33

Chapter 3 General Methodology **34**

3.1 Study design 34

3.2 Ethical approval & Participants 35

3.3 Equipment 36

3.4 Smartphone considerations 37

3.4.1 Smartphone architecture & sensors 37

3.4.2 Software development lifecycle 39

3.5 Data collection & analysis 45

Chapter 4 RApp1: Proof of concept with rheumatoid arthritis patients **49**

4.1 Introduction 49

4.2 Aim & Objectives 52
4.2.1 Aim 52
4.2.2 Objectives 52
4.2.3 Hypothesis 52
4.3 Methodology 52
4.3.1 RApp™ design 52
4.3.2 Patient recruitment and study procedure 53
4.3.3 Data collection 55
4.3.4 Data Analysis 55
4.4 Results 56
4.4.1 Patients' perception and feedback 56
4.4.2 Patients using RApp™ 62
4.4.3 Continuous monitoring (RApp™) 71
4.4.3.1 Sample summary 71
4.4.3.2 Smartphone movements 72
4.4.3.3 Fitbit steps and smartphone´s LRE 78
4.5 Discussion 81
4.5.1 Patient´s perception 81
4.5.2 Continuous monitoring 82
4.6 Conclusion 83

Chapter 5 OApp Southampton: Continuous monitoring with power saving disabled . 85

5.1 Introduction 85
5.2 Literature review 85
5.2.1 Smartphone´s battery optimisation 85
5.2.2 Smartwatch & tracker position 88
5.2.3 Rationale 89
5.3 Aim & objectives 90
5.3.1 Aim 90
5.3.2 Objectives 90

5.3.3 Hypothesis 90
5.4 Methodology 90
5.4.1 Patient recruitment and study procedure 90
5.4.2 Data collection 91
5.4.3 Data analysis 92
5.4.4 Individual phenotype 92
5.5 Results 93
5.5.1 Continuous monitoring 93
5.5.2 Smartphone & smartwatch comparison 96
5.6 Discussion 100
5.6.1 Smartphone monitoring 100
5.6.2 Smartphone & smartwatch correlation 102
5.6.3 Study considerations 105
5.7 Conclusion 106
Chapter 6 OApp Sydney: Continuous monitoring in osteoarthritis clinical trial 107
6.1 Introduction 107
6.2 Literature review 107
6.2.1 Osteoarthritis and clinical treatments 107
6.2.2 Wearable and step count estimates 109
6.3 Rationale 110
6.4 Aim & objectives 111
6.4.1 Aim 111
6.4.2 Objectives 111
6.4.3 Hypothesis 111
6.5 Methodology 111
6.5.1 Patient recruitment and study procedure 111
6.5.2 Data collection 112
6.5.3 Data analysis 113
6.6 Results 113

6.6.1 Overall participation....113
6.6.2 LRE Monitoring....114
6.6.3 Step count monitoring....117
6.6.4 Monitoring before & after injection....118
6.7 Discussion....122
6.7.1 Continuous monitoring....122
6.7.2 Step count comparison....123
6.7.3 Behaviour changes before/after injection....123
6.8 Conclusion....124
Chapter 7 RApp2: Long-term monitoring....126
7.1 Introduction....126
7.2 Aim & objectives....126
7.2.1 Aim....126
7.2.2 Objectives....126
7.2.3 Hypothesis....127
7.3 Methodology....127
7.3.1 Patient recruitment and study procedure....127
7.3.2 Data collection and analysis....128
7.4 Results....128
7.4.1 Patient´s continuous monitoring (3 months)....128
7.4.2 Continuous monitoring over 5 years....133
7.5 Discussion....140
7.6 Conclusion....141
Chapter 8 Conclusions and future research....142
8.1 General Discussion....142
8.2 Research contribution and future work....143
8.3 Summary....145
Appendix A PPI Questionnaire....146
Appendix B PPI survey summary....151

Appendix C Clinic shadowing notes 156

Appendix D Study 1: Patient feedback 157

Appendix E Study 2: SP & SW (x, y and z) 158

Appendix F Study 2: SP & SW jerk (mean) 163

Appendix G Study 2: SP & SW Phenotypes 167

Appendix H Study 3: Patient list 171

Appendix I Study 3: Phenotype characteristics 173

Appendix J NHRA Docs 174

Appendix K BSR (April 2017) Poster & Abstract 175

Appendix L Git repository and best practices 178

Bibliography 183

Table of Tables

Table 1: Smartphone app feature review 31

Table 2: Study summary 35

Table 3: Patient recruitment summary 63

Table 4: Summary Visit 1 (Introduction) 64

Table 5: Summary Visit 2 (Follow-up after six weeks) 64

Table 6: Summary Visit 3 (Closing interview) 66

Table 7: Lines of sensor readings per day 94

Table 8: Amount of files collected per day 95

Table 9: Lines of sensor readings per file 96

Table 10: Phenotype characteristics 100

Table 11: SP & SW Correlation 104

Table 12: Summary samples (SP LR estimates and Fitbit) across patients 114

Table 13: Correlation smartphone and Fitbit (daily step count) 118

Table 14: Patient recruitment summary 128

Table 15: Phenotype characteristics 131

Table 16: Smartphone devices summary 134

Table 17: Monthly phenotype characteristics 140

Table of Figures

Figure 1: Study summary 21

Figure 2: MSK impact on the bone and joint structure 21

Figure 3: Under Vs Overloading 27

Figure 4: Measure of mechanical loading (1) via implant; (2) force plate; (3) accelerometer 28

Figure 5: Smartphone architecture and sensor growth trend 38

Figure 6: Android sensors (Google, 2016) 39

Figure 7: Branching strategy 40

Figure 8: Screenshot of Android Development Tool (ADT) 41

Figure 9: App architecture & Lifecycles 42

Figure 10: App lifecycle 42

Figure 11: Settings and OApp™ navigation 43

Figure 12: RApp™ navigation 43

Figure 13: RApp™ questionnaire 44

Figure 14: RApp™ DAS28 45

Figure 15: Storage and data format 46

Figure 16: Cloud architecture 47

Figure 17: DAS28 Form and EULAR response (source: DAS-score.nl (DAS-Score.nl, 2009)) 49

Figure 18: DAS28, SDAI and CDAI score calculation 50

Figure 19: RAPID3 Score (source: American College of Rheumatology (Rheumatology, 2014)) 51

Figure 20: RApp™ Layout (PPI session) 53

Figure 21: Southampton General Hospital (Victoria House) 57

Figure 22: Pre-session survey questions 1.1 to 1.4 (9 RA patients) 59

Figure 23: Pre-session survey questions 2.1 to 2.5 (9 RA patients) 60

Figure 24: Pre-session survey questions 2.6 to 2.9 (9 RA patients).....61

Figure 25: Pre-session survey questions 2.10 to 2.11 (9 RA patients).....62

Figure 26: PPI comments pre and post an overview of RApp™.....62

Figure 27: BSR & RAPID3.....67

Figure 28: Summary Self Assessed Pain.....69

Figure 29: Summary Self Assessed Swelling 70

Figure 30: Smartphone's sensor Patient 1.....73

Figure 31: Smartphone's sensor Patient 2.....74

Figure 32: Smartphone's sensor Patient 3.....75

Figure 33: Smartphone's sensor Patient 4.....76

Figure 34: Smartphone's sensor Patient 5.....77

Figure 35: Daily Step count and Jerk (Patient 1).....78

Figure 36: Daily Step count and Jerk (Patient 2).....79

Figure 37: Daily Step count and Jerk (Patient 3).....79

Figure 38: Daily Step count and Jerk (Patient 4).....80

Figure 39: Daily Step count and Jerk (Patient 5).....80

Figure 40: Qualcomm chipset approach to power optimisation (Qualcomm, 2013).....86

Figure 41: LAU & FD 87

Figure 42: Android Power management.....88

Figure 43: Smartphone & Smartwatch 91

Figure 44: Example of individual phenotypes.....93

Figure 45: SP & SW raw acceleromter (x, y and z) over time (Participant 2).....97

Figure 46: SP & SW raw acceleromter (x, y and z) over time (Participant 3).....97

Figure 47: SP & SW raw acceleromter (x, y and z) over time (Participant 10).....98

Figure 48: SP & SW jerk (mean) over time (Participant 2).....98

Figure 49: SP & SW jerk (mean) over time (Participant 7).....99

Figure 50: SP & SW jerk (mean) over time (Participant 9).....99

Figure 51: SP & SW Phenotype (Participant 10) 100

Figure 52: SP sample count and sample rate per hour.....101

Figure 53: SW sample count and sample rate per hour 102

Figure 54: SP & SW raw accelerometer (x, y and z) over time (Participant 10).....103

Figure 55: SP&SW LR (mean) correlation 104

Figure 79: Smartphone & Fitbit 112

Figure 57: Data Sampling 113

Figure 58: LRE (mean) per hour 115

Figure 59: Count of LRE per hour.....115

Figure 60: LRE (mean) per hour (raw and interpolated).....116

Figure 61: Raw & Interpolated Phenotype (Patients 47 and 55) 116

Figure 62: Phenotype characteristics.....117

Figure 63: Daily step count correlation.....118

Figure 64: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 3).119

Figure 65: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 36)119

Figure 66: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 44)119

Figure 67: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 45)119

Figure 68: Boxplot of Step count across all patients, before and after injection 120

Figure 69: Boxplot LRE for all patients, before and after injection.....121

Figure 70: LRE and Sample rate 122

Figure 71: Smartphone's LRE (Mean) over time 129

Figure 72: SP sample count and sample rate per hour 130

Figure 73: LR (mean) estimates per hour (raw and interpolated) 130

Figure 74: Boxplot (with and without outliers) 131

Figure 75: Phenotype (patient p73) 132

Figure 76: Phenotype (patient p74) 132

Figure 77: Phenotype (patient p75) 133

Figure 78: Smartphone devices and samples recorded over time 134

Figure 79: Smartphone's LRE (Mean) over time (2018-2020) 135

Figure 80: Smartphone's LRE (Mean) over time (2021-2022) 136

Figure 81: Count of LRE per hour 137

Figure 82: LRE (mean) per hour 137

Figure 83: LRE (mean) per hour (raw and interpolated) 138

Figure 84: Year on Year Phenotype (2018 to 2022) 138

Figure 85: Boxplot Year on Year (2018 to 2022) 139

Research Thesis: Declaration of Authorship

Print name: Jimmy Caroupapoullé

Title of thesis: Smartphones for continuous assessment of physical activity in healthcare

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signature: Date:

Acknowledgements

First, I would like to thank my supervisors for sharing their wealth of knowledge and expertise. I have learned many new engineering and statistical computing concepts while expanding my understanding of healthcare challenges. This project would not have been possible without their help, support and insights.

I would like to thank Arthritis research UK for funding parts of my PhD and Microsoft for granting me access to a broad range of Azure cloud services.

I am extremely grateful to the Southampton General Hospital team. It would not have been possible to follow the best practices, obtain the relevant ethics approval and recruit patients without their support and guidance.

And last but not least, I would like to thank my wife, Aslı. Her belief in me has kept my spirits and motivation high during this process.

Definitions and Abbreviations

ACR	American College of Rheumatology
ADT	Android Development Tools
API	Application programming interface
APK	Android application package
APP	Acute-phase proteins
APR	Acute-phase reactants
BSR	British Society for Rheumatology
CRF	Clinical Research Facility
CRP	C reactive protein
CMS	Continuous Monitoring Score
CRUD	Create, Read, Update, Delete
DAS	Disease Activity Score
DMARD	Disease Modifying Anti Rheumatic
ESR	Erythrocyte Sedimentation Rate
EULAR	European League against Rheumatism
GCP	Good Clinical Practice
GH	Global Health
GPS	Global Positioning System
HAQ	Health Assessment
HAQ-DI	Health Assessment Disability Index
HREC	Health Research Ethics Committee
HTTPS	Hypertext Transfer Protocol Secure
HOOS	Hip disability and osteoarthritis outcome score

IDE	integrated development environment
IoT	Internet of Things
KOOS	Knee Osteoarthritis Outcome Score
LRE	Load Rate Estimate
MDHAQ	Multi-Dimensional HAQ
MEMS	Micro-electromechanical systems
MRI	Magnetic resonance imaging
MSK	Musculoskeletal
NHRA	National Health Regulatory Authority
NHS	National Health Service
NRAS	National Rheumatoid Arthritis Society
NICE	National Institute for Health and Care Excellence
NSAID	Nonsteroidal anti-inflammatory drugs
OA	Osteoarthritis
OS	Operating System
PRO	Patient-reported Outcome
PW	Power
RA	Rheumatoid Arthritis
RAPID3	Routine Assessment of Patient Index Data 3
REC	Research Ethics Committee
SC	Step Count
SDK	Software Development Kit
SGH	Southampton General Hospital
SQL	Structured Query Language

TNF Tumour Necrosis Factor

TRIFoRM TRust in IT: Factors, metRics, Models

USB Universal Serial Bus

WOMAC Western Ontario and McMaster Universities Arthritis Index

Chapter 1 Introduction

1.1 Overview

Musculoskeletal (MSK) conditions affect over 20 million people in the UK (ArthritisResearchUK, 2021). They account for the third most significant area of NHS England's objectives and budget mandated by the government (Gov.UK, 2018a), with a program spent of £4.7 billion following the policy paper published by the Department of Work and Pensions (Gov.uk, 2017). Musculoskeletal comprises diseases affecting the motion ability of joints (e.g., inflammatory conditions such as rheumatoid arthritis affecting the joints), cartilage (e.g., osteoarthritis) and bone density (e.g. osteoporosis) (gov.uk, 2021).

The risks of developing osteoarthritis (OA) or rheumatoid arthritis (RA) increases with age, with the highest prevalence being seen among individuals aged between 40 and 60 years old years and even higher in those aged over 70 (NAO, 2009), but it can affect any age. According to the State of Musculoskeletal Health 2021 report (ArthritisResearchUK, 2021), over 8 million people are diagnosed with OA, which represents 18.2% of the population of adults aged over 45 in England; and over 400k people are currently diagnosed with RA in the UK which represents 0.84% of adults aged over 16 in England. Women are more prevalent than men, and being diagnosed with RA is two to three times more common in women than men (O'Fallon, 2001). Smokers' risk of developing RA increases (Sugiyama et al., 2009), (Kallberg et al., 2010). Obesity and unhealthy weight/BMI are known factors that increase the risks of developing OA and RA 15% higher for RA when overweight and 21-31% when obese (Zhong et al., 2015, Feng et al., 2016).

Staying physically active and exercising generally improves well-being but the mechanical stress caused on the bones and joints that affects MSK conditions is not currently monitored outside of clinical setting, nor standardised, and therefore "healthy" load is not characterised. This project aims at assessing the feasibility in using smartphones to estimate mechanical loads on the lower limb of MSK patients in free living conditions. The current fitness solutions measure step counts and energy expenditure but do not consider the mechanical loads. Smartphone sensor data is available but continuous tracking relies on user intervention, e.g. to proactively start an app. Therefore, we built a purpose-based framework that records smartphones' sensor data to estimate loads without user intervention. This approach provides greater visibility on the data recorded, control over the sampling parameter and flexibility to define a user experience acceptable for OA and RA patients. The novelty of this project is to use smartphones, as broadly

available technology, to estimate mechanical load in free living conditions, which is not currently monitored although recognised in MSK guidelines.

1.2 Thesis Structure

The overall aim is to investigate how continuously monitoring load rates could be used as valuable fitness insight to investigate the relationship between an individual's physical activity and disease activity. The objective is to develop a smartphone framework with different front end for RA and OA (respectively RApp™ and OApp™). This project is cross-disciplined between Engineering and Life Sciences and, therefore, incorporates standard practices used in both fields of expertise in the study design.

This thesis continues with a literature review detailing MSK diseases and their relationship with physical activity and ground force reaction on the lower limbs in Chapter 2. The thesis is structured in exploratory studies, so each study's rationale is detailed in the corresponding chapters. Chapter 3 describes the study methodology developed as well as the framework of data capture and analysis. The subsequent studies (see fig 1) assess the feasibility and results, with chapter 4 (study 1) being the first proof of concept with rheumatoid arthritis patients. Chapter 5 (study 2) reviews the smartphone tracking capabilities used in optimal conditions with healthy volunteers. Chapter 6 (study 3) examines the results obtained in a clinical trial with OA patients. Finally, chapter 6 (study 4) reviews the monitoring over an extended timeframe with RA patients.

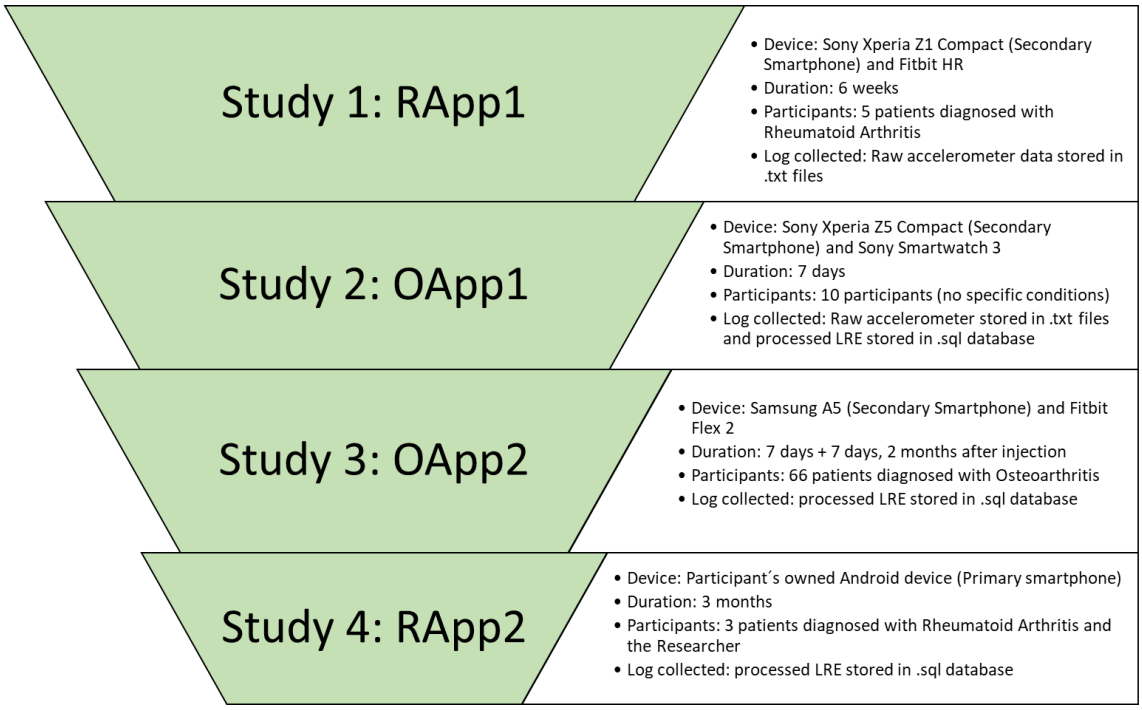


Figure 1: Study summary

Chapter 2 Literature review

2.1 Musculoskeletal conditions overview

The Musculoskeletal (MSK) system refers to the muscles and bones typically grouped as the back, upper and lower limbs (Agrawal, 2019). According to the WHO, more than 150 MSK conditions (National Academies of Sciences, 2020) affect the locomotory part of the human body of people from all demographic groups and ages. MSK conditions include osteoporosis which lowers bone density and all forms of arthritis which target the joints and cartilage.

Osteoporosis constantly affects the bone's ability to grow and rebuild, reducing bone mineral density. As illustrated in figure 2, more bone is lost than formed, so the internal structure deteriorates and becomes more porous (Australia, 2021). The weakened structure's strength increases fracture risk.

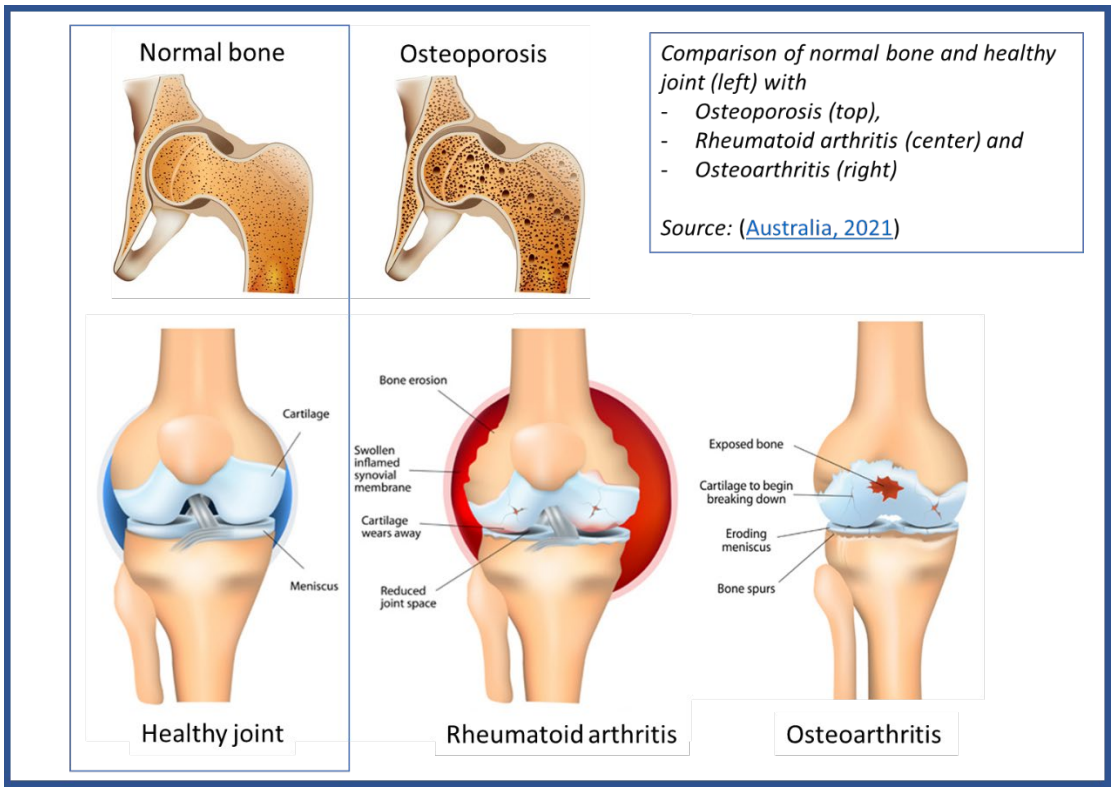


Figure 2: MSK impact on the bone and joint structure

Arthritis includes disease conditions such as lupus, gout and ankylosing spondylitis that commonly affects joints and specific organs, leading to pain, swelling, stiffness and decreased range of motion. Osteoarthritis (OA) and rheumatoid arthritis (RA) are the most common types of arthritis

and cause thinning of the cartilage. OA is a non-inflammatory disease characterised by thinner cartilage causing friction with the bone ends. RA is a chronic inflammatory in which the immune system attacks the synovia-lined joints, causing inflammation and destruction of the synovial membrane, which can lead to severe disability. The outer covering of the joints becomes the target of the body's immune system causing swelling and change in joint shape, ultimately leading to a breakdown of the bone and cartilage.

Osteoporosis does not cause pain and so is considered a silent disease. The changes in the bone structure become apparent if the fracture caused or visible change in posture, height or breathing. In contrast, OA and RA's main symptoms are joint pain and stiffness, most commonly affecting the hands, wrists, knees and feet. OA can affect any joint, while RA typically affects the joints symmetrically. Early diagnosis can help significantly prevent disease progression and bone erosion by starting medication and treatment earlier (Heidari, 2011). According to a meta-analysis study (Fautrel et al., 2017), early therapeutic intervention may also significantly reduce the risk of RA onset for arthritis patients. However, the guidelines for musculoskeletal system assessment (ArthritisResearchUK, 2019), as well as the criteria defined by the American Rheumatism Association in 1987 (Association, 1988), had limitations in predicting RA for patients with early arthritis (Alain Saraux, 2001). Besides, it is not always possible for patients to be seen by specialists, and in some cases, symptoms experienced can be due to biomechanical loading on the body (Shakoor and Moisio, 2004). OA can be diagnosed through a physical examination of the joint, looking into the creaking, tenderness, movement restriction and weakness or thinning of the supporting muscles. X-Rays can then help to check if the symptoms are due to a fracture or calcium deposit in the joint, and MRI scans of the knee or hip can identify bones or other possible joint problems. Unless a fracture occurs, Osteoporosis does not present visible symptoms and is diagnosed using dual-energy X-ray absorptiometry (DXA) scans. The results are reported as T-scores (and Z-scores) to compare the standard deviation of bone mineral density (in g/cm2) with the equivalent healthy person of the same age, sex and ethnicity (Kanis et al., 1997). Blood tests are not usually needed but will usually be conducted to evaluate other arthritis. In 2010, the RA disease classification criteria were reviewed by ACR and EULAR to consider more emphasis on laboratory values. Including serology (examination of blood serum) and acute-phase reactants, that is, a measure of plasma variations of acute-phase proteins (APP/APR) (Kay and Upchurch, 2012) throughout symptoms of more or less than six weeks.

Bone loss is a normal part of the ageing process, and no treatments to date can completely cure osteoarthritis and rheumatoid arthritis. But medications and supportive treatments are available to reduce bone loss, joint damage and inflammation. Osteoporosis patients use calcium and vitamins to supplement bone mass. Drug therapy is based on the probability of fracture using

Bisphosphonates and hormones such as Parathyroid, Selective oestrogen receptor modulators (SERMs) to maintain and slow the rate of bone loss. RA patients use disease modifying anti rheumatic (DMARD) drugs such as azathioprine, gold injections, hydroxychloroquine, leflunomide, methotrexate and sulfasalazine to address joint inflammation and subsequent damage by reducing the effects of the immune system attacking synovial joints. Anti-tumour necrosis factor (Anti-TNF) drugs such as adalimumab, certolizumab pegol, etanercept, golimumab and infliximab target TNF proteins specifically while other biological therapies including abatacept, rituximab and tocilizumab target different proteins. In the first part of the treatment, DMARDs are usually provided in the form of tablets and might be used as a combination to increase the effects. Prescription of biological therapies occurs only if DMARDs have not worked or have side effects. These newer drugs are provided by injections, which act faster by stopping specific proteins and chemicals in the blood responsible for activating the immune system from attacking the joints.

Besides addressing disease progression, patients concentrate on addressing pain and flares to help in daily activities, which can be managed with simple analgesia such as codeine (co-codamol and paracetamol). NSAID (non-steroidal anti-inflammatory drugs) such as Ibuprofen and Naproxen or alternative COX-2 inhibitors, including ascelecoxib or etoricoxib relieve both pain and inflammation but with an increased risk of internal bleeding. Corticosteroids are powerful medications for short-term pain relief, stiffness and swelling, but their usage should be limited due to side effects such as weight gain, easy bruising, muscle weakness and thinning of the skin or the bones.

2.2 Musculoskeletal diseases & physical activity

2.2.1 Personalised recommendations

The priority for patients is to manage their pain and reduce their medication. Thus instructions beyond prescriptions are expected to achieve their goals (Leach, 2018). The benefits of physical activity and regular exercise are known to improve heart and muscle strength and are promoted across most specialised online resources such as the NHS (NHS, 2018) and NRAS (NRAS, 2014). As of 2010, physical activity was not included in the top 10 recommendations developed on the matter of treating rheumatoid arthritis (Gomez et al., 2010), but in 2017, the EULAR recommendations for pain management for inflammatory and osteoarthritis (Austin et al., 2018) included physical activity and exercise interventions due to their positive effects on pain. Considering the benefits, a task force has been organised to review the literature and agree that physical activity should now be advocated as standard care for people with rheumatoid and musculoskeletal diseases (Kiltz et al., 2018, Verhoeven et al., 2016).

Nevertheless, arthritis patients are much less active than the general population (van den Berg et al., 2007). A survey of 5235 patients across 21 countries has shown that only a small proportion of RA patients exercise regularly (Sokka et al., 2008). Early survey results amongst people with arthritis (Gecht et al., 1996) have shown that physical activity is directly related to the patient's understanding of the benefits and ability to perform. Yet, studies have shown that even when the benefits are understood, the anxiety associated with pain remains a barrier for patients to stick to such programmes, and the adherence rate remains low (Vervloesem et al., 2012). According to a systematic review of 20 studies (Kirsten Jack, 2010), the barriers to treatment adherence in musculoskeletal physiotherapy are associated with motivational and psychological factors such as anxiety, depression, and social and family support. Low physical activity at baseline and worsening pain during exercise supplement the logistics challenges caused by work schedules, lack of time and financial constraints. A qualitative study (Wang et al., 2014) has identified a fear that joint damage and infection symptoms might increase with exercise when under medication to manage existing pain. Pain and drugs mediate the ability to exercise. Still, misapprehensions and conflicting information received from healthcare professionals are a source of frustration for all participants not being able to engage in exercise. "Rheumatoid cachexia" is a severe symptom of RA characterised by accelerated loss of muscle mass, a progression of cardiovascular disease and fatigue, contributing to functional limitation, disability, comorbidities, and reduced quality of life. Those symptoms might be lessened through regular exercise. A study across 39 patients with chronic heart failure has demonstrated that moderate-intensity resistance exercise training for three months produces constructive changes to skeletal muscle strength and endurance (Hare et al., 2004). Besides, a personalised exercise program can have a positive impact not only in slowing down the progress of arthritis but also on the patient's overall well-being (Jennifer K. Cooney, 2010). A study has found improvement in the microvascular and macrovascular as well as disease characteristics because of using a personalised aerobic program on a cohort of 40 RA patients (Metsios et al., 2014).

Therefore, HCP must personalise their recommendations to improve patients´ physical fitness considering the components of PA which have been reviewed (Kell et al., 2001). Muscular strength is the maximum force that can be generated, which is instrumental in performing daily activities. Endurance is the maximum amount of time that an action can be repeated over time and is typically associated with injury and falls risks. Flexibility defines the range of motion that can be covered. A well-designed exercise program is required to support an effective rehabilitation program. A case study on a soccer player with a Grade 2 ankle sprain has shown that it was possible to achieve a pain-free and short recovery from running (Kern-Steiner et al., 1999). The program was designed with specific sets and repetitions of exercises, but the crucial

element was the lab setup's ability to support a gradual gravitational increase of the weight bearing. Rehabilitation programs consider the individual's ability to recover, and a similar strategy could benefit patients with MSK conditions.

The rest of this project follows the definitions per (Caspersen et al., 1985), which defines physical activity as “any bodily movement produced by skeletal muscles that result in energy expenditure above resting levels”. This definition broadly encompasses exercise, sports and physical activities as part of daily living, occupation, leisure and active transportation. Exercise is a subcategory of physical activity “that is planned, structured and repetitive and has, as a final or intermediate objective, the improvement or maintenance of one or more dimensions of physical fitness”.

2.2.2 Physical activity and bones

The loss of bone strength caused by Osteoporosis is associated with the loss of minerals and collagen caused by ageing and inactivity. Bones adapt their density based on an individual´s environment, including physiological factors (e.g. diet) and mechanical constraints. The principle of bone adaptation is known as Wolff´s law which has been extensively reviewed through experimental and observational studies of bone changes (Ruff et al., 2006). Time in space is an example of an extreme environment causing bone loss, as shown in a study on astronauts. The space environment does not have gravitational forces, which induces a lack of mechanical load on the bones and presents a risk of developing osteoporosis (Stavnichuk et al., 2020). It has also been found that bone density might not be fully recovered after a year back from space (Gabel et al., 2022).

Bone density and remodelling can be improved as a response to the mechanical caused by exercise, but it is essential to distinguish the different types of exercises (O’Brien, 2001). Although there is not enough evidence to quantify the risks of falls and fractures due to activities, an international panel agreed in 2014 that an exercise program should include multiple components and avoid aerobic training, except for resistance or balance training (Giangregorio et al., 2014). The change in mineral content is an adaptation to force and stress caused. Hence, an effective exercise program needs to consider the type of activity, frequency, intensity and duration to cause a significant change in physical activity (Dalsky, 1987). A literature review found that resistance and weight-bearing aerobic exercises can stimulate bone osteogenesis. However, strength exercises are effective only if they cause a joint reaction greater than daily activities. Aerobic exercises are helpful only if performed with an intensity causing a significant ground reaction force (Benedetti et al., 2018). Any exercise applies an amount of mechanical loading, and besides the intensity, a study has found correlations between the frequency of load applied and

bone density growth. The bone tissues respond to dynamic loading, and higher frequencies stimulate osteogenesis, but extended exercise can be detrimental to the bones. Recent studies have observed better results in bone building when shortening the amount of time at load in a session than reducing the number of sessions (Turner and Robling, 2003).

2.2.3 Physical activity and joints

Thus, increasing load is good for bone density, but it is essential also to consider the negative impact that might incur because of excessive physical activity. In 2016, a consensus by the international Olympic committee reached a statement that the recommendations in the amount of training need to follow a methodology to prevent the risk of injury in athletes. Such recommendations consider medical results (e.g. blood lactate concentration) as well as the frequency and intensity of training (Soligard et al., 2016). Indeed, a retrospective study on endurance runners has found an increase in the risk of injury depending on the foot strike pattern and position (Daoud et al., 2012). Besides endurance, a review of the hospital admissions of OA between 1970 and 1990 has shown that Olympic male athletes (representing Finland) presented risks of OA for high frequency (i.e. endurance) sports and both high intensity (i.e. power sports) (Kujala et al., 1994). A study on a similar cohort of female ex-elite athletes found that the excessive loading in the joint caused by weight-bearing sports activities increases the risks of developing OA of the knees and hips (Spector et al., 1996). According to a review of data on degenerative joint disease done in 1994, the stress caused by excessive loading on joints might speed up the development of OA, and physios still need more information to be able to build effective programs (Panush and Lane, 1994).

The physical strain associated with sports and exercises also applies to other occupations besides athletes. A study on 1566 US army soldiers found OA to be the most common disability for US soldiers unable to return to duty (Rivera et al., 2012). A military report has defined tactical athletes as individuals with similar occupational activities and a higher fitness load, such as the military, law enforcement and rescue services (Scofield and Kardouni, 2015). Although further research is needed to identify all the factors, a systematic review of the literature has found an increased risk of developing OA in these occupations (Cameron et al., 2016). A more recent review has found the risk of developing OA to be more generally associated with any occupations requiring to perform movements with heavy physical workloads on the lower limbs, such as squats or heavy lifting (Schram et al., 2020). Biomechanical components cause OA, as shown in a study on 228 veterans, which also identified misalignment of the lower limbs and being overweight as risks (Felson et al., 2004). The effects of exercises and occupational activities require further studies, and in 2017, a consensus study was conducted to harmonise the

classification methods. The group included OA and PA international professionals that agreed to use MET-min per week for studies measuring PA. The key recommendations also raised the need to measure the intensity and duration of the joint load (Gates et al., 2017), which at the time of this study are yet to be standardised.

2.3 Physical Activity & load forces

The guidelines and fitness recommendations to stay healthy are associated with aerobic (cardiorespiratory) and anaerobic (intensity) exercise programs which can be measured in a lab or fitness environment, e.g. using equipment such as treadmills, bicycles and elliptic. However, the mechanical load on the joints can affect MSK disease progression, as illustrated in fig 3. Underloading might limit induced loss of bone density leading to osteoporosis, while overloading can affect the joint structure leading to OA. The load intensity and frequency associated with healthy joints are not characterised and thus should be measured.

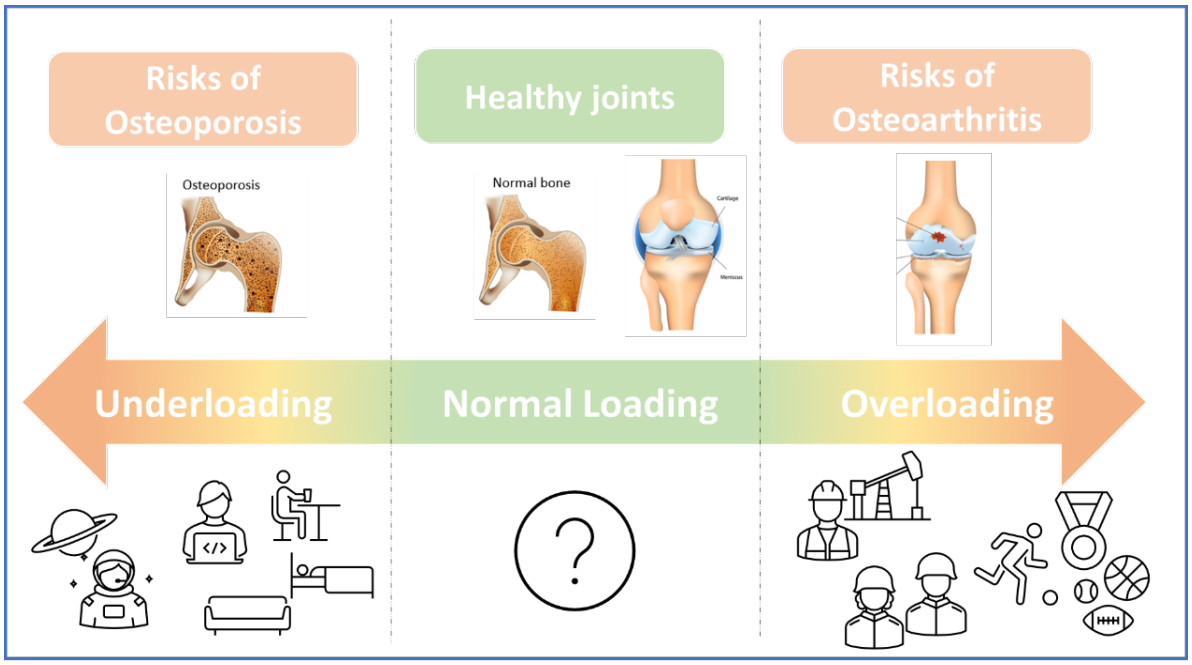


Figure 3: Under Vs Overloading

MSK studies typically focus on the medical impact of the mechanical load, e.g. before and after load changes or in the context of recovery. Typically, osteoporosis studies measure bone composition. OA and RA studies look at the joints' alignment and cartilage composition. Mechanical load´s effects have been reviewed through in vitro and in vivo studies (Griffin and Guilak, 2005). In vivo studies have been conducted on animals using implants, and implant modelling consider the biomechanical environment as suggested by a review of Wolff's law (Prendergast and Huiskes, 1995). The dynamic loading is measured as strains and can be

expressed mathematically per equation derived from the Fourier method (Turner, 1998). The strain stimulus E, measured in microstrain (με), is proportional to the characteristics of the load where k is a proportionality constant, ε is the strain´s peak-to-peak magnitude and f is the loading frequency. It can be noted that for a static load, represented with a frequency $f = 0$, $E = 0$ so no bone adaptation is induced because of strain.

$$E = k_1 \sum_{i=1}^{n} (\varepsilon_i f_i)$$

It is impractical to set up in vivo implants on humans outside a diagnosed clinical need. So in vitro methods are also used, e.g. using cartilage explant but do not provide an accurate representation of the impact on humans. Therefore, epidemiological studies are the most practical, and multiple protocols have been set to manipulate the load bearing. Astronauts have tested water immersion, parabolic aircraft flights, supine and erect cable suspension, and centrifugal methods to simulate gravity-free conditions as experienced in long space flights (Davis and Cavanagh, 1993). Other custom setups can also be put in place, allowing to adjust of the weight bearing, e.g. using an underwater treadmill with a harness in another study with astronauts (Newman et al., 1994) or a lab setup to allow gravitational alterations for the rehabilitation of a soccer player (Kell et al., 2001). A more recent study tested using mice suspended with a spring with tension adjusted between 10% and 80% of the mice´s body weight (Wagner et al., 2010).

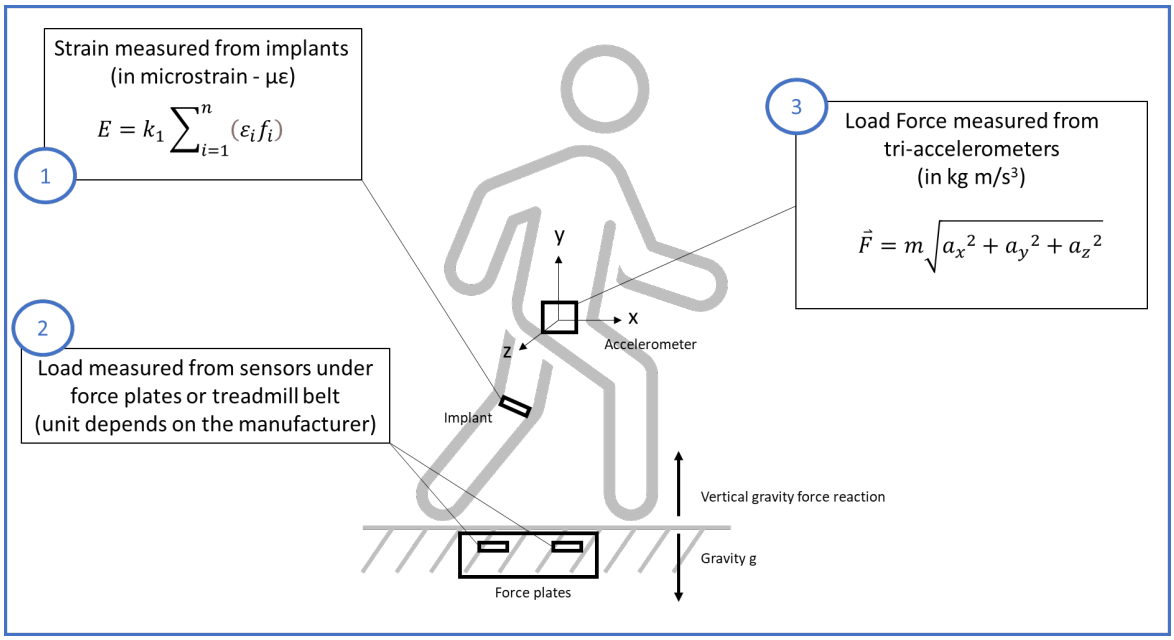


Figure 4: Measure of mechanical loading (1) via implant; (2) force plate; (3) accelerometer

The force on the lower limbs is caused by the bearing of the person´s weight reacting with the ground due to gravity, as shown in fig 4. A literature review has been conducted to identify the

main methods to measure weight bearing while standing and walking (Hurkmans et al., 2003) and found clinical examination using scales as the most common evaluation, although limited to standing position. Biofeedback systems, ambulatory devices and the use of platforms are also common. Force plates are platforms that can be used to measure the vertical ground reaction forces of the body weight in a static movement, as shown in a study with patients affected with paraplegia (paralysis of the legs) (Bernhardt et al., 2012). Treadmills can be combined with video cameras to evaluate dynamic movement, as used in a study comparing rearfoot with forefoot strikes to assess the rate of a stress injury in runners (Daoud et al., 2012a). Anti-gravity treadmills are available in the specialised market and can be used in rehabilitation and training to evaluate the effects of progressive loading (AlterG, 2012). A preliminary study has shown that the load caused in the ground reaction can be measured in dynamic loading (e.g. walking, running) from the pressure applied on treadmills with strain gauge force sensors fitted which highlighted the dependencies on the distance between the sensor and the centre of pressure (Dierick et al., 2004). However, this measuring method requires specialised equipment and is typically conducted in biomechanical laboratories, which cannot easily be replicated for day-to-day tracking.

According to the general principles of kinetic and Newton´s laws (Elvan and Ozyurek, 2020), the gravity applied to the body reacts, producing an equal ground force reaction in the opposite direction (3rd law). Unless external forces are applied, the body maintains a constant velocity (2nd law). The load force ($\vec{F}$) generated is proportional to the weight and velocity (3rd law), which can be mathematically expressed in proportion to the weight (m) and acceleration perceived on the body ($\vec{a}$).

$$\vec{F} = m\vec{a}$$

$$\vec{F} = m\sqrt{{a_x}^2 + {a_y}^2 + {a_z}^2}$$

Although the sensor setup and unit is different, studies have also compared the performance of accelerometers to force plates in static and dynamic movement. A study asking participants to stand has found the measurement from tri-axial accelerometers worn in the back to perform as well as the force platform ($P<0.05$) (Mayagoitia et al., 2002). A similar correlation was found with participants who were asked to perform different movements (walking, running and dropping a box) on a force platform while wearing the accelerometers at the wrist and the hip (Rowlands and Stiles, 2012). Accelerometry studies have considered the location of the sensors on the human body (Bouten et al., 1997) and concluded that the use of accelerometer sensors presents a viable method for long-term monitoring of the ambulatory movement of the human body (Mathie et al.,

2004). The reliability of accelerometers to measure the effects of loading and unloading generated by physical activity has been confirmed for squats movement (Bobbert, 2014) as well as with a football (Boyd et al., 2011) and a rugby team (McLean et al., 2018). Accelerometers have also been used to evaluate physical activity in musculoskeletal studies but outside the clinical environment, concluding that RA patients are more sedentary than control participants (Prioreschi et al., 2013). That correlation exists with their disease activity (Hernandez-Hernandez et al., 2014).

2.4 Smartphones Applications & remote monitoring

A review of the techniques to measure weight bearing was conducted in 2003 (Hurkmans et al., 2003), but since then, accelerometers have been embedded in devices such as smartphones and wearables. They have become more broadly available to the general public. The usage of a smartphone to record physical activity has been verified with comparative devices such as Actigraph (Eric B Hekler, 2015) as well as to send readings from sensors worn on the body (Seeger et al., 2014) and for physical activity recognition (Wanmin Wu, 2012). Moreover, systematic reviews concluded that wearables and smartphone applications could lead to an increase in physical activity (Gal et al., 2018). However, evaluation guidelines are needed to optimise the research design (McCallum et al., 2018). Smartphones are already widely available to the general population and present the advantages of being convenient, low-cost, easy to deploy, and already broadly used and part of everyone's day-to-day life (Woollaston, 2014). Business models exist to generate revenues either by charging the users directly when accessing the app, premiums, gated features and in-app purchases (Apple, 2019); or indirectly, e.g. through sponsorship, advertisement or data reselling (Mey, 2017). The commercial success can be measured with global ranking found on specialised websites such as SimilarWeb (2019).

A wide range of fitness and workout apps is available and continuously updated, including apps from leading sports brands such as Nike+ Training Club, Adidas MiCoach and Freeletics, focusing on exercise programs, guidelines and videos (Haslam, 2018). Tracker apps such as Nike+ Club and Strava track time and distances, focusing on GPS data to track geographic location rather than movement (Runner'sworld, 2018). Sensor trackers and analysers also exist but focus on sensor data visualisation rather than tracking and recording. A review of 60 studies has shown that being used within medicine is perceived as promising and exciting by clinicians (Ozdalga et al., 2012). The RA population share the same interest and a questionnaire answered by 100 RA patients showed that 94% believe that they could have a more active role in self-management and that an app explicitly developed would be helpful (Azevedo et al., 2015). The benefits of this approach have been demonstrated with an app used as part of an integrated platform (Dixon and Michaud,

2018). As seen in Table 1, the top apps made available (Veronica Hackethal, 2018) to the OA and RA population focus on providing information on the disease rather than physical activity tracking features.

Smartphone application	Feature	Type
Epocrates	Pharma and medical reference	General Health/Medicine
DynaMed	Medical and pharmacologic reference	General Health/Medicine
Johns Hopkins' Antibiotics Guide	Antimicrobial reference	General Health/Medicine
Sanford Guide to Antimicrobial Therapy	Antimicrobial reference	General Health/Medicine
Diagnosaurus	Differential diagnosis	General Health/Medicine
Taber's Medical Dictionary, Stedman's Medical Dictionary, and Dorland's Medical Dictionary	Medical dictionaries and reference	General Health/Medicine
Archimedes (Archimedes 360°)	Medical calculator	General Health/Medicine
AHRQ ePSS®	Primary care prevention	General Health/Medicine
Medscape	Medical reference, news, and education	General Health/Medicine
Massachusetts General Hospital's Pocket Medicine	Medical reference	General Health/Medicine
Washington Manual of Medical Therapeutics	Medical reference	General Health/Medicine
QuantiaMD	Medical education with interactive cases	General Health/Medicine
MedPage Today	Medical news	General Health/Medicine
Doximity	Social networking for physicians	General Health/Medicine
HealthLog	Diray to track all health measure (weight, blood pressure, workouts, sleep, diet, mood, and checkup history, nutrition, fitness, and sleep over time) through graphs and charts.	General Health/Medicine
RxmindMe	Diary and reminder - Medication	General Health/Medicine
MyMeds	Diary - Medication	General Health/Medicine
MyMedSchedule	Diary and reminder - Medication	General Health/Medicine
Mediquations	Information - calculators	General Health/Medicine
Lab Gear	Information - calculators and common symptoms info	General Health/Medicine
READY	Diary - clinician tracking of patient data	RA specific
Rheumatoid Arthritis Diagnosis and Management	Information - calculators and common symptoms info	RA specific
RheumaTrack	Diary - Track joints pain and stiffness, activities and send the data to your doctor. Includes medicine reminders.	RA specific
MyRA	Diary - symptoms	RA specific
My Pain Diary	Diary - pain	RA specific
Track + React	Diary - Analysis of symptoms and activity based on user input information (symptoms, nutrition habits, sleep, medication, fitness)	RA specific
RAVE	Diary - medication	RA specific
Hand: Rheumatology Checklist	Information - medication	RA specific
RheumaHelper	Information - calculators and arthritis info	RA specific
Body Pain Management Hypnosis	Pain management techniques (hyphosis)	RA specific
Rheumatoid Arthritis (RA) @Point of Caretm Edition	Information - Diagnosis, treatment	RA specific

Table 1: Smartphone app feature review

The amount of physical activity performed is personal and depends on the conditions of each individual thus, organised exercise programs might be helpful. However, the recommendations made by healthcare professionals remain based on an evaluation of the physical activity reported through patient feedback. A self-assessment of physical activity is not quantified or objective, as shown in a literature review (Prince et al., 2008). A review has shown that in most studies, self-

reports do not accurately estimate the absolute amount of physical activity (Sallis and Saelens, 2000). Assessing physical activity requires an objective sense of measures, and the findings on using paper diaries suggest that electronic journals with compliance-enhancing features would be more effective (Stone et al., 2003). Indeed, the results comparing self-assessments to data recorded through Actigraph have shown a weak correlation (Dyrstad et al., 2013) and a review demonstrated that there are not enough studies (Füzéki et al., 2017) on low activity recorded through accelerometers. A solution providing convenient physical activity monitoring using objective data will help better advise patients on the most appropriate program of activities to follow.

The overall physical activity hinges on an individual´s daily routine, and the perception often overlooks the living conditions (e.g., living in a two-story house rather than the ground floor) and daily chores such as housework or children's care. All these apps require user intervention (i.e. to start/stop the app) and run in the foreground (i.e., with the app always displayed), which is a barrier to continuous monitoring. Similarly, wearable such as smartwatches and insole are available but not as broadly adopted as smartphone and less likely to be used by individuals over 50 (Chandrasekaran et al., 2020). The nature of this project is to gain insight into the day-to-day level of physical activity, implying that the monitoring should be non-intrusive, seamless and wholly integrated into the user's lifestyle. The users should not be burden with smart wearable nor restricted to holding the phone in specific positions, and any app running should have minimum battery performance and storage impact. All these apps require user intervention (i.e. to start/stop the app) and run in the foreground (i.e. with the app always displayed), which is a barrier to continuous monitoring.

Bespoke apps have been developed to support specific studies. A study has developed a custom app running on LG smartphones (Android OS) to use a variety of smartphone sensors to develop a human activity recognition system able to recognize 15 activities with high accuracy (Khan et al., 2014). A study has demonstrated that data recording apps can be developed as shown with MyHealthAssistant, which retrieves data from sensors worn close to the body (Seeger et al., 2015). The performance review shows that an app can be designed to run background data processing that does not require user intervention (e.g., to start and stop the logging). Such background running apps can be developed without inducing bias in the data recorded, as shown in a systematic review of the effect of a digital intervention on the physical activity of people with inflammatory arthritis (Griffiths et al., 2018).

2.5 Project Hypothesis

In summary, the force load on the lower limbs resulting from physical activity plays a critical role in the well-being and evolution of musculoskeletal diseases. Too little might lead to osteoporosis and too much could lead to osteoarthritis. Therefore, the recommendations should be tailored to fit the patient's needs and daily lifestyle. However, the amount of physical activity is currently assessed based on patients reporting, typically biased as being self-reported, and mechanical load is measured only in dedicated facilities or via implant. Fitness trackers allow continuous monitoring of sensor movement but the commercially available solutions do not focus on people with limited movement capability and the stress and load caused on the bones and joints is not monitored nor quantified outside of clinical setting.

The hypothesis of this project is that smartphones can be used to assess the amount of physical activity performed and provide objective and valuable insight for patients diagnosed with musculoskeletal diseases. The novelty is using smartphones, as affordable access to the technology for continuous tracking, to continuously estimate the load rate on lower limbs in patients' daily routine and natural environment.

The primary objective is to validate that smartphones can be used as surrogates to estimate lower limb loads. The secondary objective is to validate the ability to continuously monitor musculoskeletal patients' physical activity outside the clinical setting and for an extended timeframe. An additional outcome is to evaluate the viability of the method for further healthcare studies to correlate load rate and physical activity, e.g., with disease activity.

Chapter 3 General Methodology

3.1 Study design

The project aims to evaluate the physical activity estimates obtained from smartphones considering the disease condition of patients recruited within the NHS. Therefore, the protocol has been approached as medical research (Kapoor, 2016), however, without intentions to experiment or treat patients and with no causality to the disease condition and activity expected as a result of taking part for the participants. The focus is to collect prospective data (load rate estimate), so this project is defined as a pilot (Leon et al., 2011) epidemiological observational study.

The objective is to validate the developed methodology and framework, which could be applied to full-scale clinical trials and experimental studies. The studies have been designed to evaluate the feasibility of the process's critical steps, the project stakeholders' management (clinical and patient), and the time and resources required for future clinical studies. A PPI session has been conducted with rheumatoid arthritis patients to validate the benefits of the study capturing actual patient feedback. Clinicians' and patients' points of view have consistently been tracked, but this is not a clinical study nor trial, so there is no need for a control group nor to randomise the cohort of participants.

The key details for each study are summarised in table 2. Study 1 (Chapter 4) served as the first end-to-end proof of concept (POC) to identify the key challenges. Using a methodical triangulation (Bekhet and Zauszniewski, 2013), both quantitative and qualitative data were collected respectively from smart devices (Smartphone and Fitbit) and self-assessed by recruited participants diagnosed with rheumatoid arthritis (questionnaire).

The consequent activities have then been defined to validate critical aspects of using smartphones to estimate load rates. Study 2 (Chapter 5) compares the load rate estimates from smartphones and smartwatches with uniform data collection (i.e. smartphone power save turned off). Study 3 (Chapter 6) compares the variations of real-life sampling, in clinical trial conditions with osteoarthritis participants, with the physical activity recorded using Fitbit. Study 4 (Chapter 7) evaluates the trends of physical and disease activity that can be recorded, in real-life conditions, from the personal smartphones (as opposed to secondary device) of participants diagnosed with rheumatoid arthritis.

Study	Participants	Devices	Sample	Data collected
PPI Session	9 patients diagnosed with Rheumatoid Arthritis	Sony Xperia Z1 Compact	1 day	Survey before session Feedback during session
Study 1	5 patients diagnosed with Rheumatoid Arthritis	Fitbit HR Sony Xperia Z1 Compact preloaded with RApp v1.0 (window=10s, interval=20s)	6 weeks	Fitbit Daily step count summary Raw accelerometer data stored as .txt files Processed Load rate mean stored in SQL database
Study 2	10 participants with no specific conditions	Sony Smartwatch 3 Sony Xperia Z5 Compact preloaded with OApp v1.0 (window=5s, interval=55s)	7 days	Smartwatch & Smartphone Raw accelerometer data stored as .txt files Processed Load rate mean in SQL database
Study 3	60 patients diagnosed with Osteoarthritis	Fitbit Flex 2 Samsung A5 preloaded with OApp v2.0 (window=5s, interval=15s)	14 days (7+7, 2 months after injection)	Fitbit Daily step count summary Processed Load rate mean stored in SQL database Google API step count
Study 4	3 patients diagnosed with Rheumatoid Arthritis	RApp v2.0 installed from Google Play (window=5s, interval=15s)	3 Months	Processed Load rate mean stored in SQL database

Table 2: Study summary

3.2 Ethical approval & Participants

Although this project is not a clinical study, NHS approval and additional documents, including the patient information sheet and patient consent form (Study 1 and 4), have been reviewed, along with the study protocol, and approved by London Stanmore's ethics committee (REC reference: 16/LO/0182; IRAS ID: 192803). The Faculty of Engineering and Environment Ethics committee at the University of Southampton has approved studies 1 and 4 (ERGO ID: 18061). Study 2 uses datasets generated as part of a project (Nazirizadeh, 2018) approved by the University of Southampton's Faculty of Engineering and Environment Ethics Committee (no.30213, see Appendix B). The Northern Sydney Local Health District HREC has approved study 3 (Reference number: LNR/17/HAWKE/370). No ethical approval is required for activities solely focused on app development and testing.

Several variables and formulas are used in clinical studies to accurately define the minimum sample size of participants required (Charan and Biswas, 2013). However, this project gathers prospective data, so the number of participants recruited in each study has been defined to focus on the particular point of interest evaluated. The participants in the PPI session were adults diagnosed with rheumatoid arthritis, recruited from an existing research group for a biotherapy project. Inclusion criteria were defined for studies 1 and 4 to recruit participants aged above 18 years old, diagnosed with rheumatoid arthritis and willing to take part in the study for the entire duration (6 weeks for study 1 and 3 months for study 4). No aspects of disease activity were of

interest for study 2. So, inclusion criteria were defined to recruit healthy participants who were willing to participate in the study, who were above 18 years old and who were without neurological, systemic illnesses or other physical disabilities that may have limited their mobility. The cohort for study 3 were patients over 50 with symptomatic, radiographic knee osteoarthritis recruited as part of clinical research before and after knee injection (Yu et al., 2022). No exclusion criteria were specifically included but the scope of study includes the use of smartphones which is naturally not appealing to individuals with negative perception on technology.

RA patients were recruited from the rheumatoid arthritis population treated at Southampton General Hospital, and patients were screened during weekly clinics at the hospital. Recruiting patients as part of a research study requires GCP (good clinical practice) training at Southampton General Hospital. Access to patient notes (i.e., NHS paper-based records) requires obtaining a "Research passport" (see NHRA docs) to track the consent date, follow up and end of participation in the study. Patient participation is also recorded electronically through EDGE (Edge, 2018). All face-to-face interactions with patients (consent, follow up and end of study) are done within the Clinical Research Facility of the Southampton General Hospital in the presence of at least one nurse.

3.3 Equipment

Smartphones

Smartphones are used for their convenience (low-cost and broadly used) of access to accelerometer sensors. Android offers the most extensive range of devices and has the broadest portion of users, which makes it easier to obtain the technology. Therefore, iPhones have intentionally not been included in this project. Android-based devices that include Samsung S7 Edge, HTC M8, and HTC M8 mini have been used for development purposes, and Sony Xperia Z1 Compact, Sony Xperia Z5 Compact and Samsung A5 have been provided to recruited participants (Study 1, 3 and 4).

Smartwatch

Android also provides a range of wearables that includes smartwatches. In this project, we used the Sony Smartwatches 3 (Study 2) to record accelerometer data using Google Android Wear 2.0 API.

Fitbit

Fitbit devices are wristbands that measure daily step counts and physical activity. Considering the recent commercial grasp it has amongst the broader population, these were used as a reference and electronic, physical activity baseline. Fitbit uses proprietary algorithms that are not publicly available, and the review of their performance is not part of this project. In the context of this study, we accept these as being a reliable measure of performance already validated in other studies ((Adam Noah et al., 2013) ; (Diaz et al., 2015); (Dontje et al., 2015); (Takacs et al., 2014).

Fitbit HR (Fitbit, 2016) were used for Study 1 and Fitbit Flex 2 in Study 3. As of the beginning of this study, Fitbit provided access only to processed data (e.g., step count) but opened their APIs on 2018 models (Versa and Ionic). It should be noted that while the use of Fitbit's APIs is not part of the scope of this project, raw Fitbit accelerometer data could be considered for future work and could bring commercial value, e.g., in the context of comparative data analysis paired with data recorded via smartphones.

3.4 Smartphone considerations

Data collected using commercially available devices (Fitbit and smartwatch) are used as a benchmark, and disease activity data are obtained directly from patients using self-assessed questionnaires.

The primary focus of work is on evaluating the physical activity estimated using smartphone accelerometer data. So, designing a specific app provides a better user experience that increases the retention of the app. This approach also dramatically simplifies the ability to obtain ethics approval and expand the possibilities for further development, e.g. for commercial purposes by adhering more strictly to the UK regulation and guidelines on medical devices that have been extended to the medical apps (Gov.uk, 2018b). This section details the implementation of the monitoring framework resulting in two smartphone apps that are Rapp™ and Oapp™.

3.4.1 Smartphone architecture & sensors

The architecture of the smartphone combines hardware and software components at various levels (see fig 5). As well as the screen and overall casing, the hardware consists of integrated sensors that are used for multiple functions, such as adjusting the operating conditions and user experience (Abiresearch, 2019), and the trend shows that more sensors are included to expand the use case to fitness, gaming or security (Qualcomm, 2014). Chipsets manufacturers provide their proprietary software stack (middleware) to manage the multiple technologies supported, such as the radio protocol (e.g. Qualcomm or Apple), the Wi-Fi (e.g. Texas Instruments) and Bluetooth.

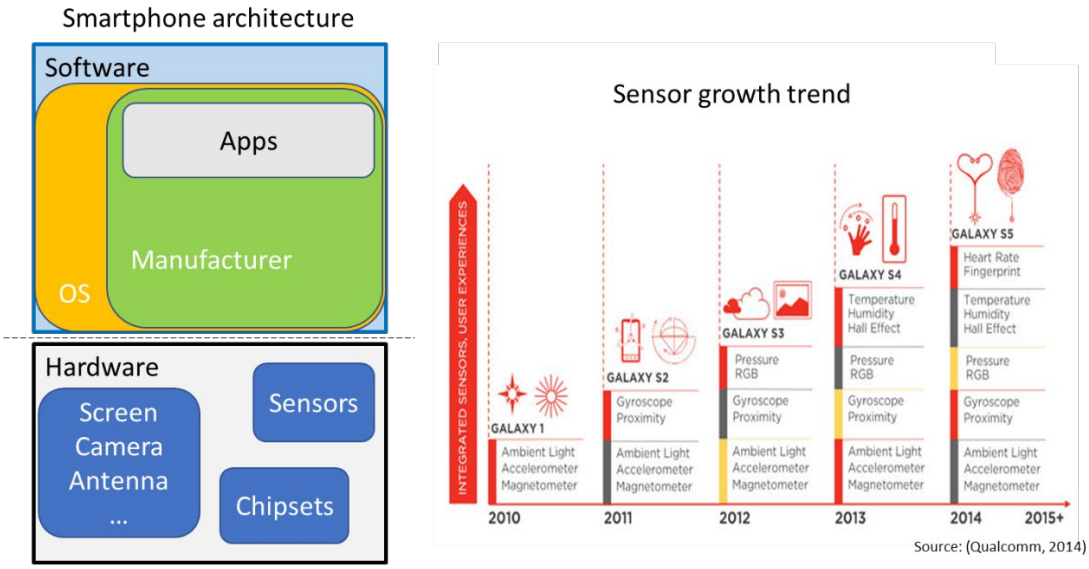


Figure 5: Smartphone architecture and sensor growth trend

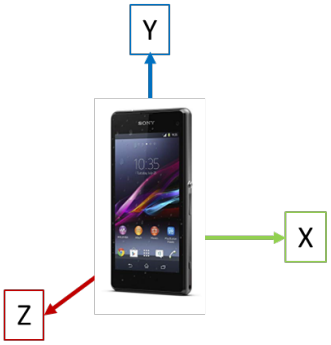
The overall user experience and front end are based on the OS used (e.g., Android, Apple, Blackberry) and manufacturers (e.g. HTC, Sony, Samsung...) design and build their devices based on the hardware components supported, e.g. on Android (Google, 2018a); Apple combines the requirement and support of iOS along the iPhone specifications (Apple, 2018). Smartphone manufacturers typically add a proprietary software layer to customise further the user experience, such as “HTC Sense” or “Samsung Experience”, which must comply with the OS provider regulations (e.g., obtain Google Technical Acceptance). A smartphone app is specifically developed to be compatible with the targeted OS that can access other software (i.e. other apps) and hardware components but is limited to the API capabilities. For this project, the smartphone apps were designed only for Android, but the principles apply to any other OS (e.g. Apple, Blackberry, and Windows).

Sensor data

Google’s Android platform provides a Hardware Abstraction Layer (HAL) representing the sensors as virtual devices that can be accessed programmatically via API. As detailed in figure 6, *Base* sensors are a single type of physical sensor, such as the accelerometer, gyroscope and magnetometer; and *Composite* sensors are processed data, e.g. for activity (e.g. step count), attitude (e.g. rotation vector) or interaction (e.g. wake up gesture) (Google, 2018b).

In this thesis, we focus on the accelerometer sensors to calculate the load rate and refer to these as “raw”, but are readings calibrated using temperature compensation, online bias calibration and online scale calibration, not raw output from the physical sensors. As per fig 6, the format of the readings includes both the gravity and rate of change of velocity along the three sensor axes, represented as time (t) and coordinates (x, y and z). Composite sensors combine sensors and

processed data, but Android considers each sensor independently. The step count reading has been later added (see study 3) and does not alter the overall logging (Google, 2016).



Description	Unit	Sensor event data	Description
TYPE_ACCELEROMETER	m/s²	SensorEvent.values[0]	Acceleration force along the x axis (including gravity).
		SensorEvent.values[1]	Acceleration force along the y axis (including gravity).
		SensorEvent.values[2]	Acceleration force along the z axis (including gravity).
TYPE_GRAVITY	m/s²	SensorEvent.values[0]	Force of gravity along the x axis.
		SensorEvent.values[1]	Force of gravity along the y axis.
		SensorEvent.values[2]	Force of gravity along the z axis.
TYPE_GYROSCOPE	rad/s	SensorEvent.values[0]	Rate of rotation around the x axis.
		SensorEvent.values[1]	Rate of rotation around the y axis.
		SensorEvent.values[2]	Rate of rotation around the z axis.
TYPE_GYROSCOPE_UNCALIBRATED	rad/s	SensorEvent.values[0]	Rate of rotation (without drift compensation) around the x axis.
		SensorEvent.values[1]	Rate of rotation (without drift compensation) around the y axis.
		SensorEvent.values[2]	Rate of rotation (without drift compensation) around the z axis.
		SensorEvent.values[3]	Estimated drift around the x axis.
		SensorEvent.values[4]	Estimated drift around the y axis.
		SensorEvent.values[5]	Estimated drift around the z axis.
TYPE_LINEAR_ACCELERATION	m/s²	SensorEvent.values[0]	Acceleration force along the x axis (excluding gravity).
		SensorEvent.values[1]	Acceleration force along the y axis (excluding gravity).
		SensorEvent.values[2]	Acceleration force along the z axis (excluding gravity).
TYPE_ROTATION_VECTOR	Unitless	SensorEvent.values[0]	Rotation vector component along the x axis (x * sin(θ/2)).
		SensorEvent.values[1]	Rotation vector component along the y axis (y * sin(θ/2)).
		SensorEvent.values[2]	Rotation vector component along the z axis (z * sin(θ/2)).
		SensorEvent.values[3]	Scalar component of the rotation vector ((cos(θ/2)).[1]
TYPE_STEP_COUNTER	Steps	SensorEvent.values[0]	Number of steps taken by the user since the last reboot while the sensor was activated.

Source: Google, 2018b

Figure 6: Android sensors (Google, 2016)

Gyroscope data could be helpful for smartphone positioning but not used in this project, as accelerometer data provide the same insight. GPS data could also be helpful, e.g. to estimate sunlight exposure or to track the distances and routine. However, GPS is not a sensor and uses a different reporting mechanism (like the camera, touchscreen, and fingerprint). GPS requires satellite coverage which limits the capability to gather data indoors and raise practical issues with low mobility and sedentary users. Besides, the tracking of users 'location might raise ethics and privacy concerns, so GPS data was not used, and the project focused on estimating mechanical load from accelerometers data.

3.4.2 Software development lifecycle

The research team continuously measures the results and performance of the apps at various stages of the development, which allows verification of the correct implementation of features and fixes used for extended period, unlike studies limited to specific timeframes and limited by protocol. The traditional Waterfall methodology follows sequential phases from requirements to design, implementation, verification, and maintenance. Considering that this project is a pilot and that various requirements might change, it has been decided to follow an Agile methodology which permits flexibility in code changes and be reactive to issues found.

For research purpose, the code does not need to be optimised but follows an Agile methodology using Sprints. Each sprint defines the functionalities and blocks of code required, which are then tested and validated (or not) as part of each study.

As shown in figure 7, we use the main branch that includes all the code changes; the feature branch is to work on specific features that might include bugs or take longer to verify; the release

branches are the code version deployed to users, i.e. used for the study. Git is a code repository that allows functionality, development and to revert through code branches (see Appendix L).

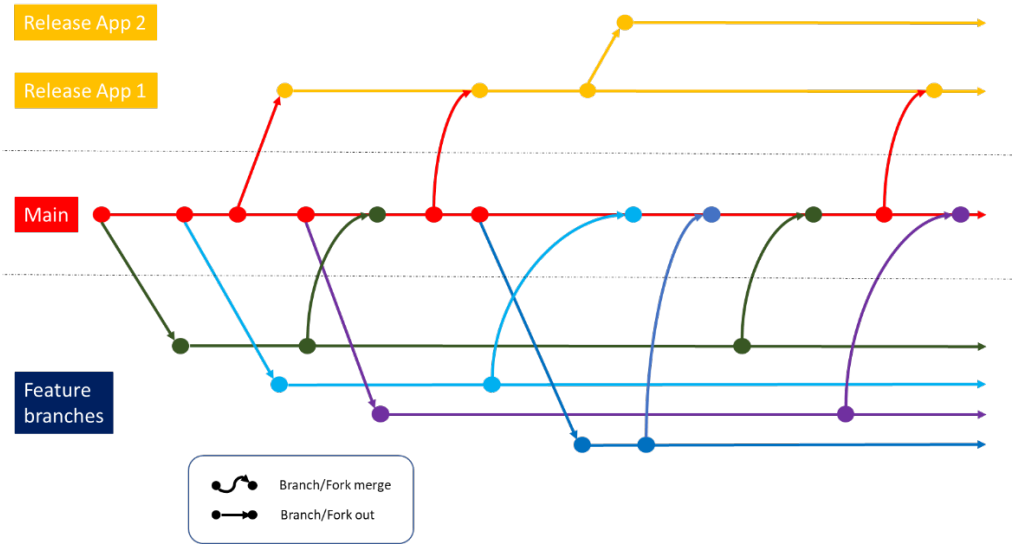


Figure 7: Branching strategy

The version of code changes is tracked with two digits (e.g., 1.5) indicating major feature release with the first digit and less significant changes or bug fixes with the second digit. Rapp™ went through multiple versions to develop the user interface for disease activity score (DAS), questionnaires and the first implementation of continuous monitoring. The first version (i.e. Rapp™ v1.0) has been used for the patient and public involvement (PPI) session and improved in study 1. Rapp™ v1.0 has then been forked out to an Oapp™ branch where modifications have been made to the user interface, including fixes to the continuous monitoring and add support for Android Wear (i.e. smartwatch); resulting in Oapp™ v1.0 which has been used for study 2. Following the observations with both variants (study 1 & 2), monitoring of step count has been added along layout optimisations for Oapp™ v2.0 (study 3) and Rapp™ v2 (study 4).

App coding environment

Throughout this project, an app refers to a program that developed for smartphones. Advanced integrated development environment (IDE) such as Xamarin allow building mobile apps across platforms (i.e., Android, Apple and Windows) using C# as coding language. However, these solutions rely heavily on APIs and consequently dismissed to ensure more control over the sensors and data layer used.

Android is a Linux based Operating System that considers each app as different users. Android apps are coded in Object-Oriented Language (e.g., Java, C++) and compiled through Android SDK into Android packages (APK) that can be installed on Android devices. Any integrated development environment (IDE) can be used, and Eclipse was initially used but migrated to

Android Development Tools (ADT), as seen on figure 8, to benefit from in-depth features such as debugging tools, code version tracking and integration with GitHub.

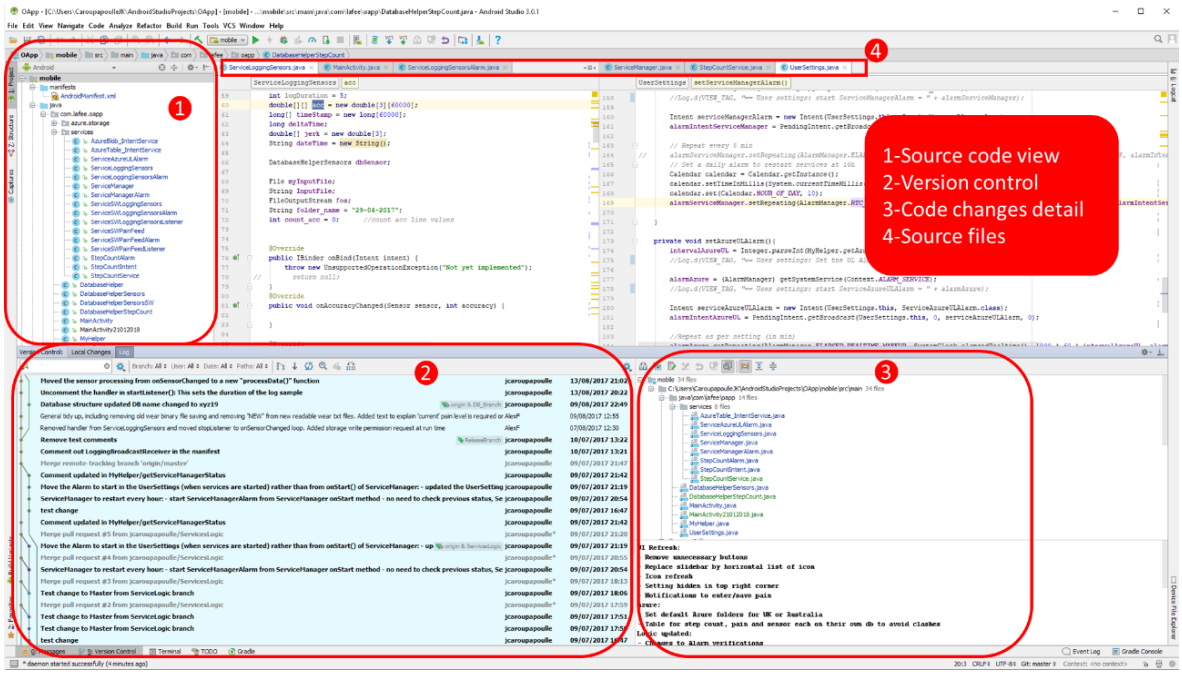


Figure 8: Screenshot of Android Development Tool (ADT)

Apple follows a similar approach but with critical differences such as coding language being Objective-C based with the possibility to use Apple's coding language (Swift), and the IDE (Xcode) runs exclusively on Apple OS devices (e.g. iMac, MacBook) to be able to compile the code loaded onto iOS devices (i.e., for iPhone, iPad...).

App architecture

Smartphone apps are typically implemented in blocks of code that define specific behaviour and entry points of the app for the user or the system. Figure 9 details the components used for this study. The Android API defines four fundamental types of components that each have their lifecycle and own use case. *Activities* usually are the visual components run in the foreground such as the main view or any of the child views that the user would navigate to; *Services* are used for tasks running in the background that do not require user interventions; *Broadcast receivers* allow interaction with the app by the system, outside of the regular flow triggered by the user; and *Content providers* manage the file system.

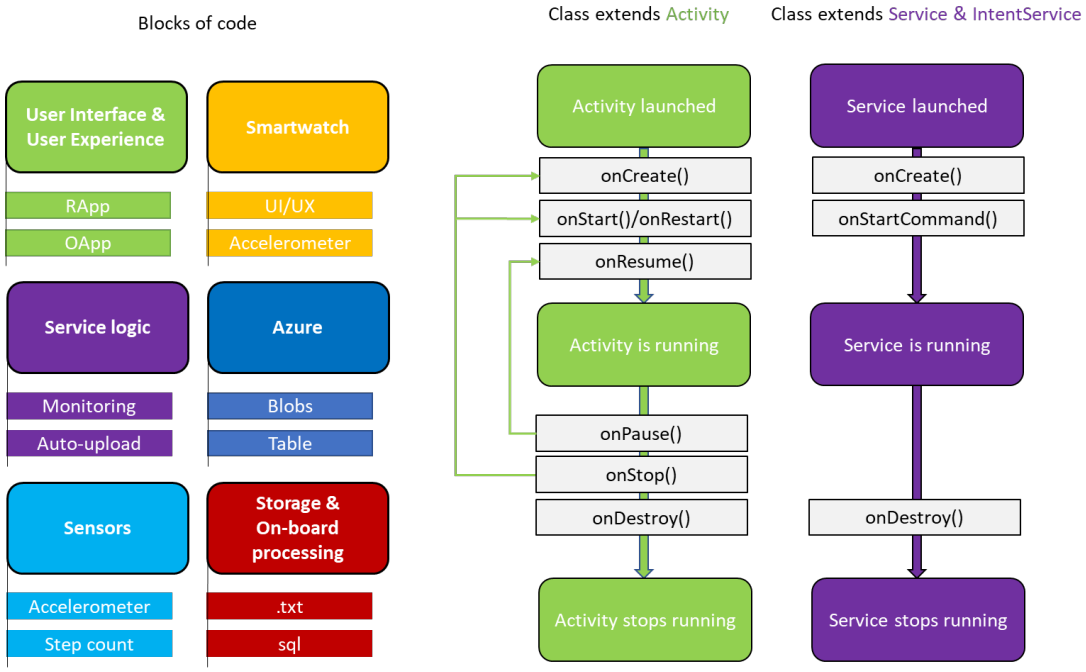


Figure 9: App architecture & Lifecycles

The continuous background monitoring and data upload are coordinated through a series of Android Services as detailed in figure 10. The **ServiceManager** handles the coordination of services and once started, verifies the settings (e.g., window, interval) to define the monitoring configuration and trigger **SensorLoggingServices** which oversees collecting the samples within the specified window, CPU frequency and specified format (i.e. raw and SQL entry). The scheduling of sampling (i.e., interval) and uploads is managed through *Alarms* which will trigger the corresponding *Services*.

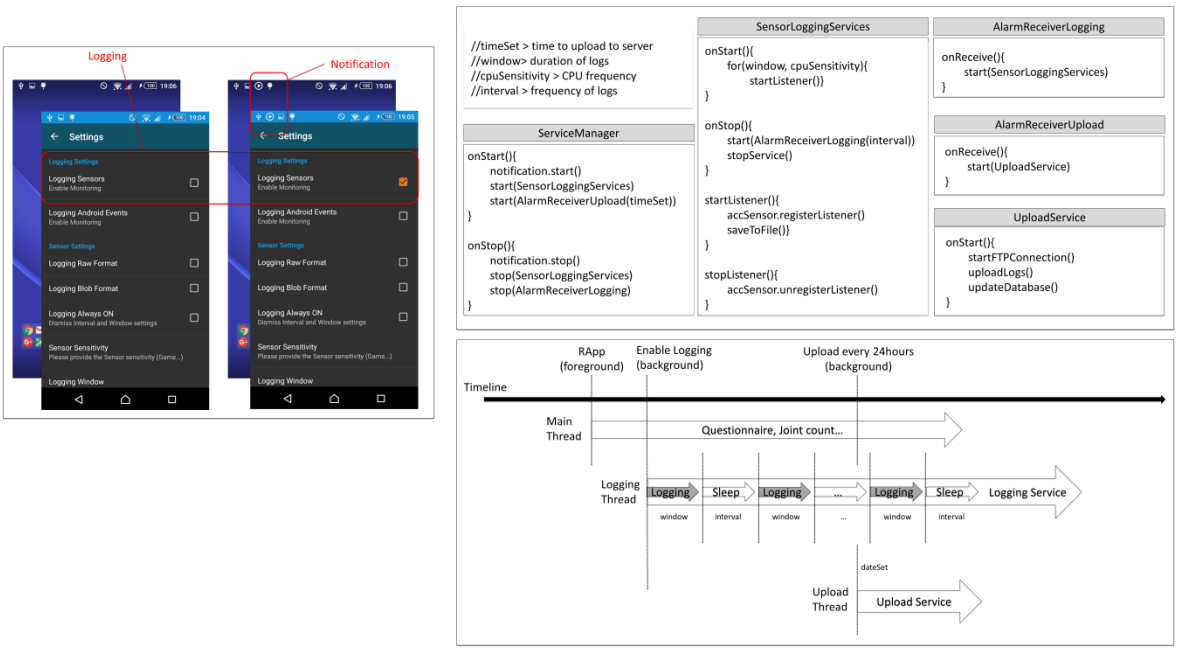


Figure 10: App lifecycle

The user interactions are defined as foreground activities designed considering the User Interface (UI) and User Experience (UX) that are respectively the visual layout and overall navigation. RApp™ and OApp™ have been developed as two variants of the same app with front end designed respectively for OA and RA patients. The **Settings** activity for both is hidden to the user and is accessed by clicking three times on the right corner of the banner. OApp™ has been developed for patients with Osteoarthritis with a single **Home** view that allows recording the intensity of overall pain on an ad-hoc basis and from either the smartphone or the smartwatch as per figure 11.

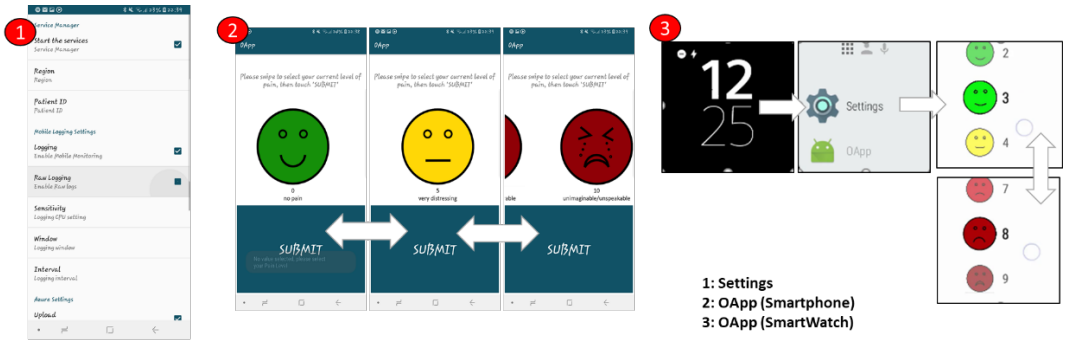


Figure 11: Settings and OApp™ navigation

RApp™ focuses on RA patients and launches an Activity upon opening the app built with *Fragments* allowing users to swipe between Home and Pain tab. The **Home** tab lists the last 30 days of **Questionnaire** activity in a *List* of *Buttons* and each day opens either the daily summary (if recorded) or the daily **Questionnaire** activity (if not filled).

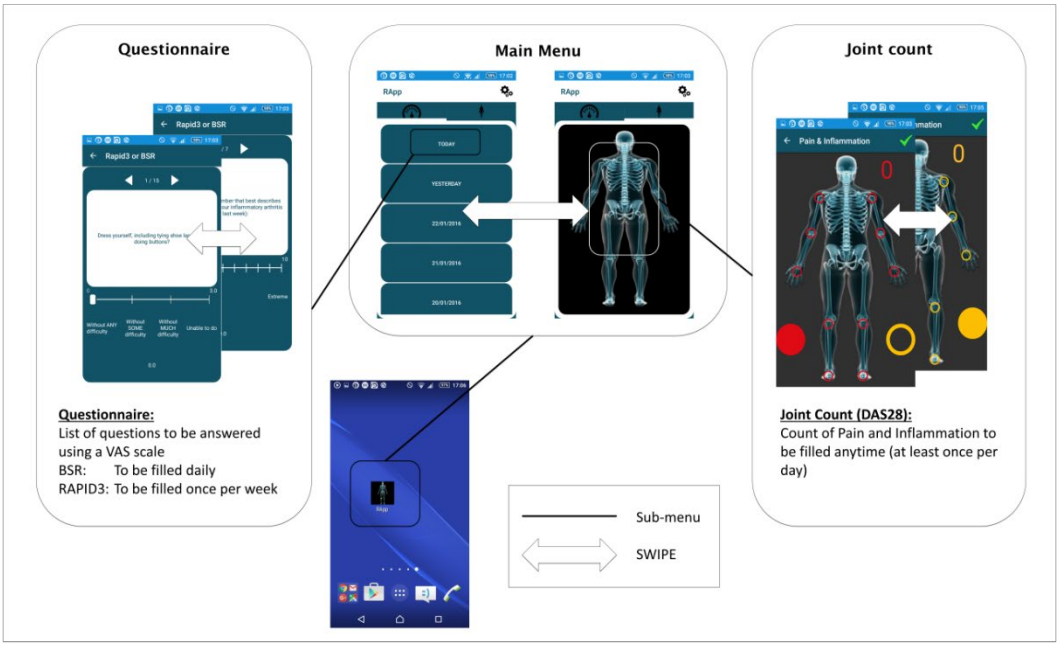


Figure 12: RApp™ navigation

Patient Reported Outcome Measues (PROMs) are self reported questionnaires by patients to self-assess their health status. The user interface included PROMs that do not require clinical

dependancy (such as blood tests). The **Questionnaire** activity prompts the user with the questions listed either as part of a questionnaire based on the British Society of Rheumatology (**BSR)**, for the daily report, or based on the routine assessment of patient index 3 (**RAPID3)** questionnaire, for weekly reports (see figure 12). The user response is recorded through a visual analogue scale (VAS) which is widely used and recommend for medical assessment (Harrison et al., 2009). Preloading scores with the last value entered has been considered but dismissed not to influence the patient when answering. Randomising the order of questions has been found to prevent patients from responding without reading but has not been incorporated to retain the integrity of using validated questionnaires.

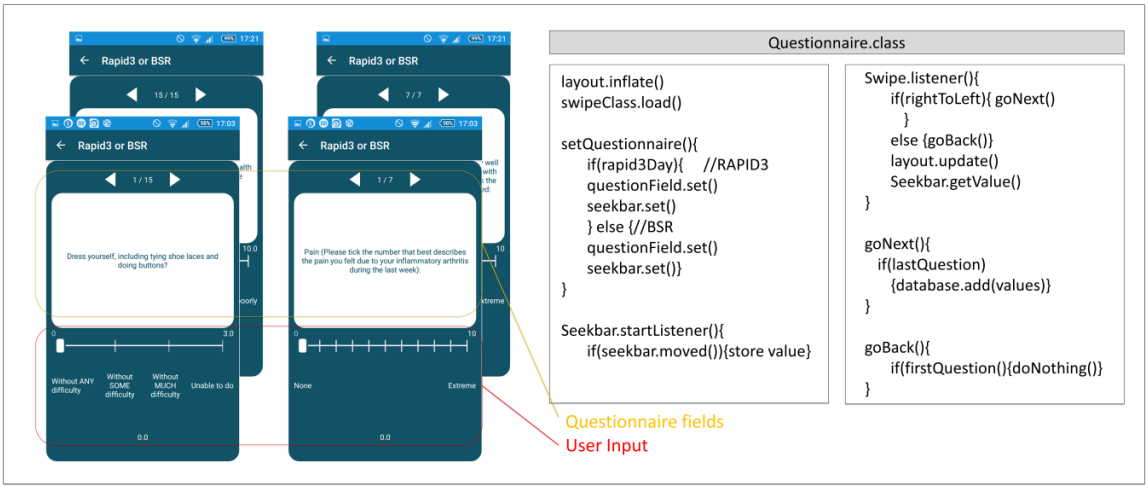


Figure 13: RApp™ questionnaire

The Disease Activity Score is a measured on 28 specific joints (DAS28) typically used by clinicians has been designed to record the joint count of pain and swelling (Ozlem Pala, 2006). The **Pain** tab launches the **PainAndInflammation** activity, which follows the model of DAS28 paper forms used at the hospital with a skeleton (see figure 14) and joints highlighted. However, the layout of the app includes only the shoulders, elbow, hands, knees, and feet to focus on large joints. Patient's daily sensitivity and perception (e.g., chronic pain, used to a certain level of pain) might help to correlate specific patterns of physical activity to flares. Subsequently an intensity factor, which is not a validated criterion, has been introduced to the layout allowing to quantify, from 0 to 4, the pain and swelling on each joint.

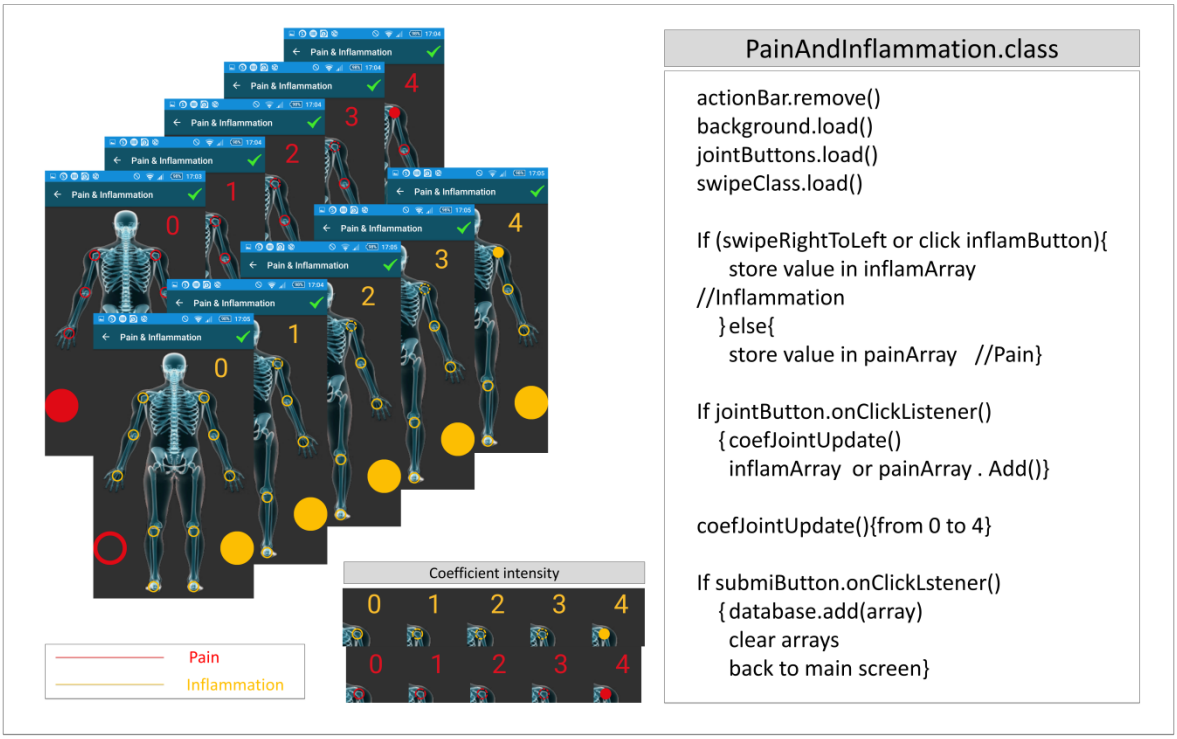


Figure 14: RApp™ DAS28

3.5 Data collection & analysis

Data format

The app is tracking user input but also a significant amount of raw accelerometer data. As detailed in figure 15, the design choice for Study 1 has been to save each sample of accelerometer data as individual text files stored on the smartphone's memory, and that can be accessed via USB connection. For the following studies, the accelerometer data are processed for load calculation, and the storage uses SQL databases which significantly reduces the amount of storage required and improved data access security as the database can only be accessed by the app (unlike data stored on the smartphone's internal memory). This solution also addresses Big Data challenges as a significantly lower amount of data is uploaded via secured HTTPS protocol to Microsoft Azure storage tables (instead of raw files uploaded as Azure Blobs).

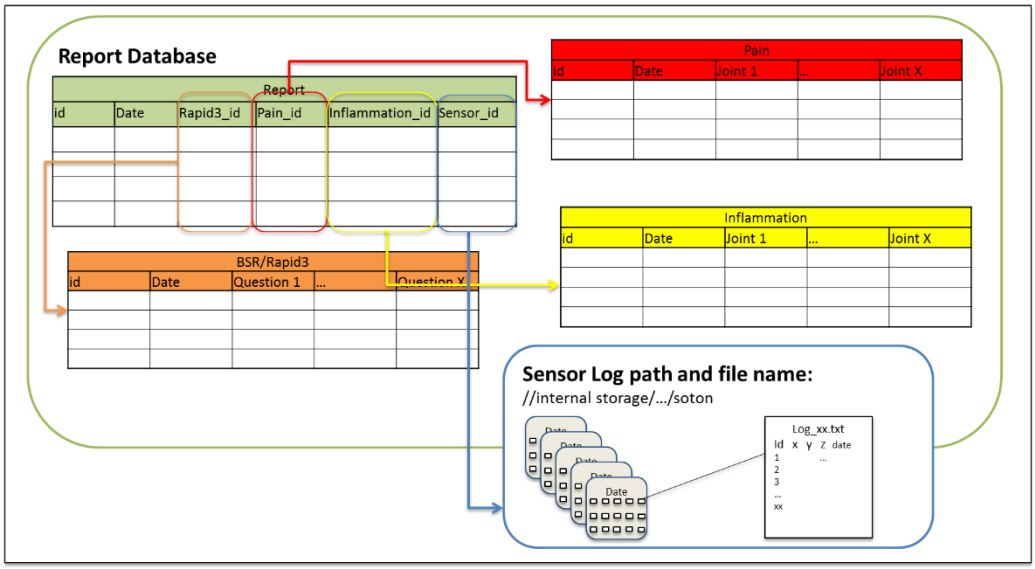
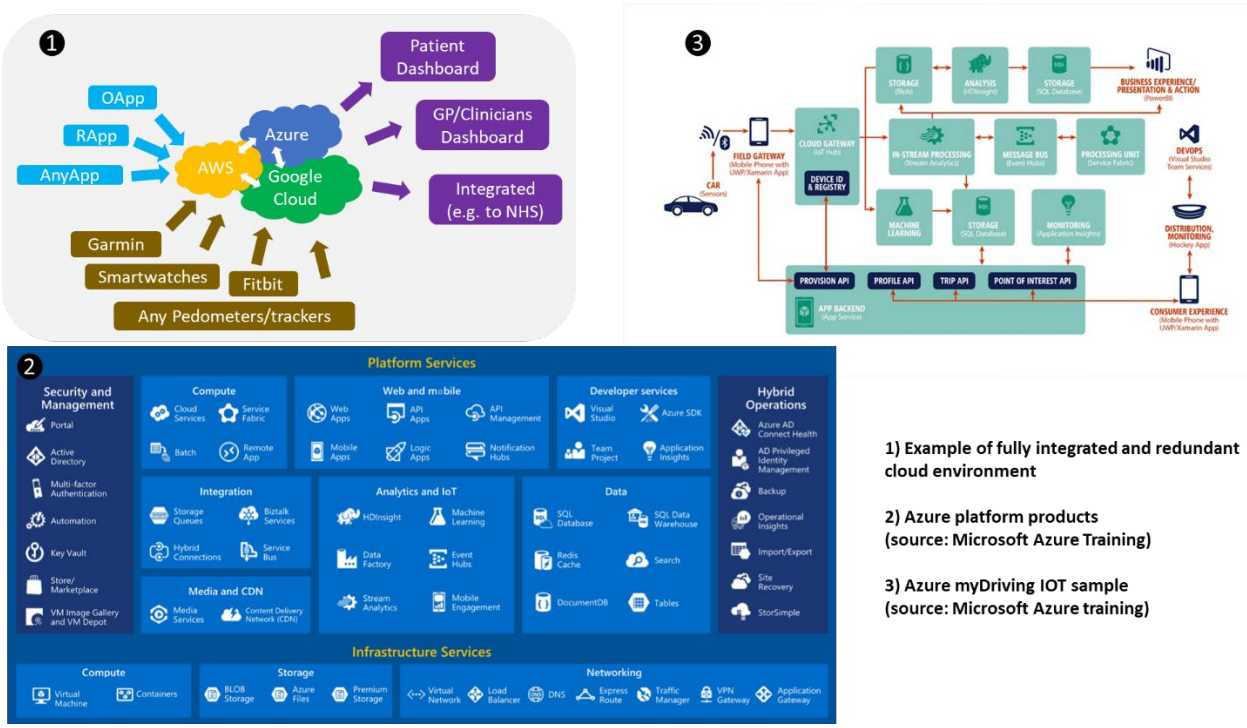


Figure 15: Storage and data format

Data storage

The data recorded are stored on the smartphone (or SD card) which relies on the read/write speed to hardware storage and generates a large amount of data for an extended period which causes storage size limitations and corruption (see study 1). A remote solution is preferred to prevent data loss caused by smartphone might malfunction or patients, e.g., not losing the device. There are many factors to consider for online solutions such as data redundancy, backup, security, data privacy regulations, scalability and deciding which components to implement. Cloud solutions offered by Amazon (AWS), Google (Google Cloud) and Microsoft (Azure) provide a range of products and services that natively address those concerns. Making use of these services is not limited to a single provider and multiple combinations of services (see figure 16.2) and providers could be used, e.g. to address cost-saving or to obtain a scalable architecture built with no single point of failure as shown in figure 16.1.



1) Example of fully integrated and redundant cloud environment

2) Azure platform products (source: Microsoft Azure Training)

3) Azure myDriving IOT sample (source: Microsoft Azure training)

Figure 16: Cloud architecture

At the point of writing, Azure is the cloud provider with the most certifications (Microsoft, 2017a). Azure's recognition in the health industry is a particularly good fit for this project considering that the data stored might include patient data (Microsoft, 2017b), which is particularly sensitive in terms of approval by the NHS. The geographic location used for storage can be specified (e.g. specified to the UK only) and defined to segregate resources (e.g. across study and applications).

Microsoft has granted this project with an allowance to experiment with the various components available. Figure 16.3 showcases the elements used for an end to end solution using IoT (internet of things) devices tracking GPS locations (Microsoft, 2017d). The solution uses an IoT hub that acts as gateway for the data uploaded and Azure App Services to support the various API calls. All data are stored in storage that facilitates the processing, analysis operations and real-time visualisation and services for both business and consumer experience.

This project uses the same principles applied to the sensor data but mainly focused on using the cloud-based blob and table storage component with most of the analysis done offline using MATLAB. Raw sensor data recorded are stored as text files, in the internal storage of the phone, and then uploaded as Azure Blobs, which is the preferred type for files (Microsoft, 2017c). Data processed by the smartphones (i.e., calculated load samples) are stored, in the internal memory, using Android SQL (Structured Query Language) database. Azure SQL is available, but Azure table storage is more flexible and allows the same functionalities without the constraints associated with relational databases.

Further capabilities might include integration with other platforms, e.g. to aggregate other data sources such as Fitbit or Garmin and could be achieved through commercial APIs. The value of data visualisation could be optimised and made available to more than one system, e.g. through the admin dashboard, purpose-fit layout for individuals and integrated with hospital records for clinicians (Dixon and Michaud, 2018).

Data processing

Analysis of the data collected is a core part of the project and is being reviewed in relevant section. Fitbit data were exported already aggregated as JSON file from Fitbit´s online portal. Smartphone´s data were exported directly from phone storage (Study 1&2) and from Azure storage (Study 3&4) using Azure storage explorer. Data processing was done through MATLAB and Python. Power BI was used for additional flexibility on the data visualisation. Data integrity in sample extraction, file corruption and sampling errors are reviewed throughout the project.

Smartphones continuously generate sensors data which were captured in sample chunks. The sampling for each study defines a window size (w), which represents the length or duration covered in the sample; and an interval (i), which represents a gap or timer before capturing the next samples. Accelerometer data are generated with a time (t) and coordinates along each axis (x, y and z) with a value that ranges from -10 to +10 m/s2 on each axis. The load rate estimates (LRE) are defined as the rate of change in load with respect to time, which is the body mass times the physical quantity jerk. The infinitesimal calculus of the load rate is defined as:

$$\dot{f} = \frac{df}{dt} = m\,\frac{da}{dt}$$

The estimated mean load rate magnitude (in kg m/s^3) is defined as:

$$\frac{\widehat{\Delta f}}{\Delta t} = \frac{\sum_{t(i=1)}^{t(i=n)} \quad m \times \sqrt{\left(\frac{a_{x,t=i+1} - a_{x,t=i}}{\Delta t}\right)^2 + \left(\frac{a_{y,t=i+1} - a_{y,t=i}}{\Delta t}\right)^2 + \left(\frac{a_{z,t=i+1} - a_{z,t=i}}{\Delta t}\right)^2}}{n}$$

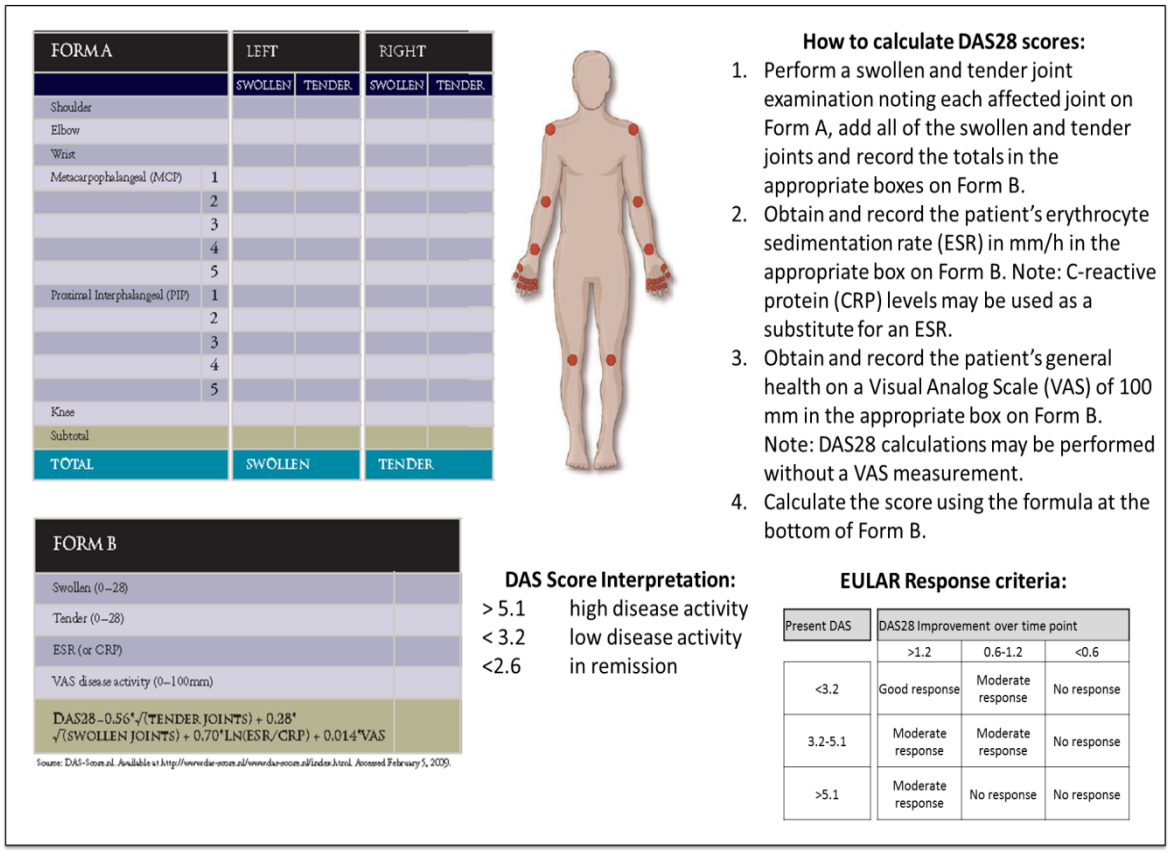
where a_x = is the acceleration in x direction, a_y = is the acceleration in y direction, a_z = is the acceleration in z direction, n = the number of data samples at interval Δt (i.e., 1/sample frequency).

Chapter 4 RApp1: Proof of concept with rheumatoid arthritis patients

This chapter covers the steps undertaken, considering the input from rheumatoid arthritis patients, to identify and assess the feasibility and key challenges of the project.

4.1 Introduction

The symptoms of RA patients develop gradually; therefore, monitoring the disease activity is critical and achieved using score-based assessment. The score calculation considers the lab results of blood samples, a joint assessment by the clinician, and a questionnaire-based assessment. The aim for patients is to become less dependent on medication, but the trend of the score drives the dosages. Increasing or higher scores imply that the patient's condition is getting worse, which can signify that a change might need to be introduced, e.g., an increase or new medication should be used. A trend in remission might show that a patient is responding well, e.g., that a particular treatment is working or could be lowered.



FORM A		LEFT		RIGHT	
		SWOLLEN	TENDER	SWOLLEN	TENDER
Shoulder					
Elbow					
Wrist					
Metacarpophalangeal (MCP)	1				
	2				
	3				
	4				
	5				
Proximal Interphalangeal (PIP)	1				
	2				
	3				
	4				
	5				
Knee					
Subtotal					
TOTAL		SWOLLEN		TENDER	

How to calculate DAS28 scores:

1. Perform a swollen and tender joint examination noting each affected joint on Form A, add all of the swollen and tender joints and record the totals in the appropriate boxes on Form B.
2. Obtain and record the patient's erythrocyte sedimentation rate (ESR) in mm/h in the appropriate box on Form B. Note: C-reactive protein (CRP) levels may be used as a substitute for an ESR.
3. Obtain and record the patient's general health on a Visual Analog Scale (VAS) of 100 mm in the appropriate box on Form B. Note: DAS28 calculations may be performed without a VAS measurement.
4. Calculate the score using the formula at the bottom of Form B.

FORM B	
Swollen (0–28)	
Tender (0–28)	
ESR (or CRP)	
VAS disease activity (0–100mm)	
DAS28 = 0.56*√(TENDER JOINTS) + 0.28* √(SWOLLEN JOINTS) + 0.70*LN(ESR/CRP) + 0.014*VAS	

Source: DAS-Score.nl. Available at http://www.das-score.nl/www.das-score.nl/index.html. Accessed February 5, 2009.

DAS Score Interpretation:

> 5.1	high disease activity
< 3.2	low disease activity
<2.6	in remission

EULAR Response criteria:

Present DAS	DAS28 Improvement over time point		
	>1.2	0.6-1.2	<0.6
<3.2	Good response	Moderate response	No response
3.2-5.1	Moderate response	Moderate response	No response
>5.1	Moderate response	No response	No response

Figure 17: DAS28 Form and EULAR response (source: DAS-score.nl (DAS-Score.nl, 2009))

The Disease Activity Score 28 (DAS28) is a clinical index calculated by clinicians using a form similar to figure 17 (DAS-Score.nl, 2009). The variables consider the measure of blood markers ESR (erythrocyte sedimentation rate) or CRP (C reactive protein), the GH (Global Health) of the patient and an assessment counting the pain and swelling across 28 joints. The trend of disease activity is evaluated using the EULAR criteria by comparing the current to the previous score. An improvement lower than 0.6 is considered non-responsive and moderate otherwise. For DAS lower than 3.2, a score improvement above 1.2 is regarded as a good response.

Several disease scorings exists, and the DAS28 scoring is the most widely used but considers a weighted scoring and has to be performed in a clinical setting to run the blood tests. The SDAI is an alternative that does not need to weight each variable and is calculated as a simple addition (Smolen et al., 2003); the CDAI does not consider the response of APP/APR, which means that blood tests are not required and provides the ability to conduct a disease activity evaluation out of clinical environments. CDAI and SDAI are less accurate and do not replace DAS28, but their validity has been demonstrated through various studies (Aletaha and Smolen, 2005). Recent research has shown a positive correlation between DAS-28-CRP, CDAI and SDAI at initial evaluation and SDAI and CDAI performing better than DAS for remission criteria (Dhaon et al., 2017).

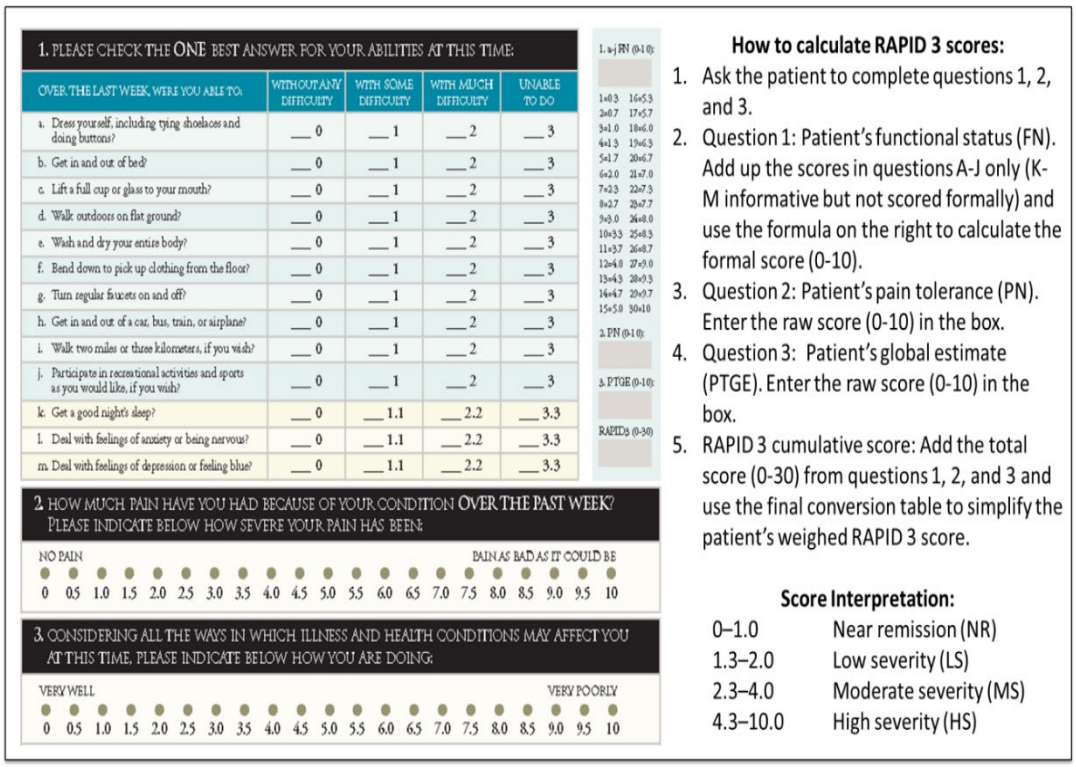
- $DAS28 = 0.56 \times \sqrt{TJC28} + 0.28 \times \sqrt{SJC28} + 0.36 \times \ln(CRP + 1) + 0.014 \times GH + 0.96$
- $SDAI = TJC28 + SJC28 + PaGH + PrGH + CRP$
- $CDAI = TJC28 + SJC28 + PaGH + PrGH$

- TJC28: Count of tender joints (0-28)
- SJC28: Count of swollen joints (0-28)
- CRP: Concentration of C-Reactive Protein (in mg/dl, between 0 and 10)
- PaGH: Patient global health assessment (from 0=best to 10=worst)
- PrGH: Provider (care provider) global health assessment (from 0=best to 10=worst

Figure 18: DAS28, SDAI and CDAI score calculation

Clinicians routinely use PROs (Patient-reported Outcome) which is recognised by the American College of Rheumatology (ACR) as being a core component for providing the patient's perspective on their disease activity (Gossec et al., 2016). The GH value (Global Health) used in the DAS28 calculation is measured through either PGA (Patient Global Assessment) or PTGA (Patient Global Assessment of Disease Activity), which can be used interchangeably (Khan et al., 2012). The PGA is one of the most used PROs that can be recorded either by the care professional (PrGH) or the patient (PaGH) by asking to grade their RA disease activity with a single value, typically between 0 and 10 (or 0 to 100) using a VAS (Visual Analogue Scale).

The principle of using questions answered by the patients has been expanded, and the HAQ (Health Assessment Questionnaire) is the first assessment developed to be completed by the patient rather than the doctor through a series of questions. It is widely used to assess the physical functionality (Bruce and Fries, 2003). Further questionnaires have been developed to focus on other matters, such as the HAQ-DI (Health Assessment Disability Index), which has been designed to measure the quality of life. The MDHAQ (Multi-Dimensional HAQ), also known as RAPID3 (Routine of Patient Assessment Index Data), has been developed to include physical function, pain, and global estimate, which are the three core data for RA (Pincus, 2007). The RAPID3 score is calculated by filling in a form as per figure 19 (Rheumatology, 2014) and has been compared to other disease activity indexes (Anderson et al., 2011). Studies have shown that similar quantitative information to DAS28 and CDAI can be obtained (Pincus et al., 2010) with equivalent values to DAS28, CDAI, and SDA in patients with lower disease activity (Kim et al., 2014). The severity of the disease activity is interpreted using the cumulative score (or weighted score in figure 19) and considered near remission for scores lower than 3, low for scores between 3 and 6, moderate for a score between 6 and 12 and high for a score above 12 (up to 30).



1. PLEASE CHECK THE ONE BEST ANSWER FOR YOUR ABILITIES AT THIS TIME:

OVER THE LAST WEEK, WERE YOU ABLE TO:	WITHOUT ANY DIFFICULTY	WITH SOME DIFFICULTY	WITH MUCH DIFFICULTY	UNABLE TO DO
a. Dress yourself, including tying shoelaces and doing buttons?	___ 0	___ 1	___ 2	___ 3
b. Get in and out of bed?	___ 0	___ 1	___ 2	___ 3
c. Lift a full cup or glass to your mouth?	___ 0	___ 1	___ 2	___ 3
d. Walk outdoors on flat ground?	___ 0	___ 1	___ 2	___ 3
e. Wash and dry your entire body?	___ 0	___ 1	___ 2	___ 3
f. Bend down to pick up clothing from the floor?	___ 0	___ 1	___ 2	___ 3
g. Turn regular faucets on and off?	___ 0	___ 1	___ 2	___ 3
h. Get in and out of a car, bus, train, or airplane?	___ 0	___ 1	___ 2	___ 3
i. Walk two miles or three kilometers, if you wish?	___ 0	___ 1	___ 2	___ 3
j. Participate in recreational activities and sports as you would like, if you wish?	___ 0	___ 1	___ 2	___ 3
k. Get a good night's sleep?	___ 0	___ 1.1	___ 2.2	___ 3.3
l. Deal with feelings of anxiety or being nervous?	___ 0	___ 1.1	___ 2.2	___ 3.3
m. Deal with feelings of depression or feeling blue?	___ 0	___ 1.1	___ 2.2	___ 3.3

1. a-j FN (0-10):
1=0.3 16=5.3
2=0.7 17=5.7
3=1.0 18=6.0
4=1.3 19=6.3
5=1.7 20=6.7
6=2.0 21=7.0
7=2.3 22=7.3
8=2.7 23=7.7
9=3.0 24=8.0
10=3.3 25=8.3
11=3.7 26=8.7
12=4.0 27=9.0
13=4.3 28=9.3
14=4.7 29=9.7
15=5.0 30=10
2. PN (0-10):
3. PTGE (0-10):
RAPID3 (0-30)

2. HOW MUCH PAIN HAVE YOU HAD BECAUSE OF YOUR CONDITION OVER THE PAST WEEK? PLEASE INDICATE BELOW HOW SEVERE YOUR PAIN HAS BEEN:

NO PAIN — PAIN AS BAD AS IT COULD BE

0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10

3. CONSIDERING ALL THE WAYS IN WHICH ILLNESS AND HEALTH CONDITIONS MAY AFFECT YOU AT THIS TIME, PLEASE INDICATE BELOW HOW YOU ARE DOING:

VERY WELL — VERY POORLY

0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10

How to calculate RAPID 3 scores:

1. Ask the patient to complete questions 1, 2, and 3.
2. Question 1: Patient's functional status (FN). Add up the scores in questions A-J only (K-M informative but not scored formally) and use the formula on the right to calculate the formal score (0-10).
3. Question 2: Patient's pain tolerance (PN). Enter the raw score (0-10) in the box.
4. Question 3: Patient's global estimate (PTGE). Enter the raw score (0-10) in the box.
5. RAPID 3 cumulative score: Add the total score (0-30) from questions 1, 2, and 3 and use the final conversion table to simplify the patient's weighed RAPID 3 score.

Score Interpretation:

0–1.0	Near remission (NR)
1.3–2.0	Low severity (LS)
2.3–4.0	Moderate severity (MS)
4.3–10.0	High severity (HS)

Figure 19: RAPID3 Score (source: American College of Rheumatology (Rheumatology, 2014))

Protocols and tools for monitoring of RA disease activity are already well established but generally must be run in a clinical environment which is restrictive. The assessment of joints measures their state at a given time (the patient´s visit) but not the level of mobility routinely undergone as part of the patient´s lifestyle. The motivation for this study is to evaluate how smart devices could

complement the clinician's visibility of a patient´s physical activity routine, beyond the details reported by a patient, through remote monitoring.

4.2 Aim & Objectives

4.2.1 Aim

This study aims to investigate the feasibility of using smartphones as a way to monitor RA patients' physical activity in actual life conditions and the correlations that could be drawn with their disease activity.

4.2.2 Objectives

The objectives for this study are:

1) Evaluate patients' perception of using smartphones to monitor their experience with the disease.
2) Assess whether smartphone sensor (accelerometers) can be used to monitor patients' physical activity without proactive user interventions continuously.

For this study, RA patients have been engaged in two parts:

1) Feedback and review through PPI session organised with RA patient screened from Biotherapy group
2) Data capture, in natural conditions of usage, recruiting 5 RA patients over six weeks.

4.2.3 Hypothesis

If patients use a smartphone, it is possible to use the smartphone´s capabilities to monitor their physical and disease activity remotely, outside of a clinical environment.

4.3 Methodology

4.3.1 RApp™ design

Blood tests require a clinical environment and healthcare professionals. However, patients can perform joint counts and validated questionnaires by themselves. Diaries and productivity apps already exist (see chapter 2). So, the front end of the app has been designed specifically to allow capturing these self-assessments on a more frequent basis. OA and RA have different

requirements, and RApp™ is a proof of concept (RApp™ v1.0, see chapter 3) used to get feedback on the UI, which includes DAS28 and RAPID3 for self-assessment, by RA patients. OApp™ provides a UI to record pain using a VAS. Additional features such as social media, productivity and gamification might make the app more engaging and increase the user retention rate but have been dismissed as not directly serving the purpose of this project.

An indication of successful app development would be allowing the evaluation of the correlation between disease activity and physical activity outside of a clinical setting. The app developed continuously monitors the load rate to draw an overall physical activity baseline, not specific to an activity type or exercise program. Subsequently, the framework could be applied to a broader range of diseases where continuous physical activity monitoring might be a key indicator. Patients' impressions and feedback are consistently captured along this project, on the methodology as well as the design of the user interface to aim at constructing an intuitive experience for patients.

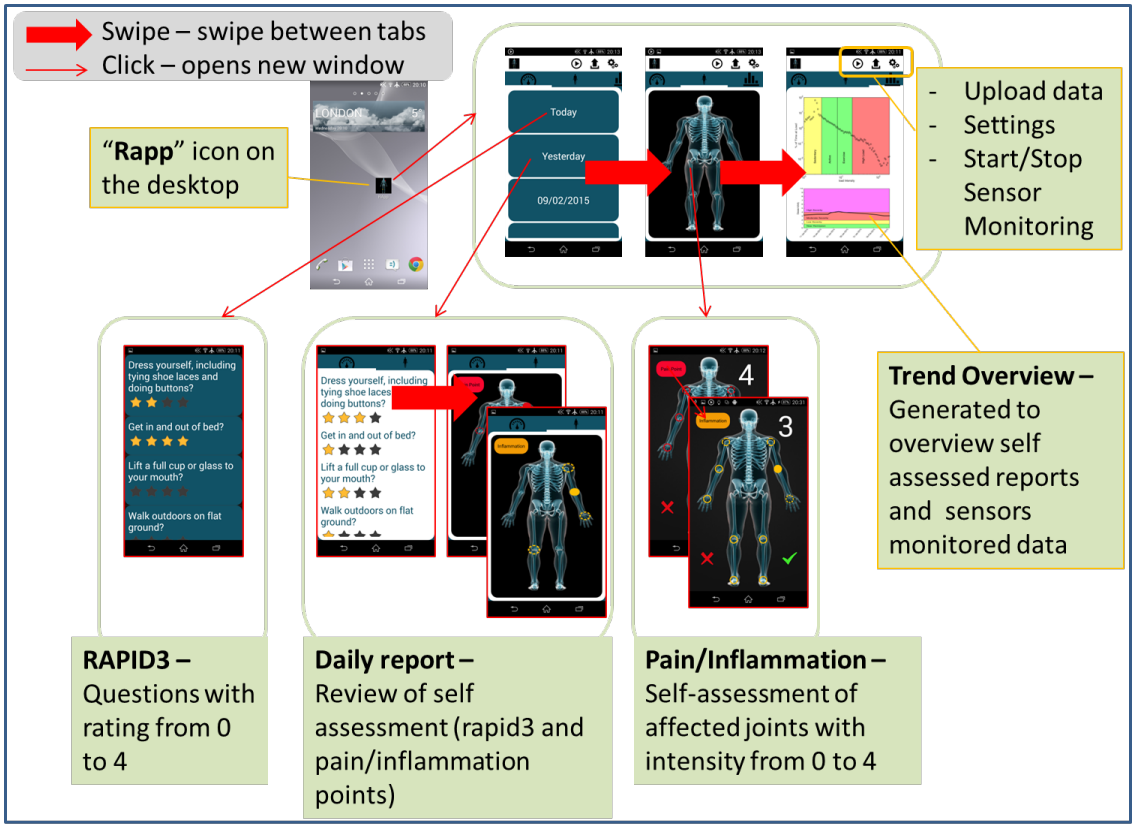


Figure 20: RApp™ Layout (PPI session)

4.3.2 Patient recruitment and study procedure

As described, the front end of RApp™ has been explicitly developed for RA patients and includes DAS28 joint count and validated questionnaires. Following INVOLVE recommendations

(healthtalk.org, 2017), patient and public involvement in research is a way to drive the investigation "with" rather than "to". So, a PPI group session was conducted for the first part of this study to gather RA patients' impressions. Participants recruited for a Biotherapy study have been used for convenience and representation of a mixed population of RA patients with no specific interest in smartphones or technological aspects of the project. The group was composed of 9 patients (8 females and one male), aged 36 to 65 years old and diagnosed with rheumatoid arthritis for 4 to 23 years. A questionnaire (see Appendix A) was sent to participants before the session to capture their perception before seeing RApp™. The first 20 min spent passing a Sony Xperia smartphone preloaded with the beta version of RApp™ around the group, without introduction or presentation, to gather raw feedback on the usability. The remaining 20 min focused on reviewing the comments to draw several conclusions and refine the app requirements.

The second part of this study is a pilot, observational and quantitative, focusing on comparing the trends of physical activity recorded via smartphones and Fitbit. Each participant wore a Fitbit HR wristband and used a Sony Xperia Z1 Compact preloaded with RApp™ v1.0. The Faculty of Engineering and Environment Ethics Committee at the University of Southampton approved this study (ERGO Ethics ID: 18061). The protocol and additional documentation, including the patient information sheet and patient consent form, have been reviewed and approved by London Stanmore's ethics committee (REC reference 16/LO/0182). Inclusion criteria were defined to recruit participants diagnosed with rheumatoid arthritis, willing to participate in this study and above 18 years old. Participants were five adults diagnosed with rheumatoid arthritis (female n=2, male n=3; aged 26 to 67 years old) recruited from the University Hospital Southampton NHS Foundation Trust database. The clinicians and nurses screened participants during Monday clinics, focusing on patients already or newly diagnosed with rheumatology arthritis. Patients interested were provided with a copy of the Participant Information sheet to allow a minimum period of 24 hours for review. Interested patients were invited for an initial interview at the Clinical Research Facility (CRF).

This first face-to-face meeting was focused on introducing the project and gathering initial thoughts about using the technology and about attitudes toward trust in technology and healthcare more generally. Once the consent form was signed, participants were provided with instructions and contact details in case of questions or issues. A loaned Fitbit wristband and smartphone preloaded with RApp™ v1.0 were then handed over to be used over the recruitment period. Each patient visit was logged in the patient note stored at the hospital, with the follow-up meeting and closing session scheduled to ensure that the six-week timeline also fitted with the patient's diary. Follow-ups were used to back up collected data, address questions, and gather continuous feedback. The closing meeting was used to gather participants' input on the overall

user experience. It included questions about post-experience attitudes to trust in the App and its impact on care delivery.

4.3.3 Data collection

For each patient, the data recorded on the first day has been dismissed and considered for the patient to get familiar with the instructions and for consistency of the datasets. Participants completed daily reports through the user interface of RApp™, specifically in the morning, usually when symptoms (e.g., pain) are most distinct. The daily report includes a DAS28-based joint self-assessment of pain and swelling, an MDHQ based questionnaire (BSR) during weekdays and RAPID3 weekly to fill in using VAS for each question. Patients were also invited to record joint assessments whenever experiencing significant pain or inflammation through the skeleton-based user interface, which allows highlighting of pain and swelling for each joint as well as the level of intensity with a scoring of 0 to 4 (see figure 20).

The monitoring aimed at covering a third of the continuous duration in a day (i.e., 20 sec per min) set to record data samples of 10 seconds (window=10) every 20 seconds (interval=20). The samples have been stored in raw format on text files on the internal storage of the phones along with the daily self-reports and uploaded every day at 23h59 in a daily batch of data to Azure Blob storage. Additionally, the files and reports stored in the smartphone were extracted mid-way manually and at the end of the recruitment period during the interviews with patients. Fitbit daily step count has been exported as an excel spreadsheet through Fitbit's online portal.

4.3.4 Data Analysis

For each patient, we evaluate the performance and accuracy of the app logging in the context of continuous monitoring for the recruitment duration (6 weeks). The daily amount of samples collected is compared to the amount expected to be collected, which, considering the settings used, should be two load rate values per min (i.e. two files per min, 2880 files per day). Periods without samples recorded indicate app and smartphone issues but recorded null values suggest that the smartphone is in a static position (e.g., on a table). The integrity of each sample is assessed by comparing the duration collected to the window set (i.e., each sample should contain ~10 sec of accelerometer data). Samples with a significantly wider window include multiple samples but have been considered corrupted, as there is no clear way to weigh each sub-sample. The Fitbit report is used as a baseline to identify peaks and drops of activity that could be used as data points of direct correlation. It should be noted that Fitbit recommends doing 10k steps per day (ROSENBAUM, 2019), which is widely encouraged through user interface and notifications.

However, the number of steps performed depends on each patient´s routine and no specific amount of steps was recommended nor set as expectation. With the Fitbit report limited to daily step count, the load rate values have been calculated through MATLAB for each sample and summed into daily means, allowing comparison of the trends of physical activity tracking over the recruitment period. The visualisation of data has been implemented through Power BI dashboards.

As detailed in the literature review, scores are used to evaluate disease activity, and, subsequently, the daily reports are used to identify significant peaks of disease activity, pain and inflammation reported by the patients that can be compared to the fitness baseline recorded. RAPID3 score calculation allows interpretation of the disease activity (high for a score above 12, moderate for a score of 6.1-12, low for a score of 3.1-6 and in remission for a score below or equal to 3). The BSR questionnaire is a lighter version of MDHAQ, and so we considered the score as the total sum from the questions, which is sufficient to draw a baseline while not being an accurate scoring. Likewise, the DAS28 score allows the interpretation of the disease activity (active for a score above 5.1, low for a score under 3.2 and in remission for a score under 2.6). The user interface of RApp™ allows reporting of pain and inflammation with a level of intensity but with the caveat of being limited to 10 joints (instead of 28). Therefore, the score used in this study might identify peaks and monitor the most sensitive joints but is not a validated DAS28 scoring. Joints 1,3,5,7, and 9 are the right shoulder, elbow, hand, knee, and foot; and correspondingly, joints are the left shoulder, elbow, hand, knee and foot.

4.4 Results

4.4.1 Patients' perception and feedback

Interacting with diagnosed RA patients is critical to understand how to capture their perception of this study. One of the RA clinics at Southampton General Hospital (see figure 21) is conducted on Monday mornings, where patient visits are typically scheduled and prioritised depending on patients' needs, from routine check-ups to injections or to be diagnosed.

The first observation while shadowing the interaction of healthcare professionals is that patients diagnosed with RA often have other family members, parents, spouses, or partners, involved in the disease monitoring and treatment. Secondly, it has been observed that medications and prescriptions are thoroughly agreed upon between the clinician and the patient [see Annexe C]. The efficiency of medicines and dosages is assessed through a review of the blood test results but also based on the patient's inputs and referred to as an "experiment" asking questions such as

"Should we try to lower/increase the dosage?" or "Should we try a different injection?". Patients appear to be informed, and their preference might overrule the most efficient remedy. During one of the visits, it was observed that one of the patients wanted to stay on tablet rather than getting injections, which may have been more effective due to potential side effects of hair loss. These observations were significant indications that patients´ engagement with RA disease is not limited to the time spent in the clinic. It affects their day-to-day routine as well as their social circle. In contrast, anyone has experienced catching the flu, getting healed and then moving on. For RA patients, extensive self-learning and proactive symptom monitoring occur daily. So, these observations were particularly encouraging and supported the thought that RA patients would be willing to use their smartphones as a tool to help their disease monitoring.

Figure 21: Southampton General Hospital (Victoria House)

The PPI group session allowed interaction with RA patients in the context of a clinical study. The first part of the pre-session questionnaire (figure 22) evaluates patients' perception towards their clinical visits and follow up and was worded in partnership with TRIFoRM (TRust in IT: Factors, metRics, Models) (Hooper et al., 2015). When RA patients were asked if they could recall their disease activity (1.1 and 1.2), including pain and swelling, 100% of patients indicated that they could remember details on the whole period, or at least the last month, since the previous visit. When asked if they would like their GP/Rheumatologist to be able to monitor their overall status between visits (1.3), 89% of patients responded that it would be useful for patients newly diagnosed or with high disease activity as the gaps between visits can be long. Still, it is not useful for patients in remission or with disease activity already under control. When it comes to monitoring physical activity (1.4), 100% of patients expressed interest, e.g., to see the impact on their disease activity, but not as an invasive metric that could dictate their day-to-day.

The second part of this questionnaire (fig 23 to 25) focuses on patients' relationships with smartphones. 89% of patients already own a smartphone (2.1), more than half running on Android OS (2.2), that they typically always carry with them or at least whenever being out of their house (2.3). 86% of patients use a smartphone above the minimum size, larger than 4.7 inches (2.4), with apps for managing their everyday life (2.5) that include email, calendar, and social media (2.6). When asked if an app could (2.7) and whether they'd be willing to use it to support their arthritis (2.8), 100% were keen to or neutral based on not having enough details on the usage. Specifically, the feeling is closely related to the disease activity, and not all patients would necessarily be willing to have phones continuously tracking their physical activity (2.9), nor might they be able always to carry it (2.10) as it also depends on the individual circumstances, e.g., not possible when the workplace does not allow mobile. 86% of patients would likely (or very likely) to use a smartphone to record their pain and activity (2.11), assuming that the user interface is easy to use to supplement face-to-face consultations and not replace them.

The PPI session allowed us to capture patients' feedback and impression on the app UI layout as much as usage point of view. A short introduction of RApp™ and the summary of answers to the survey sent before the session is in Appendix B and the most relevant patient quote for each question is listed in figure 26. The most frequent comment was that questionnaire and pain entry are great, but the overall layout and navigation should be more explicit. A "fun" factor and improved personalisation capabilities should be included for RApp™ to be perceived as an "app" rather than a medical tool. The wording of the questionnaire (Rapid3), in American English and aimed at consultations, are confusing for daily usage. Patients clearly expressed that they would need a reason to use an app, and several suggested medication tracking as improvements. RA patients with active disease are likely to be more responsive to using an app between visits but might not want to think about and monitor their activity if they are healthy or in remission. After seeing and testing the app, PPI participants understood how it could be used, which validated recruiting patients to use it, and further feedback was captured as part of the face-to-face interviews, which overall also received positive feedback.

Figure 22: Pre-session survey questions 1.1 to 1.4 (9 RA patients)

Figure 23: Pre-session survey questions 2.1 to 2.5 (9 RA patients)

Figure 24: Pre-session survey questions 2.6 to 2.9 (9 RA patients)

Figure 25: Pre-session survey questions 2.10 to 2.11 (9 RA patients)

Comments pre overview:
"I was expecting to receive more reminders to rest, do more or not"
"It feels a bit too clinical or being at school"
"The App should be more fun"
"It should be more user friendly"
"Feels too much like a diary"
"I'm already using myclinicaloutcomes.co.uk to track"
"I would like being able to press a button when it hurts and when it stops"
"How would patient benefit from using the app is not obvious if already at advanced stage. It seems more appropriate for patients early diagnosed"
"Could we personalize the body" (George Clooney mentioned as a joke)
"Questions are not great and too long"
"What happens if I can't take my mobile at work, would it still work remotely? This should be clearly taken into consideration by anyone taking part in the project"

Comments post overview:
"Great but I didn't see that"
"I didn't realize we could switch between the pain and inflammation"
"There is no save button to record your today's activity"

Figure 26: PPI comments pre and post an overview of RApp™

4.4.2 Patients using RApp™

The next part of this study consisted in getting the app used by RA patients. The smartphone application (RApp™) has been preloaded on smartphones we loaned to the participants to use as

a secondary smartphone, preventing them from transferring their personal accounts and SIM card. Table 3 shows the patient details and participation timeframes. The cohort included females (patients 1 and 4) and males (2, 3 and 5). The range of age was wide, with patients 1 to 5 being, respectively, 67, 57, 39, 45 and 29 years old at the time of the study. Patient 1 was recruited for six weeks, as expected, from 05/08/2016 to 16/09/2016, but duration varied due to schedule conflicts to meet patients and retrieve devices. As a result, patient 2 has been recruited from 06/09/2016 to 06/12/2016 (3 months); patient 3 from 13/10/2016 to 06/12/2016 (8 weeks), although no reports were completed after 06/11; patient 4 from 23/11/2016 to 16/01/2017 (7 weeks), although no Fitbit data recorded until the 07/12; and patient 5 from the 13/12/2016 to 02/03/2017 (10 weeks).

Patient			Smartphone		Recruitment		
ID	F/M	Age	Xperia ID	Use	Start	End	Duration
1	F	66	1	Secondary	05/08/2016	16/09/2016	6 weeks
2	M	57	2	Secondary	06/09/2016	06/12/2016	3 months
3	M	39	3	Primary	13/10/2016	06/12/2016	8 weeks
4	F	45	4	Secondary	23/11/2016	16/01/2017	7 weeks
5	M	29	5	Secondary	13/12/2016	02/03/2017	10 weeks

Table 3: Patient recruitment summary

Interview results

The patients recruited were provided with contact details in case of issues. There were no specific challenges in following the protocol or using RApp™ throughout the recruitment period nor during the follow-up interviews designed with questions to drive patient feedback that were approved for the study (see Annexe D).

Table 1 summarises the feedback when introducing the study. All patients expressed that RApp™ appeared easy to use (1.2) and agreed that it would not affect their treatment in any negative way (1.3), so the feeling toward using it was positive (1.4). The participants recruited were confident that internet-based technologies work correctly (1.5) and generally comfortable that it achieves their goal (1.6) and what they expect it to do (1.7). Table 5 and 6 summarise the feedback gathered during the follow up and last visit. At the conclusion of the study, the ease of use was highlighted alongside other feature improvements on the user interface. It should be mentioned that while patients expressed that they would not find an app useful when in remission, Fitbit's report and step count tracking has been well received and encouraged engaging in more walking activities.

Visit 1: Introduction

Patients id	2) How easy does RApp appear to use?	3) Do you think using RApp could affect your Rheumatoid Arthritis and its treatment in anyway?	4) How do you feel about using RApp?	5) Generally how confident are you that Internet based technologies actually do what they claim to do?	6) Generally how helpful do you find Internet technologies in supporting you to achieve your goals?	7) Generally how predictable do you think Internet technologies in carrying out what you expect them to do?
1	1/10	0/10	2/10	7/10	8/10	8/10
1	It seems ok	No	Apprehensive of using it	Love using internet and can find anything with Google.	Very good	Very good
2	3/10	0/10	10/10	8/10		9/10
2	seems ok		Good as could help other people (being part of a research)	Confident	Don't use muh tech, know it's there if need it, amount used would be 5/10. Don't rely much on tech but likes it when need it.	High expectations of technology
3	0/10			5/10	7/10	7/10
3	Very easy.	No	No worries, even if not for me it's still beneficial.	that they do what they claim.	Not too used, Satnav and FB.	Very reliable
4						
5	7/10			10/10	10/10	10/10
5	Very easy. Questions are not clear.	In the future, possibly.	Up for it, good, would be better if could use on any phone.	Very confident. The way it will be in the future.	Extremely helpful.	Very predicatble.

Table 4: Summary Visit 1 (Introduction)

Visit 2: Follow-up

Patients id	Other comments:
1	- Fitibit is encouraging to do 10k steps - Does more in the morning, tired in the afternoon - Would like to see the reporting even when feel good. For eg, it is rewarding to see when doing 5k steps
2	- Iphone has already an health app formonitoring steps. Fitbit seems to track more steps (806 through iphone vs 1251 for Fibit) but wear Fitbit longer. - Could fill the form every day. - No logs for period of 3rd to 21st of Sept 2016 but happy to extend participation in the study.
3	- Would use as ok for now. - Not too intensive. - Just bought a new phone but will stick to Xperia - When switch between P&I, it should be 0 (not the last value entered). - The dates should show the reports.
4	- The scale goes back to previous screen. - Questionnaire summary, swipe to go back takes to the beginning. - P&I, enter pain then tick does not go to Inflamation. - Missing joints: Ankle, hips, neck, back. - Questions (RAPID3): "physical well-being". Label "without any", difficulty… so scale is not clear. Should be on the scale. - Skeleton: which is left or right (is it mirrored image of you?). Maybe highlight if left or right. Click on the joint is nice but should be clearer. - P&I: 0 to 5 is nice to record.
5	- Not complicated. - The weekly questionnaire doesn't seem related to RA. - Phone keeps running out of battery (1.5 day), sometime dead battery. - Plays drum 3 times per week (for 2-3 hours). Could affect thte fitbit reading. Don't keep the phone.

Table 5: Summary Visit 2 (Follow-up after six weeks)

Patients id	1) How easy was it to use RApp?	2) How easy was it to understand RApp?
1	1/10	2/10
1	Easy once explained.	Questions every day is too much. Long questionnaire 3 times per week would be better. Same for P&I but to keep open for when flare happen. Using became part of routine and was not hard to do, takes 2 min.
2	1/10	1/10
2	Very straigth forward.	Self explanatory. Flows on its own. Very easy. Not complicated in any ways.
3	0/10	0/10
3	Easy	Easy but reports not always working. Good to have: to see the difference after stopped the medication (2 weeks before) as had MTX stopped for 2 weeks then back on lower doses. Could see the trend. Not used when was in Venice.
4	0/10	8-9/10
4	Easy	
5	8/10	8/10
5	Easily used. P&I: doesn't always work when press the buttons. Q: Easy to do. Never drag the VAS, just tap. Difficult and hard to know how to address, likely to always be the same. Knee was really bad in that period. If was run by specialist, would listen the recomendation in form filling. - Liked RAPID3 scoring better. Easier to do than 0 to 10 as 0 to 10 is very difficult to know how to gage. 10 implies disability, 7 hurts but can live with it, it depends on people. Scale from "without any difficulty" to "unable to do" is more useful for patient perspective. Explanation is very important to know how to answer.	Questions were not that easy.

Patients id	3) Did using RApp make you feel positive about your Rheumatoid Arthritis?	4) Did using RApp make you feel negative about your Rheumatoid Arthritis?	5) Could you highlight 5 positives comments?	6) Could you highlight 5 negative comments?
1	8/10			
1	Encourage to move more, did lots of steps. Positive as need to move with RA.	No, you got to live with it.	Tech is the way forward.	Can't think of anything negative. No hardship in using RApp.
2	Didn't make feel any different.	Didn't make feel any different.	Info callected can be good. Not difficult to use.	Nothing negative. Would be good on iPhone. Not all joints are included (for eg, elbow)
3	As positive as can be, it's got its purpose. Would like to see the trend. Ability to look back would be usefull for GP&nurses.	Not at all.	Good app and good for RA. Ability to monitor. Easy to use and to understand. People would use it. Layout is good and self explanatory for users.	Not fully working.
4	- Not in pain so don't think about pain. So more of a negative because had to think about it while feel well. - Report everydat is too much when no pain. - But if was in pain, it would be a comfort and positive.	See above.	- Easy to use. - Notification would be good for questionnaire, medication. - Reporting (Fitbit style). - Simplified UI would be good.	- If not in pain, make think about it. - Enter infos but don't get anything from it.
5	Positive. Scale helps put things in contrast.	Didn't feel negative at all.	- Easy to use. - Made feel good, felt as was being recorded, good psychologically. - Valid doctor Q, helps understand RA as a disease. - Light weight app, doesn't take long to load.	Xperia is bad, would use it on his phone. Couldn't see results/reporting.

Patients id	7) Any additional comments?	8) Would you use RApp again?	9) How well do you think this app achieves your goals?	10) How much do you think you can rely on it?	11) How much do you trust the app's performance?	12) How secure do you think the processes behind the app are?
1				10/10	9/10	10/10
1	Battery was draining even without using the phone. Last for about 2 days.	Yes	App side is more directed to clinicians. Need to incorporate more to be patient focus (for eg, include steps).	Perfect, only problem was upload to server.	Trust that it's doing its job with taking the information. Same as when submiting something online.	Very secure I would think. Security is very important. Anything that I don't understand, I don't touch. For eg, don't accept the friends requests from people I don't know.
2				10/10	7/10	
2	Nothing.	Yes, no personal use but for other people. Could be used to show history to clinicians. Would use it if on App store (even for nominal payment).	Don't know how to answer.	Reliable except for logging. Battery charge lasts only for a day (vs iphone that last 3 days)	Good. Logging not working and if battery is flat, needs to re-enable.	Fine
3						
3	When is it available, please keep up to date when can be used.	Yes	Good.	No problems, no issues, no crashes.	Haven't seen the report but fine.	Fine, no personal data. If someone wants to see what I've done for a day, that's ok.
4				9/10	9/10	9/10
4	Son has diabeteses and uses medical dev ice to record and upload the results. Then goes every 3 months at the clinic. Also self manage a bit by checking once per week.	It depends on disease activity, so not really. But would use if newly diagnosed.	See above.	Completely	No reasons to questions it. Trustworthy.	Hadnt thought about it so must mean that trust it. Don't put anything confidential, it doesn't matter who gets hold of it (no credit card details).
5		10/10		10/10		10/10
5		Yes, not without stats to use for doctor recommendations.	If there were a doctor involved, it would achieve taking the results.	Never crashed. Logging to re-enable when phone dies.	See above.	Trust it

Patients id	13) Do you think anyone could pretend to be you or access your data?	14) How predictable do you think the app when handling your input or recording your activity?	15) What are your thoughts about how much you can trust the processes behind the app?	16) Do you think that the people who might provide the app have your best interests at heart?	17) How confident are you that the app will continue to provide benefits in the future?	18) Tell us whether or not you trust the purposes to which the app is being used to help your care.
1	No	Very predictable, the questions are always the same.	Can't trust completely.	Yes, of course. Otherwise would not be able to sell it.	Yes, only will get better as this is for research.	Yes, trust completely also because done through the hospital. Credibility of Medical body is important.
2	People get your data. If nothing to hide, it doesn't matter.	Fine for recording. Can't see any problems.	Nothing specifial.	- Yes, if not charged for a lot of money. - If funded by NHS, for well being. - If Apple, for the money. Even if 1gbp, for the money.	Very confident.	Yes, definitely to help.
3	No.	Can't put so many different questions.	No thoughts.	Yes. If payable, would use it if tie with a doctor/GP.	Confident.	Trust it.
4	Yes. Anyone in the family could but not sure anyone would want to. There is no PIN or Password.	Don't understand the questions.	Hadnt thought about it. Would trust it.	Yes, cant think of any other reasons to do it.	Yes, I think it will after a few teaks: usefull for patients and clinicians. Makes you feel more in control and involved.	I do trust the purpose.
5	Nothing personal and believe it's secured. If someone does, doesn't really care.	Predicatable, questions are the same every day. Mixing might be good. No reward.	See above.	Guess so. They want to keep making money. Yes and No.	More info is always better.	Knew what was going into and that wouldn't actually being used for the diseases. Don't trust it yet and "benefits you, not me". The bigger picture is that when it goes live, it will help.

Table 6: Summary Visit 3 (Closing interview)

BSR & RAPID3 Self-assessments

All participants expressed that the questions were too redundant for daily purposes, but most completed the questionnaires. As shown in Fig27, patients 1, 2 and 5 completed the BSR (daily) and Rapid3 (weekly) questionnaires as expected, while patients 3 and 4 were intermittent. Yet, all patients utilised the DAS joint count to record pain and swelling.

Patients 3 and 4 completed the questionnaire more sporadically with BSR under 5 and RAPID3 under 2, indicating disease activity in remission. Patient 1's BSR questionnaire was answered daily and reached a maximum score of 28.2 with an average of 17.5; RAPID3 can be calculated as being above 7.6 for the last three weeks, which indicates moderate disease activity. Patient 2 also responded daily to the questionnaires over a longer time frame of recruitment, with the BSR averaging at 32 and reaching a peak of 39; RAPID3 be calculated for most weeks as being consistently above 10, indicating moderate disease activity. Patient 5 scored the highest BSR, above 40 on most days, and RAPID3 was calculated between 10 and 19 over the recruitment period, which indicates a moderate to high disease activity.

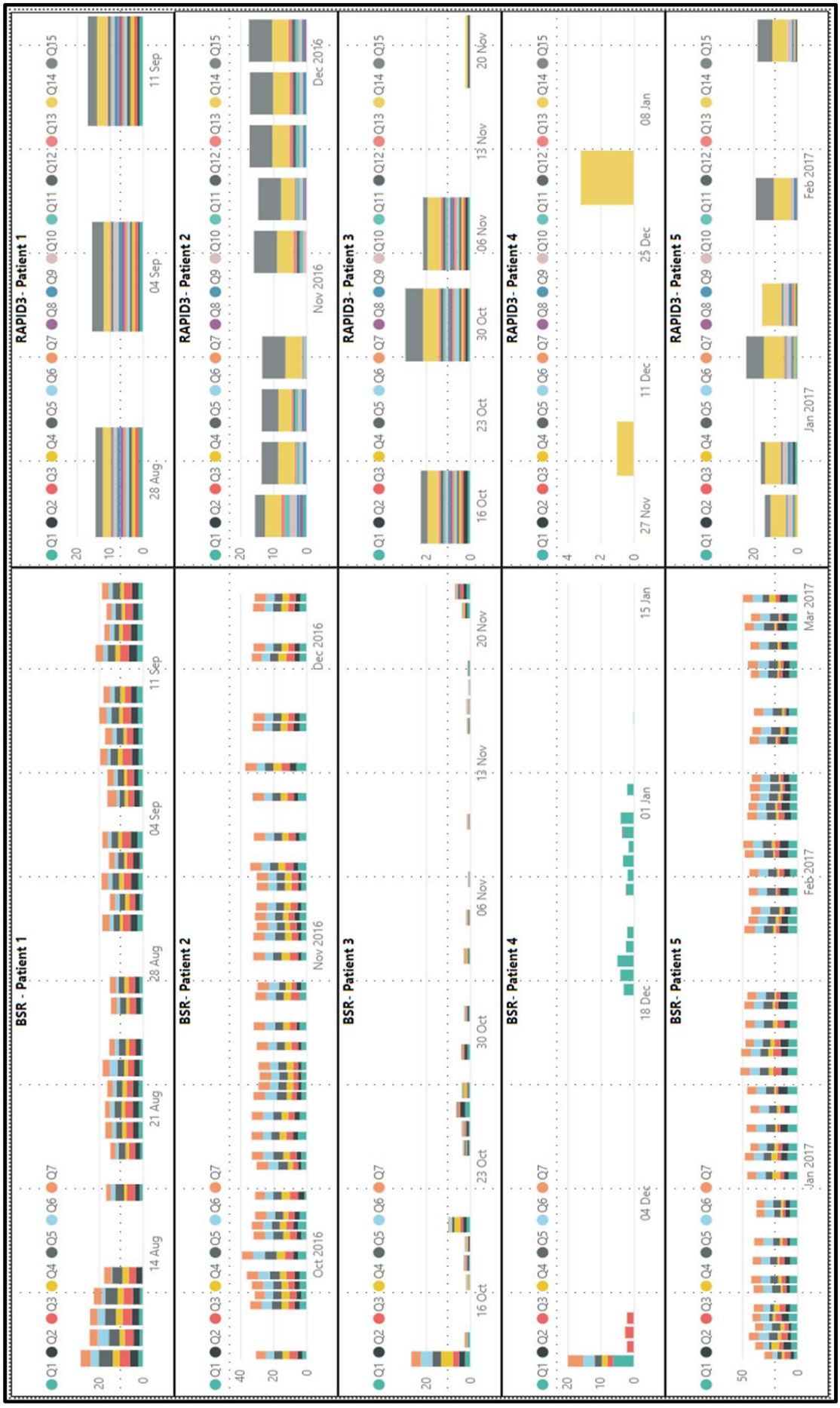


Figure 27: BSR & RAPID3

DAS28 Joint counts

Figures 28 and 29 show all joints' joint count of self-assessed pain and swelling (between 0 and 5). While the responses to patient 1's questionnaire might indicate moderate disease activity, the patient recorded pain on the left hand, knee and foot, scoring a max DAS score of 4 with an average of 0.64, which would indicate disease activity in remission.

Patient 2 recorded pain on both the left and right hand, knee and foot, consistently scoring a DAS score of 10 throughout the entire duration. No pain or inflammation was reported on the shoulders and elbow, but pain on the hand, knee and feet peaked at an intensity of 3 and consistently scored at two as well as being scored as inflated as 2 for the entire period. Patient 3 recorded pain on both shoulders, elbow and hands, consistently scoring a max DAS score of 5 with an average of 3, which indicates low disease activity. No pain or inflammation was reported on the knees and feet, but pain on the hands peaked at an intensity of 4 and consistently scored at one, as well as being scored as inflated for the entire period.

Patient 4 recorded pain only on the left knee with an average score of 0.62 during the recruitment period, which indicates a disease activity in remission. Still, a peak of pain has been highlighted through the intensity factor on weeks 5 and 6. Patient 5 recorded pain on the left knee and both hands and feet through the DAS, scoring with an average of 8.65, indicating an active disease activity over the recruitment period.

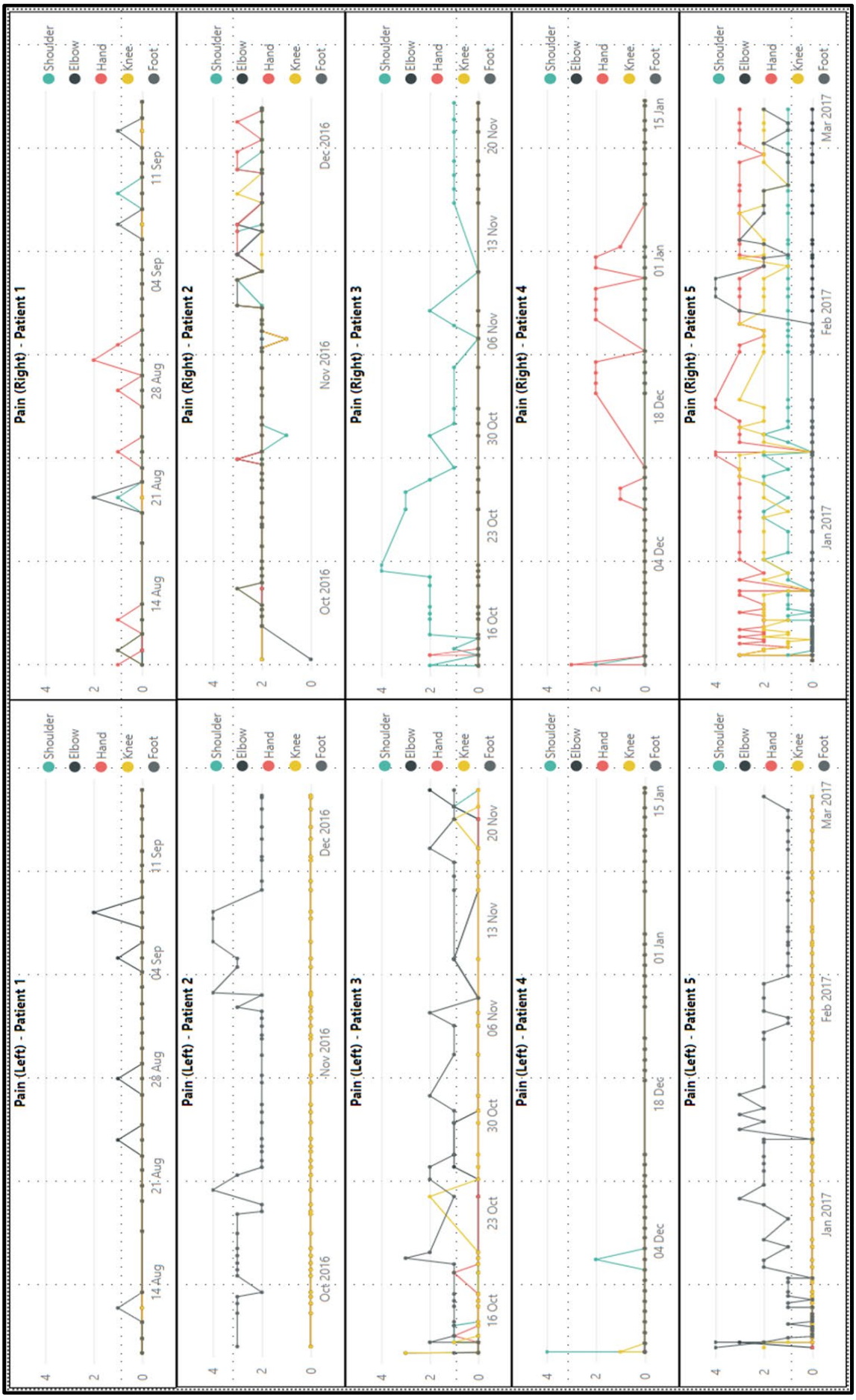


Figure 28: Summary Self Assessed Pain

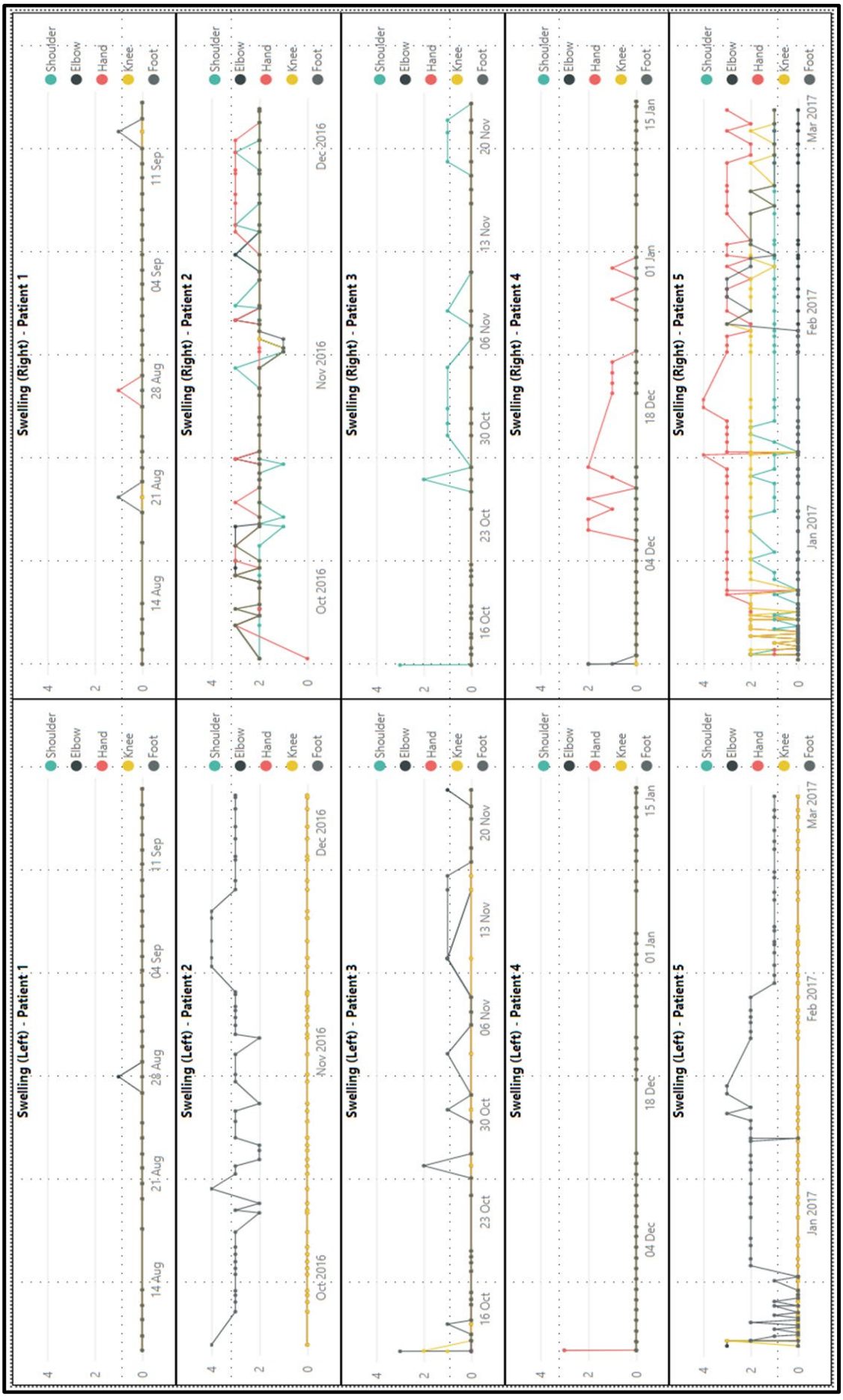


Figure 29: Summary Self Assessed Swelling

4.4.3 Continuous monitoring (RApp™)

4.4.3.1 Sample summary

Table 7 summarises the recruitment period with a number of readings captured. Patients 1 to 5, respectively generated 11.9, 24.8, 30.9, 31.8 and 56.1 million lines of sensor events (t, x, y, z). The difference in the range of data generated is mainly due to the recruitment timeframe but investigating the sampling rate performance is critical to accurate load rate estimates.

All patients showed significant variations in the number of logs generated per day. Patients 1 and 2 averaged 304.9 and 330 thousand lines daily, while patients 3, 4 and 5 averaged more than twice this quantity, with 753.8, 662.9 and 825.5 thousand lines daily. The number of readings was also dispersed across the day with standard variations of 107.8, 158.7, 156.1, 188.8 and 224.3 thousand. On their highest days, over a million lines were generated beside patients 1 and 2, which peaked at 0.5 and 0.7 million readings, the lowest days being when no data was generated.

Patient	Recruitment		Amount of Readings (per day)				
ID	Start	End	Total	Average	Min	Max	Std Dev
1	05/08/2016	16/09/2016	11891673	304914.7	6550	555294	107845.1
2	06/09/2016	06/12/2016	24791018	330546.9	8783	710874	158715.1
3	13/10/2016	06/12/2016	30905571	753794.4	269565	1008348	156179.7
4	23/11/2016	16/01/2017	31820357	662924.1	494	1063115	188839.6
5	13/12/2016	02/03/2017	56149771	825731.9	94461	1178147	224340.4

Table 7: Lines of sensor readings per day

Table 8 shows the daily amount of .txt files generated to store the lines of raw sensor readings. Patients 3 and 5 averaged 779 and 793 files per day which is more than twice that generated by patients 1, 2 and 4, with respectively 225, 244 and 573 files per day. The number of files generated daily was dispersed with standard deviations of 158, 174, 216, 234 and 337. On their highest days, a maximum of 639, 927, 1590, 1570 and 1984 files were generated.

Patient	Amount of files (per day)				
ID	Total	Average	Min	Max	Std Dev
1	8776	225.03	12	639	158.48
2	18317	244.23	3	927	174.07
3	31966	779.66	226	1590	216.5
4	27522	573.38	1	1570	235.43
5	53957	793.49	139	1984	337.64

Table 8: Amount of files collected per day

The number of sensor readings per file also had much variation. Table 9 shows the average number of readings per file is above 1100 for most patients except patient 3, with 966 lines per

file and high standard deviations of 1432, 1453, 879, 1054 and 949. The highest number of readings per file was seen with patients 1, 2 and 3, respectively at 13795, 13557 and 12252 lines.

Patient	Amount of Readings (per file)				
ID	Total	Average	Min	Max	Std Dev
1	11891673	1355.02	117	13795	1432.64
2	24791018	1353.44	25	13557	1453.11
3	30905571	966.83	8	12252	879.94
4	31820357	1156.18	14	8655	1054.85
5	56149771	1040.64	3	9726	949.45

Table 9: Lines of sensor readings per file

4.4.3.2 Smartphone movements

The top plot of figures 30 to 34 shows the smartphones' accelerometer over the x, y and z axis, and the bottom plot shows the corresponding jerk calculated. Quantifying the phone's movement requires looking at all axes so the jerks can identify the movement's peak.

The period of blanks corresponds to periods where no samples were collected and occurred only intermittently for patients 2 and 3 but more frequently for patients 1, 4 and 5, where multiple days of data were missed. The sensor values vary from -10 to +10 m/s^2 on each axis, and a more prominent colour (e.g. orange) means more prominent movement over the corresponding axis (e.g. z). The activity on the z-axis is more prominent than on the x and y-axis because of the force of gravity that constantly applies to the phone. A higher average of z means that the phone is at constant acceleration, i.e., not actively used, as can be seen for patients 1 and 3. The variations on the x and y axis are less common and signify that the phone is in motion, either interacted with, which would typically be through short and sharp interactions, or as a surrogate of the user's movement while carrying the phone, which is typical as part of a pattern of movements.

The average jerk values observed (in $kg.m/s^3$), 9.58, 9.69, 9.67 and 9.72, are expected, considering the force of gravity of 9.8, which is where the phone is inactive and not moving. The movement is characterised by values greater and lower than 9.8 with the latter, not to misinterpret as less movement but rather as the phone´s jerk countering gravity´s force. The max values, however, highlight the highest load which reaches 14.63, 26.95, 36.91, 28, and 31.66. The overall shape of the jerk shows the patterns of movement with patient 1 as the least active, with jerks rarely above 13. In contrast, the other patients exhibit patterns of jerk density, indicating routine of more frequent activity.

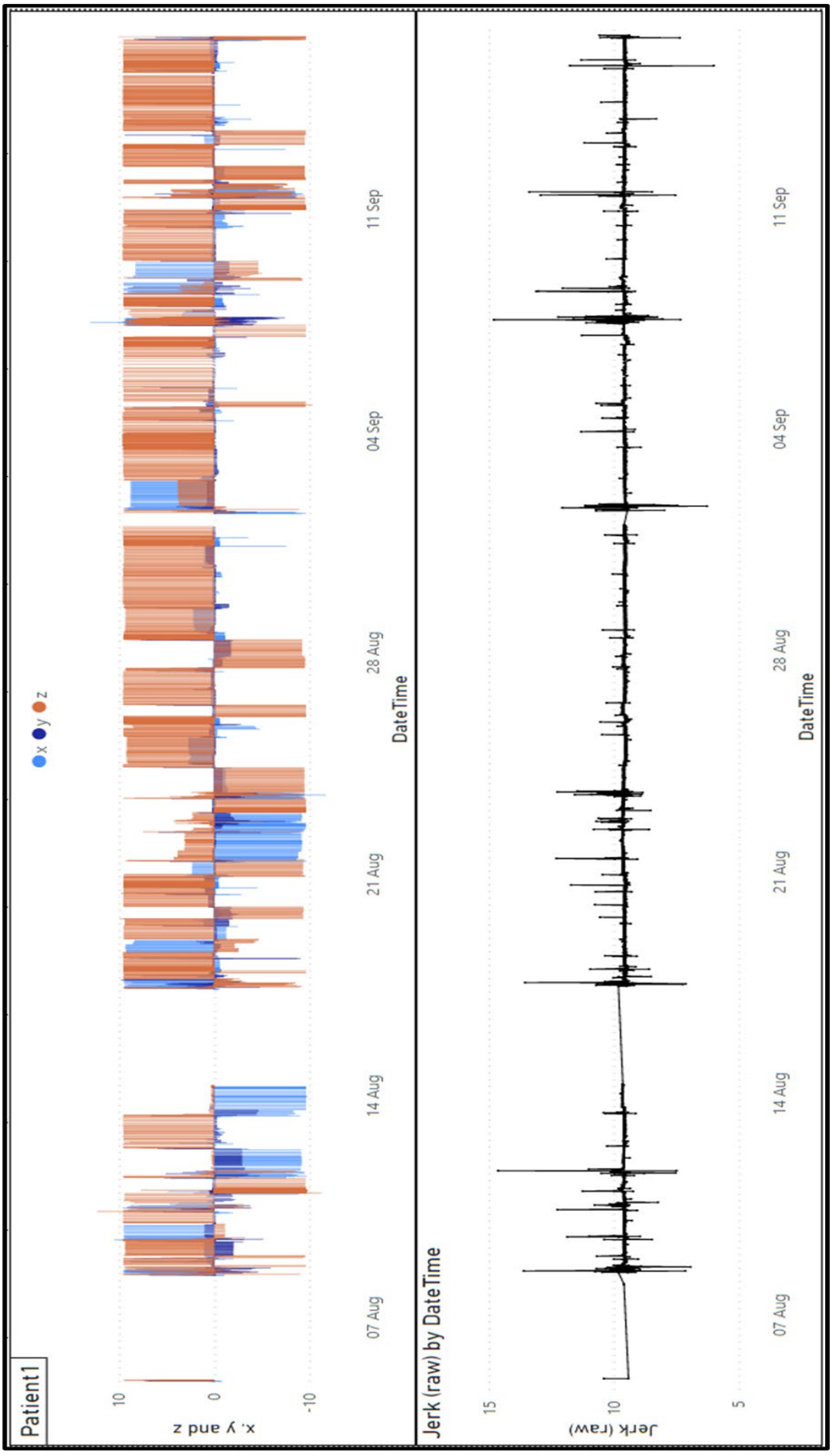


Figure 30: Smartphone's sensor Patient 1

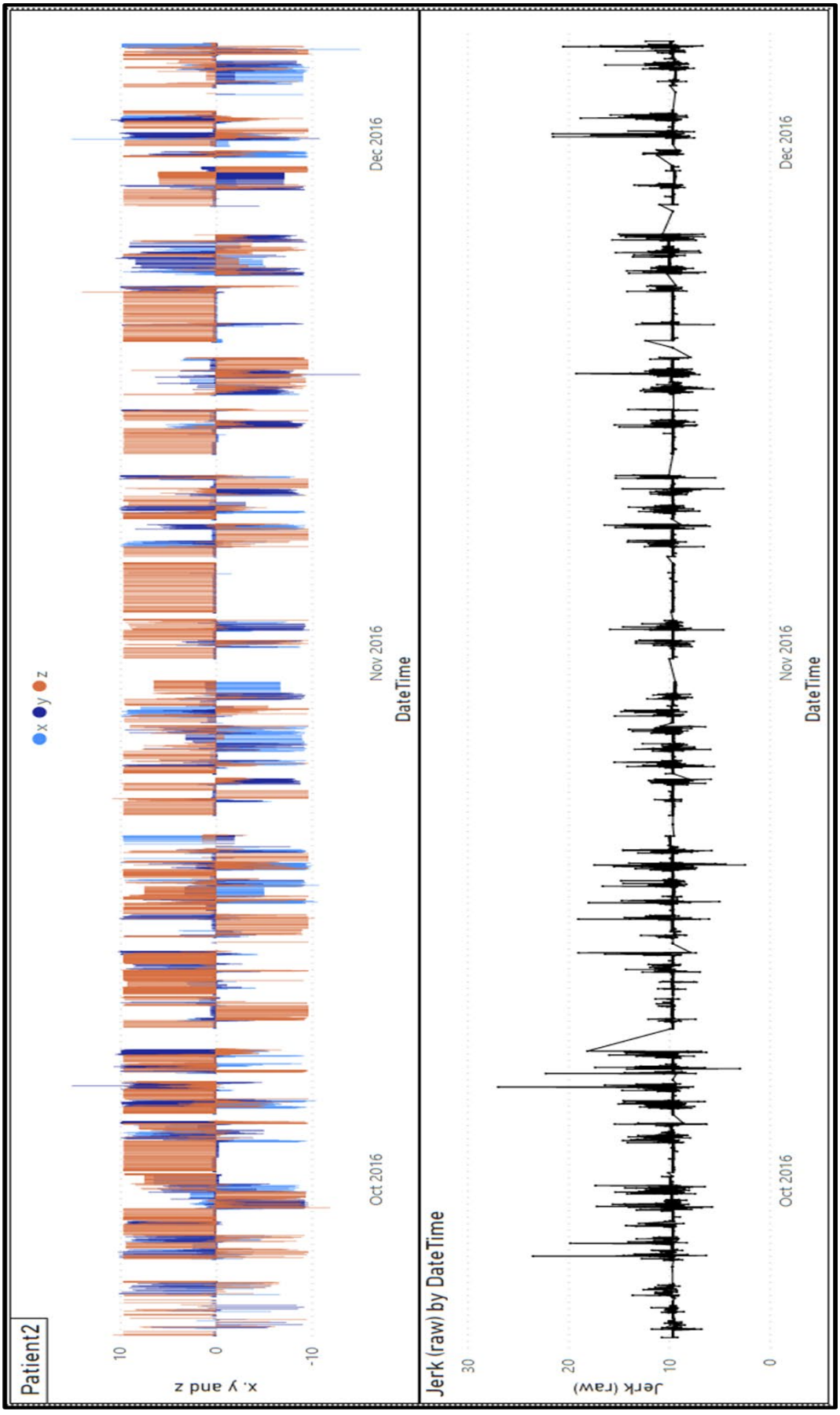


Figure 31: Smartphone's sensor Patient 2

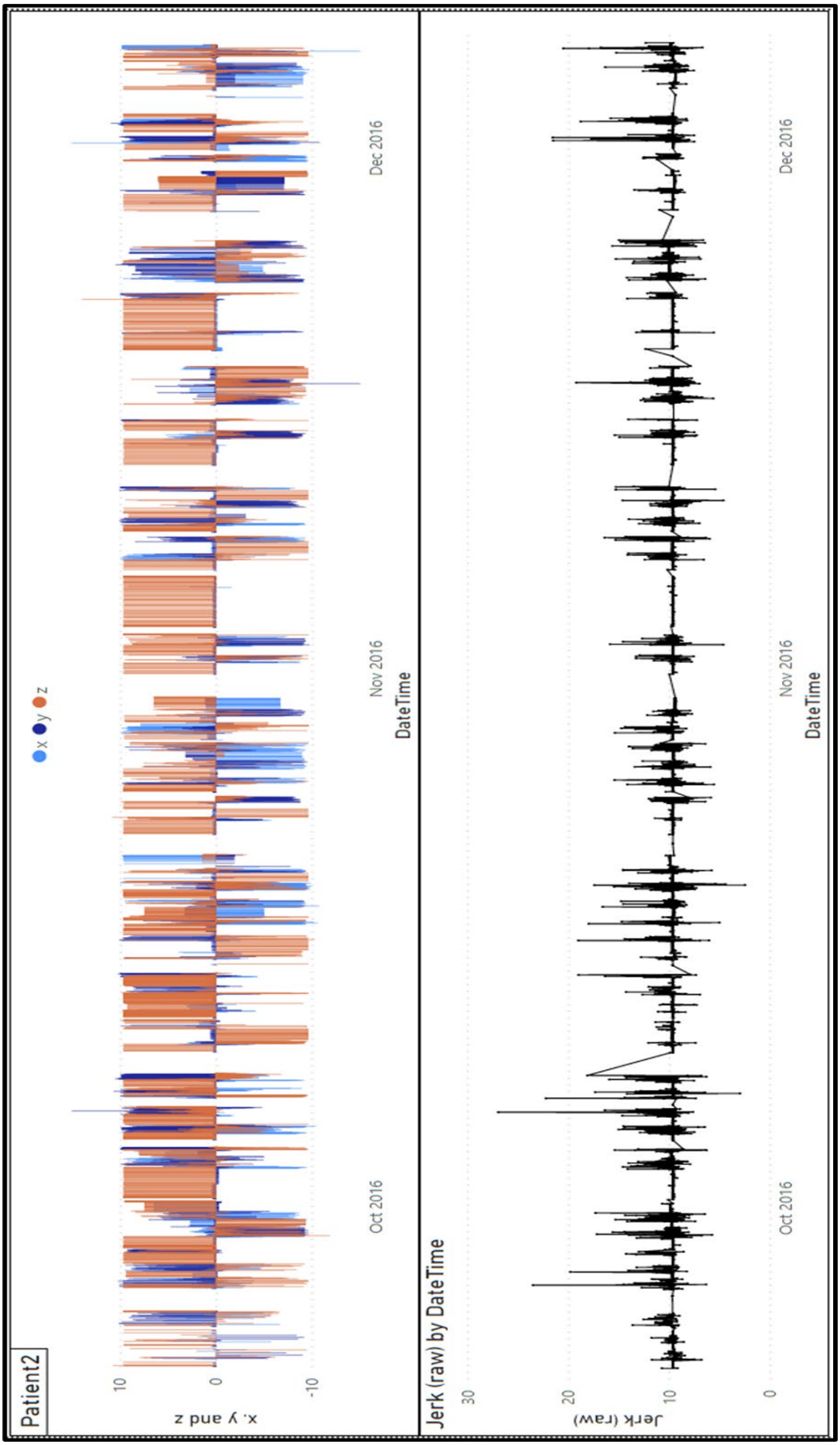


Figure 32: Smartphone's sensor Patient 3

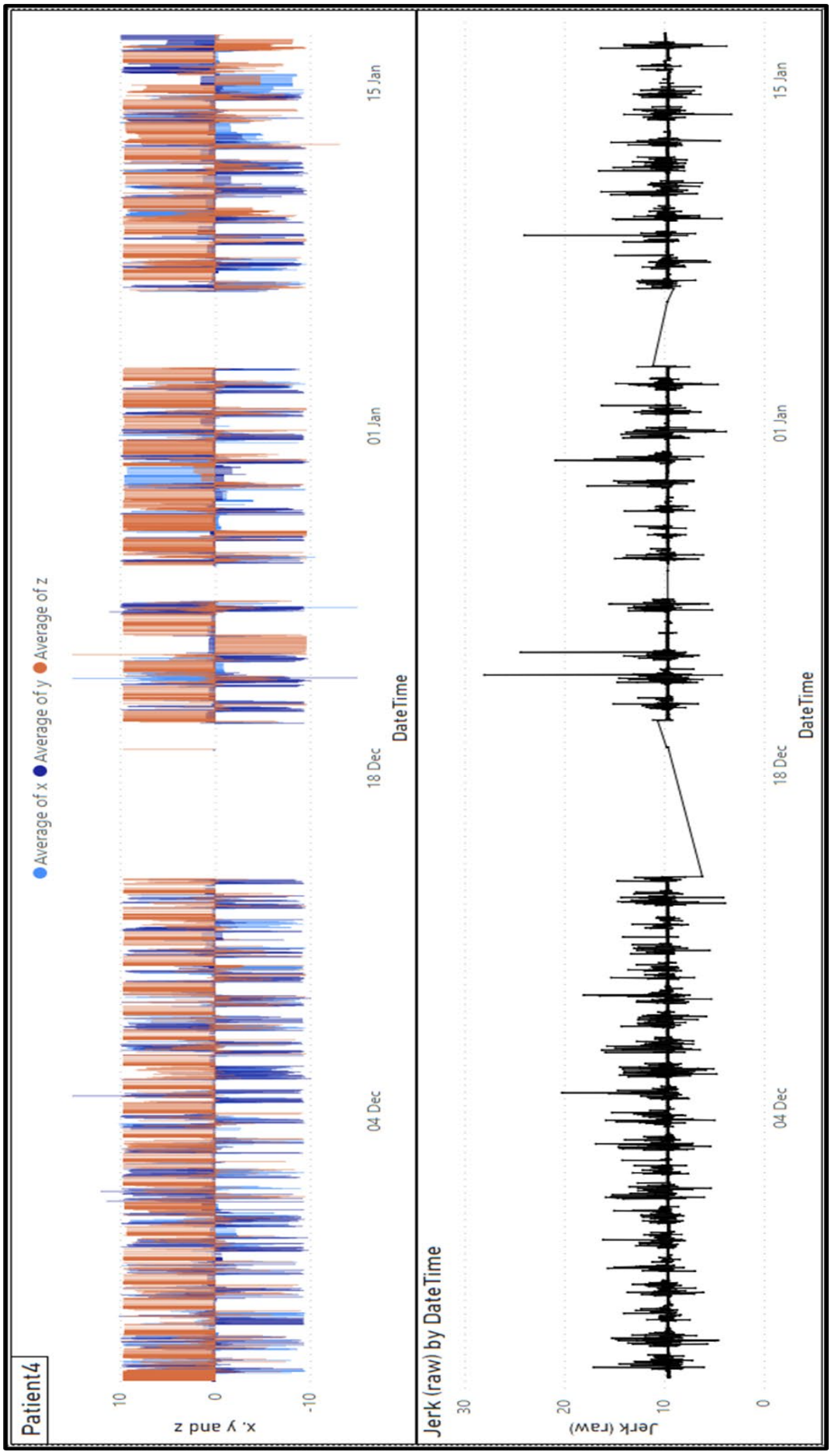


Figure 33: Smartphone's sensor Patient 4

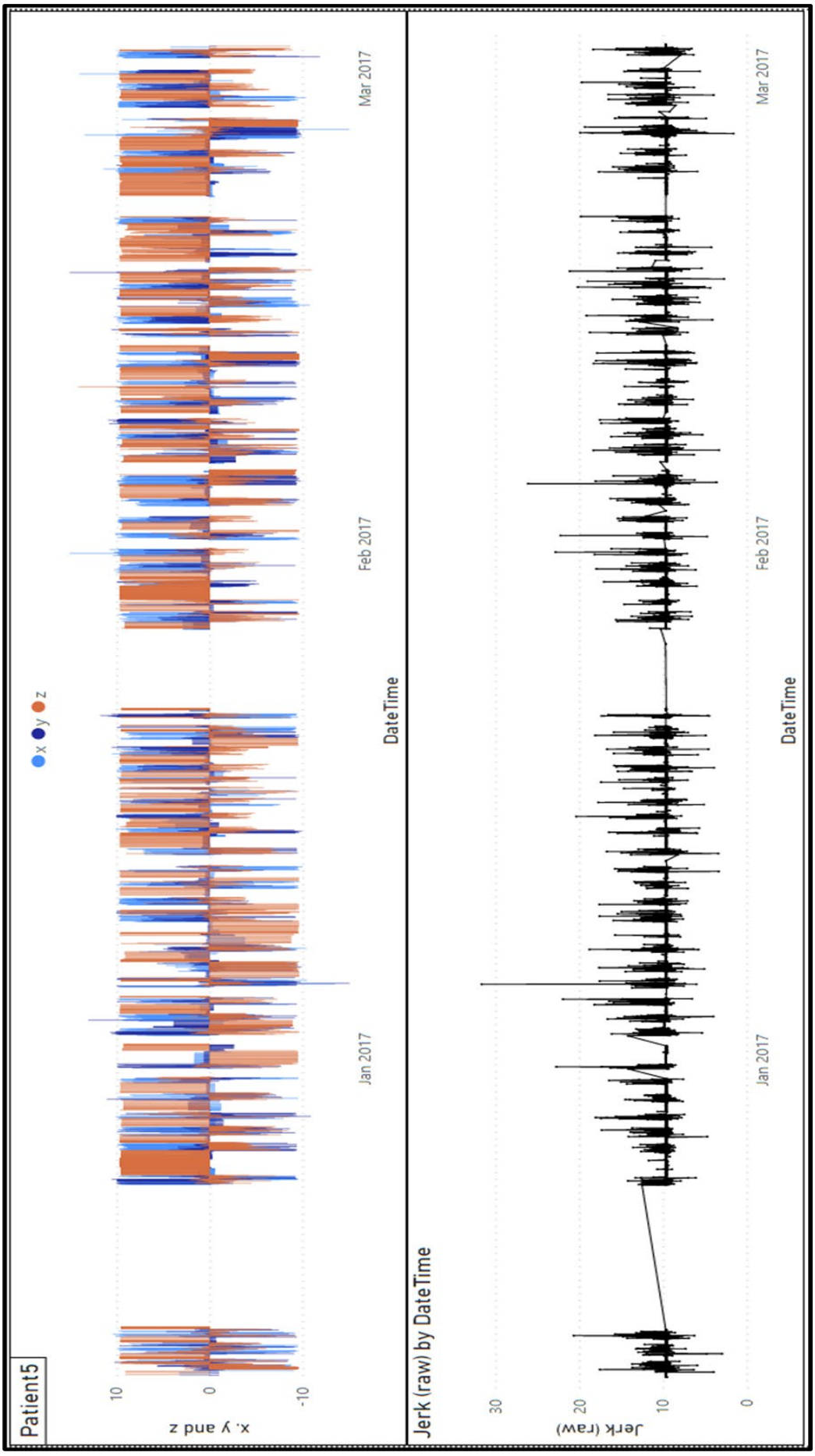
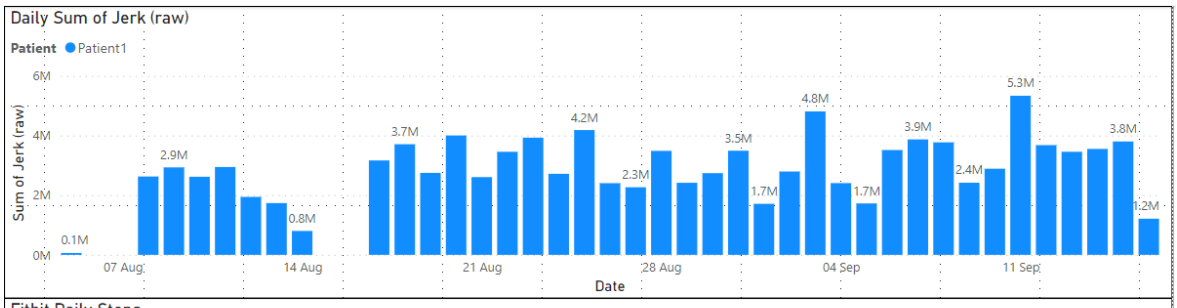


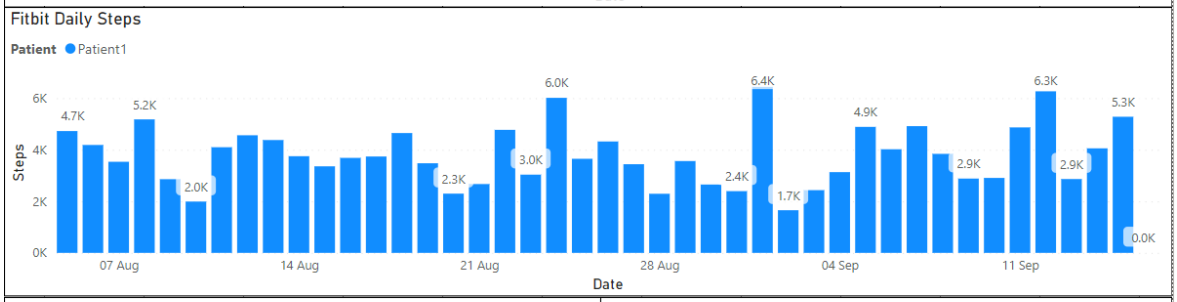
Figure 34: Smartphone's sensor Patient 5

4.4.3.3 Fitbit steps and smartphone´s LRE

Figures 35 to 39 show the daily amount of steps recorded through Fitbit and the daily sum of load rate recorded by the smartphone. All patients considered their lifestyle physically active apart from patient 1, who was considered as not particularly active and relatively sedentary. This is confirmed by looking at Patient 1's daily average of 3,711 which is significantly lower than other patients, averaging at 11044, 17404, 12478, and 9146 steps per day. Patients 1, 2 and 3 have max daily steps similar across months while patients 4 and 5 show more spikes of daily steps. The standard deviation of 1253 steps shows that patient 1 is the most consistent, with patient 3 and 5 having the highest standard deviations at 7441 and 5229 steps.

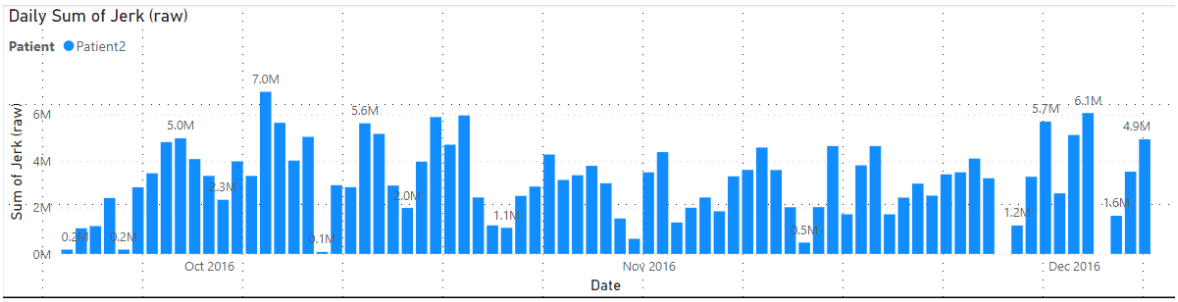
The jerks have been aggregated as daily sum to compare with Fitbit´s daily sum of steps. As expected, the daily sum of jerks is proportional to the recruitment period, and Fitbit's daily sum of steps. However, the sum of jerks varies much more across patients than the sum of steps seen between patients 3 and 5, both totalling 73 thousand steps but with the sum of jerks at 299 and 545 thousand, almost double each other. It is also noticeable that patient 2 takes two months to generate as much jerk as patient 3 did in two months.

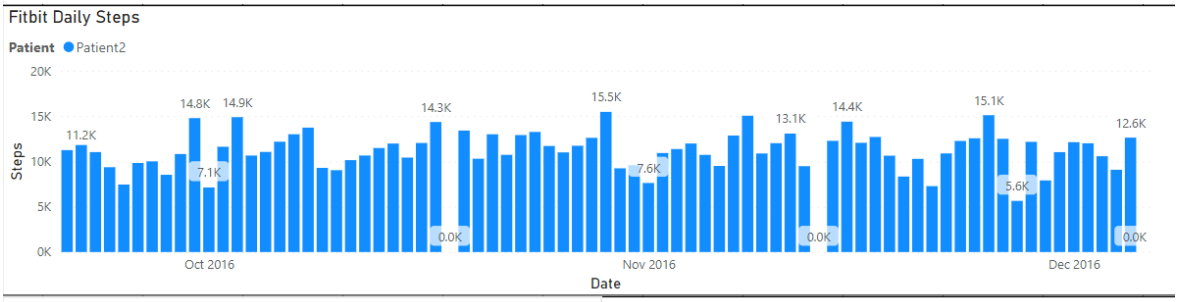




Patient	Sum of Steps	Average of Steps	Max of Steps	Standard deviation of Steps	Sum of Sum of Jerk (raw)	Average of Average of Jerk (raw)	Max of Max of Jerk (raw)
Patient1	**159594**	**3711.49**	**6435**	**1253.80**	**113,931,215.55**	**9.58**	**38.57**
August	99191	3673.74	6010	957.51	62,987,313.26	9.59	38.57
September	60403	3775.19	6435	1634.37	50,943,902.29	9.58	37.92
Total	**159594**	**3711.49**	**6435**	**1253.80**	**113,931,215.55**	**9.58**	**38.57**

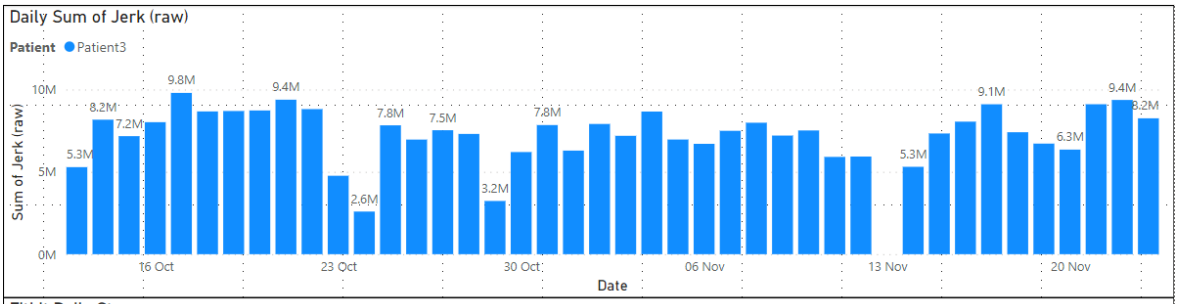
Figure 35: Daily Step count and Jerk (Patient 1)

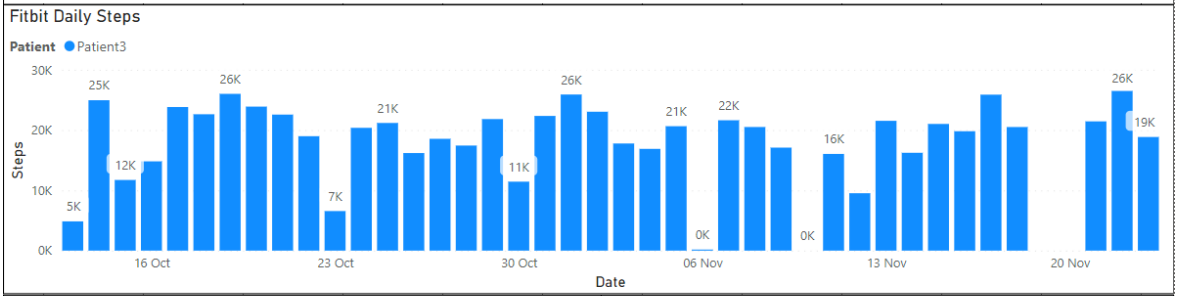




Patient	Sum of Steps	Average of Steps	Max of Steps	Standard deviation of Steps	Sum of Sum of Jerk (raw)	Average of Average of Jerk (raw)	Max of Max of Jerk (raw)
Patient2	**1016056**	**11044.09**	**15895**	**3052.75**	**240,268,575.45**	**9.68**	**66.61**
September	287961	11518.44	15895	2986.16	25,289,851.49	9.73	48.08
October	348400	11238.71	15462	2733.74	106,911,316.24	9.69	53.51
November	323306	10776.87	15098	2962.02	86,764,007.12	9.64	66.61
December	56389	9398.17	12622	4365.91	21,303,400.60	9.68	55.27
Total	**1016056**	**11044.09**	**15895**	**3052.75**	**240,268,575.45**	**9.68**	**66.61**

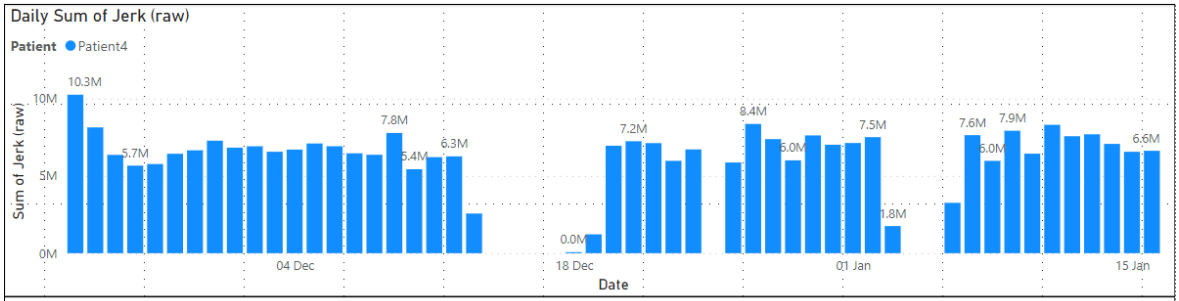
Figure 36: Daily Step count and Jerk (Patient 2)

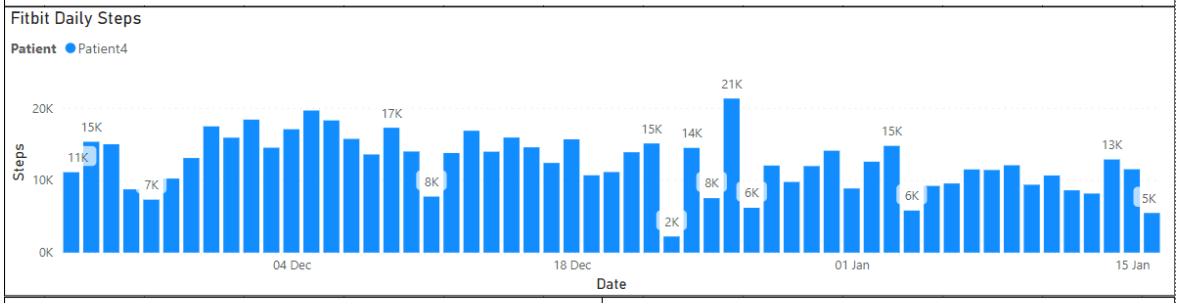




Patient	Sum of Steps	Average of Steps	Max of Steps	Standard deviation of Steps	Sum of Sum of Jerk (raw)	Average of Average of Jerk (raw)	Max of Max of Jerk (raw)
Patient3	**730979**	**17404.26**	**26490**	**7441.84**	**299,612,428.33**	**9.69**	**60.26**
October	350130	18427.89	26009	5940.34	136,925,106.73	9.69	59.79
November	380849	16558.65	26490	8390.43	162,687,321.59	9.69	60.26
Total	**730979**	**17404.26**	**26490**	**7441.84**	**299,612,428.33**	**9.69**	**60.26**

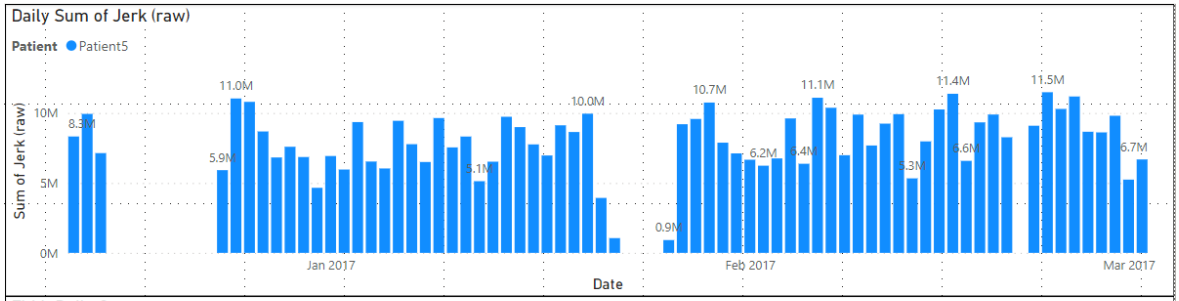
Figure 37: Daily Step count and Jerk (Patient 3)

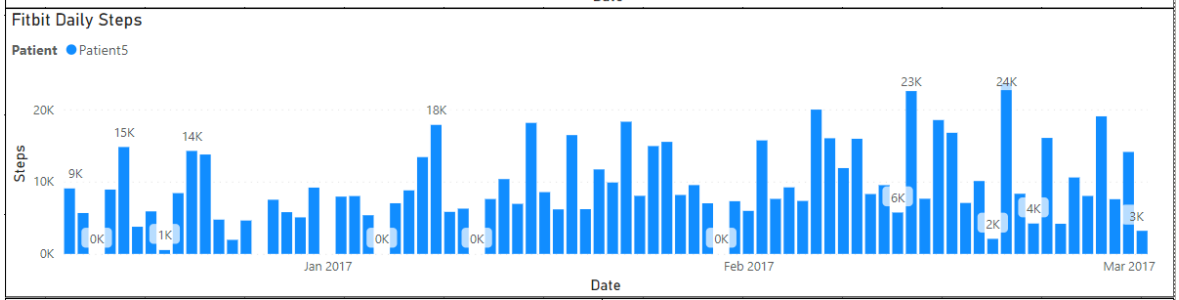




Patient	Sum of Steps	Average of Steps	Max of Steps	Standard deviation of Steps	Sum of Sum of Jerk (raw)	Average of Average of Jerk (raw)	Max of Max of Jerk (raw)
Patient4	**686342**	**12478.95**	**21305**	**3831.69**	**307,692,490.71**	**9.67**	**66.61**
January	162518	10157.38	14759	2425.95	91,476,332.88	9.67	53.66
November	98270	12283.75	17457	3289.12	56,570,647.84	9.67	55.35
December	425554	13727.55	21305	3983.52	159,645,509.99	9.67	66.61
Total	**686342**	**12478.95**	**21305**	**3831.69**	**307,692,490.71**	**9.67**	**66.61**

Figure 38: Daily Step count and Jerk (Patient 4)





Patient	Sum of Steps	Average of Steps	Max of Steps	Standard deviation of Steps	Sum of Sum of Jerk (raw)	Average of Average of Jerk (raw)	Max of Max of Jerk (raw)
Patient5	**731737**	**9146.71**	**23653**	**5529.80**	**545,900,304.78**	**9.73**	**61.65**
January	271123	8745.90	18297	5082.93	207,264,574.06	9.73	61.65
February	319473	11409.75	23653	5797.19	238,966,447.12	9.72	58.72
March	17309	8654.50	14114	5459.50	11,919,847.27	9.77	49.18
December	123832	6517.47	14796	4380.29	87,749,436.33	9.73	54.10
Total	**731737**	**9146.71**	**23653**	**5529.80**	**545,900,304.78**	**9.73**	**61.65**

Figure 39: Daily Step count and Jerk (Patient 5)

4.5 Discussion

4.5.1 Patient´s perception

Patient's perceptions recorded during the PPI session and during the recruitment period while using the app allow us to conclude that using smartphones appears to be a popular proposition, and developing apps could support a broad range of features in categories such as social media, games, or productivity. User retention in using an app relies on the ability to respond to a need and requires careful consideration on the way that the phone is used. Therefore, the scope of the framework used to capture data is constantly evaluated by tracking patients and clinicians' feedback. The features of RApp™ have been implemented and reviewed considering users' perspectives (patients and clinicians) but prioritising the purpose of data collection to define the components implemented. Features not functional (e.g., social media and UI personalisation) for this project were dismissed but could be added for a commercial version.

Patient input recorded during face-to-face interviews at the beginning, middle and end of the recruitment period helped to drive the dialogues. The questionnaire used (see Appendix D) and comments can be found in Appendix E. This feedback provides insights not only on the front end but also on the usability by rheumatoid arthritis patients. The layout of RApp™ should improve, but no functional issues, crashes, or app freeze (e.g., "app not responding") were reported, which is perceived as good app responsiveness. The background monitoring of activity received praise, which is a positive sign that patients could be willing to use RApp™ outside of the study context. Patients in remission do not necessarily want to think about their disease activity, and an app should not give rise to the feeling that it controls the patient's life. The app is a way to observe patients' behaviour but should not influence or motivate them. Therefore, data capture should be seamless to participants and available for remote monitoring. Still, reporting, incentives, reminders (e.g., medication) and the addition of instructions (e.g., daily exercises) are not included in the current version of the application.

The range of scores (DAS and questionnaire) observed across the group varies from moderate to high disease activity, but the individual trends are similar. A patient might be used to a certain amount of pain, with the perception attributed to the patient's awareness but it is also based on the resistance built over time. While the consistency of the scores does not provide significant insights, RApp™ allows tracking of specific joints and areas of concern. The correlation between the scores resulting from questionnaires and DAS highlights the reliability of patient self-assessment in evaluating significant changes in their disease activity. The lack of substantial

change in pain and disease activity does not allow obtaining enough evidence to draw a conclusive correlation against the physical activity recorded.

At the time of recruitment, all participants were diagnosed with rheumatoid arthritis for more than two years and were comfortable in filling the questionnaires provided and joint self-assessments. However, each patient highlighted that filling out an assessment daily could only be helpful for newly diagnosed patients and is too frequent when the disease activity is stable. The BSR questionnaire being completed daily allows us to monitor the patient closely but does not add much value as no scoring can be attributed to it. The RAPID3 questionnaire filled weekly provides a better representation as it allows for a validated score, but the wording of the questions is confusing and not designed for weekly usage. The user interface to record pain and swelling is helpful but does not include enough joints to calculate a validated DAS28 score. Identifying the joints from the left or right is not trivial but assigning an intensity is a valuable feature to quantify the values reported.

4.5.2 Continuous monitoring

The file corruptions and inconsistency in the number of samples generated indicate that both the window and interval are affected, implying issues with the smartphone's service logic and timers. Smartphone’s services also appear to stop running with the phone low in battery, not in charge, restarting or following updates. The activity detected by the phone triggers internal mechanisms such as power save mode and resources prioritisation, and smartphones used by patients as secondary devices most likely result in extended periods of inactivity, disrupting the service schedule. Patient 003 used the smartphone as the primary device, which resulted in significantly more samples recorded. It should be noted that if used as the primary device, inactivity recorded could be interpreted as a period of physical inactivity of the patient rather than issues in sampling. There has been no sign of corruption in storing the user self-assessments in SQL databases. Azure uploading once daily could take up to an hour, but the files uploaded are identical to those extracted from the phone’s storage. The number of logs collected confirms that the smartphones can record a trend of physical activity. Still, the number of samples collected is less than half of the amount expected, which could be due to issues in writing text files or specific to file management and phone storage.

The Fitbit devices are continuously on the user´s body, worn at the wrist, which provides an efficient way to record a reliable fitness baseline. However, Fitbit´s recommendation to perform 10k steps per day did not apply to all participants recruited. Specifically, 10k daily steps at least twice the daily amount for patient 1 while being much lower than the amount of daily steps for

patient 3. The load rate values cannot easily correlate with step count, which raised challenges in the protocol. Besides, the smartphone is a secondary phone, which is significantly different in body location. A number of wearable devices for data collection and connected care is available (Åkerberg et al., 2016) but at the time of this study, raw accelerometer data recorded from Fitbit were not accessible. So load rate values at the wrist could not be calculated, which prevents side-by-side comparison and no common baseline between the smartphone and Fitbit. The only unit of measure available is the step count estimated through Fitbit algorithms which are not publicly disclosed (Fitbit, 2018). Furthermore, Fitbit only provides daily statistics, and consequently, the load rate recorded by the smartphone had to be converted to daily summaries, preventing data analysis into hourly, or peak-based activity timeframes.

Passive monitoring and ambient processing are valid concepts. Still, they require considerations of the data sampling, the timeframe of sampling (i.e., days, months, years), and the way the device is handled (i.e., in hand, pocket, handbag, desk or not with the user) and generally used (i.e., as primary, or secondary). Besides, changes in a patient s' behaviour and habits might incur following significant medication changes. Still, it is not possible to correlate physical activity with disease activity without significant variations of joint pain or inflammation, and multiple factors such as ad-hoc illnesses, weather, seasons, work, and holidays might affect the self-assessments reported by patients.

4.6 Conclusion

In this study, we see that using smartphones to continuously monitor physical and disease activity is a reasonable and technically feasible approach. The user interface and type of questionnaires could be further optimised and tailored to specific disease conditions (e.g., back pain), but the overall feedback and reactions from all participants were positive and demonstrated that a cohort of rheumatoid arthritis patients is comfortable using smartphones, as part of their day to day. Nonetheless, patients do not want to be burdened with continuously filling out questionnaires or being forced to interact with their phones. Fitbit shows that it is possible to monitor physical activity passively but requires being worn on the wrist and at the time of this study is not able to quantify load forces and is limited to step count estimated daily via proprietary algorithms.

This study shows that using smartphones represents a reasonable approach to patients' continuous monitoring of physical activity and load forces. The amount of self-reporting convenient for patients was introduced and some of the challenges associated to passive monitoring were raised. The subsequent studies have been designed to further assess these points. Study 2 addresses the technical challenges and introduces a comparison with load rate

values obtained from smartwatches. Study 3 introduces step count obtained from smartphone to compare the trends of physical activity obtained from Fitbit devices. Study 4 evaluates a smartphone's ability to continuously monitor physical activity through a user's primary phone (as opposed to secondary), in real-life conditions and over an extended timeframe.

Chapter 5 OApp Southampton: Continuous monitoring with power saving disabled

This chapter covers the steps undertaken to compare the load force estimates between smartphone and smartwatch, when disabling power saving.

5.1 Introduction

As described in the previous chapter, monitoring physical activity without the patient's intervention is possible. Still, continuous monitoring requires a lot of storage space which cannot scale to an extended timeframe. The performance and tracking accuracy also rely on internal algorithms (i.e., power saving), so we need to consider sample-based estimates and the smartphone's processing capacity. Moreover, Fitbit's step count data is aggregated daily, preventing side-by-side comparison of smartphone's load rates with the wrist-worn device. So, for this study, we compare estimates of load rates from smartphones and smartwatches. RA patients were recruited as a sample of convenience for study 1, but the methodology is generic. So, for this study, participants were not required to be diagnosed with any specific conditions.

5.2 Literature review

5.2.1 Smartphone´s battery optimisation

Smartphones offer processing capabilities like a microcomputer, but with a significant challenge for all manufacturers: the battery drain. The smartphone's battery supports lighting of the screen, support of protocol on the cellular network (e.g., GSM, GPRS, UMTS, LTE), additional wireless and radio stacks (i.e., GPS, WI-FI...) and applications with their background activities (e.g. OS tasks, email push and notifications). Qualcomm is one of the major smartphone chipset manufacturers. It allows battery optimisations at the hardware level for their Snapdragon range (see fig 40). It releases further optimisation at the application level with "Qualcomm GURU", which learns the user habits to trigger and free resources. Handheld manufacturers aim to develop slicker designs with smaller and thinner handsets while increasing the power requirements. So, besides optimising the battery, shortening the charging times is also an area of focus, releasing technologies such as the proprietary Quick charge from Qualcomm (Qualcomm, 2022).

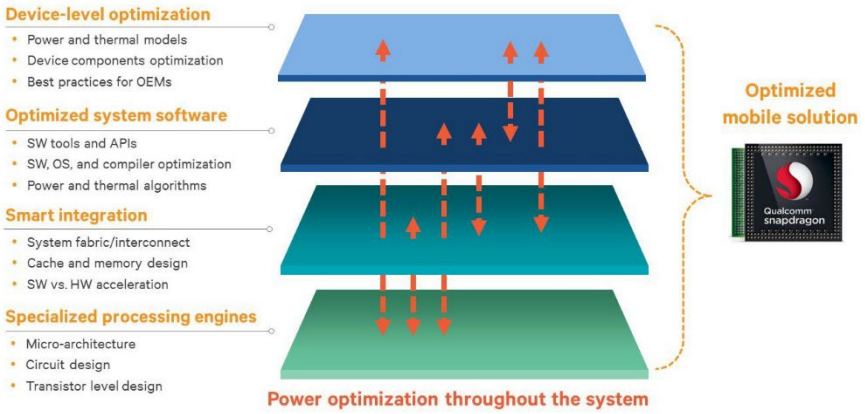


Figure 40: Qualcomm chipset approach to power optimisation (Qualcomm, 2013)

The behaviour of smartphones on the cellular network has been defined following the evolution of radio technology from GSM (2G) to LTE (4G). The mode of operation of smartphones is standardised through the 3rd Generation Partner Project (3GPP), which covers the technological standards and protocols that cellular network providers and manufacturers aim to comply with. These typically use timers defined on the network provider's SIM card to ensure that the connected mobiles use only the network resources allocated. Fig 41(a) shows the smartphone´s location area update (LAU), a protocol for the smartphone to update its location on the cellular network. This operation occurs when registering on the cellular network (i.e., when the phone is turned on) when the user's location changes and periodically as defined by the timer (ETSI, 2016a). Fig 41(b) shows the fast dormancy protocol (ETSI, 2016b), designed to optimise the network resources and the smartphone´s battery consumption. Radio resource control (RRC) is a protocol defined at the network layer to define the network resources allocated as well as the state of the phone. When a phone is inactive (i.e., no active data connection), it is considered by the RRC to be in an idle mode with the lowest battery consumption. The network allocates radio resources when the smartphone needs to support an active data connection (e.g., when the user is browsing). It sets the phone to enter Cell_DCH, a dedicated channel state with the highest battery consumption. Fast dormancy is triggered following the inhibit timer defined by the network to check if the connection is inactive. Network resources are then released, and the state of the phone is set to Cell_FACH (or Cell_PCH), which has a much lower battery consumption than Cell_DCH. If the data connection remains inactive, the phone returns to idle mode.

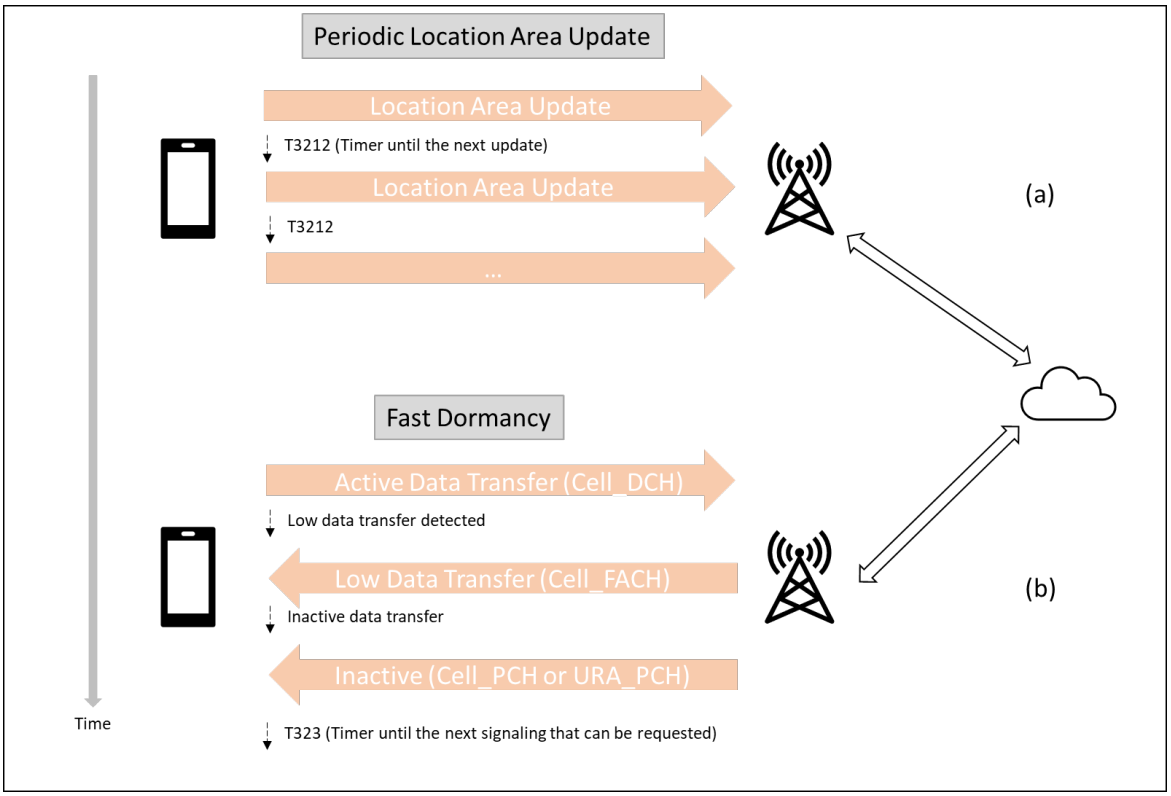


Figure 41: LAU & FD

Timers are also used, e.g., to dim the screen or shut down applications to release resources when inactive for extended periods. Android power management evaluates the user´s usage pattern to free resources affecting the usage of the CPU or battery (Google, 2022b). The CPU resources are allocated based on whether the app is active (e.g., an app that has just been launched) and frequency of usage (frequently or never). Further battery optimisation might also be implemented directly by the manufacturers, such as “Power save mode” for HTC (HTC) or “Stamina Mode” found in Sony’s Xperia range (Sony, 2019). Fig 42 shows the effect of the doze mode for apps on standby (i.e., not actively used by the user). The algorithm considers the battery's state and will not be in effect if the device is on charge. When charging, the device will be regarded as not used by the user if the screen is off and entered in app standby mode, which defers all background activities by up to 24 hours. Doze mode prevents using network and CPU-intensive resources while periodically opening a window of resources (every 9 min or less, depending on manufacturer´s implementation) that allows background operations to be performed (e.g., email, social media, etc..).

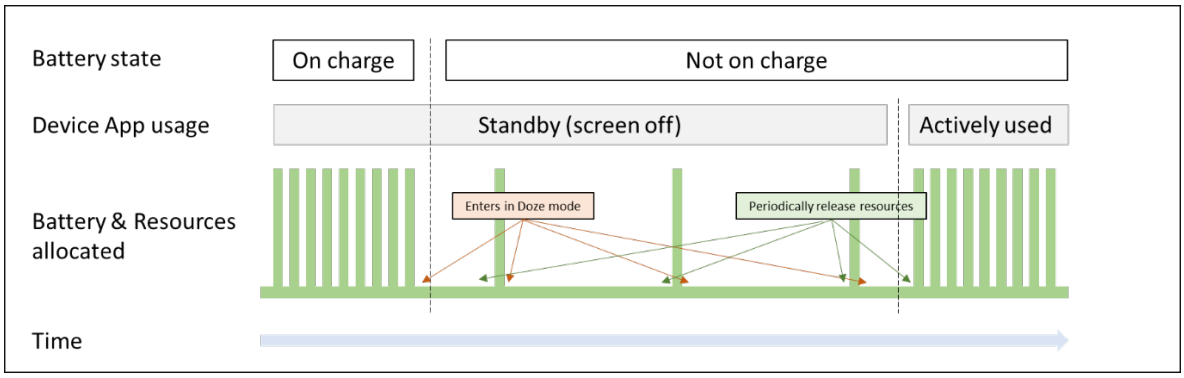


Figure 42: Android Power management

Android power management represents a challenge for any app running background services and so is a challenge for this project that aims at estimating jerks without user intervention. The framework (i.e., RApp™ and OApp™) records raw accelerometer data and uses the storage and CPU of the phone to process data. Therefore, considerations should be made on the phone´s data sampling constraints to estimate load rates.

5.2.2 Smartwatch & tracker position

Load rate estimates (LRE) are affected by power saving, the device handling, and the sensors' location on the user. The way that smartphones are handled affects the data sampling and is specific to each individual day-to-day usage. Moreover, smartphones are typically held in multiple locations on the body. A survey and pilot study on women´s risk perception carrying their smartphones (Redmayne, 2017) showed that 96% kept their phones on standby throughout the day. The device location changed based on usage, characterised as passive, active (e.g., on call) and asleep. When passively used, the device was noticeably kept in hand, skirt/pocket or against the breast for 86% of the 197 participants (women aged 15-40).

The impact of the smartphone location on the body has been reviewed in a study comparing the step count recorded from an Android smartphone (Galaxy Nexus GT-I9250) to the video recording of 27 participants over a straight-line route at a different speed (Brajdic and Harle, 2013). According to this study, step count estimates undercount when worn in the back trouser pocket and overcount when carried in a handbag but comparable in all other positions such as being held by a hand, by hand with interaction (e.g., typing a message) or on the front trouser pocket. This study also highlights that none of the step count algorithms was 100% reliable and recommends using probabilistic methods.

Accelerometer-based systems have been reviewed and can be integrated for unsupervised monitoring of free-living subjects (Mathie et al., 2004). The accelerometric nature of the human

body´s movement considers the amplitude and frequency range. These can be used to estimate metabolic energy expenditure, physical activity, postural sway, gait, fall detection, and postural orientation and activity classification. However, calculating load rates raises multiple challenges when using accelerometer data in healthcare. This has been discussed in a study with data collected from 49 participants (with ActiGraph GT3x+) to compare activities performed in a lab and free-living conditions for seven days (Karas et al., 2019). Noticeably, the study highlights the impact of the data size and sampling frequency while recommending storing raw and processed data, confirming the findings from study 1. The smartphone placement in free usage is dynamic and potentially kept away from the body, which furthers the complexity. Still, a study validated that smartphones can be used without constraints to estimate step count using a Fast Fourier-based algorithm (Kang et al., 2018).

As of this project, continuous load rate estimate has not been explored in free-living conditions. A benchmark of reference should be defined to validate smartphone estimates, and wearable devices offer possibilities as they have been used in other studies. A systematic review of the literature published between 2013 and 2017 identified key research themes (Shin et al., 2019). The wearable used in the various studies included wrist-worn devices such as Actigraph, Misfit, Fitbit, Nike+, Fuelband and Jawbone. Most studies focused on the technical aspects of PA or incorporating these devices in medical settings, e.g., for monitoring. The other prominent themes were wearable technology adoption, behaviour change, and self-assessment.

In recent years, wearable devices have become more readily available to the public. A study has shown a solid correlation to measure steps with research-grade devices using BodyMedia SenseWear, and ActiGraph GT3X+ as references. The consumer devices included the Fitbit wristband, Jawbone UP, Misfit Shine, Nike Fuelband, Striiv Smart Pedometer and Withings Pulse (Ferguson et al., 2015). The study recruited 21 participants and was limited to 2 days of sampling, but it has shown the performance of wristband wearables to measure PA and step count in free-living conditions. A systematic review further gathered the validity for using consumer-level devices to estimate steps and PA. The studies reviewed found high reliability in measuring step count and generally reliable between Jawbone and Fitbit devices (Evenson et al., 2015).

5.2.3 Rationale

Smartphone handling affects sensors' recording ability, so power save mode is disabled to uniform sensor sampling. Smartphones are typically handled dynamically and not always worn or carried in the same positions, so we compare them with the load rate obtained at a static location

on the body (at the wrist). Load rates are not a standardised unit of measure, so we compare estimates using raw accelerometer data obtained from smartphones and smartwatches.

5.3 Aim & objectives

5.3.1 Aim

This study aims to assess whether smartphones are as reliable as smartwatches to evaluate participant load forces trends.

5.3.2 Objectives

The objectives of this study are:

1) Ascertain the ability of a smartphone to capture accelerometer data continuously
2) Ascertain the smartphone´s processing capability of raw accelerometer data to estimate load rates
3) Compare the load forces estimated from smartphones and smartwatches to assess the impact of the device location on the body

5.3.3 Hypothesis

If smartphone's sampling and processing of accelerometer data is manipulated to be uniform and unbiased, load forces estimated are comparable to smartwatch's estimate.

5.4 Methodology

5.4.1 Patient recruitment and study procedure

This study is observational and quantitative, focusing on accelerometer data recorded via smartphones and smartwatches. Each participant maintained a physical activity diary as a baseline for activity classification. Each participant wore a Sony Smartwatch 3 and used a Sony Xperia Z5 Compact, both preloaded with OApp™.

The data collected for this study have been collected as part of a project approved by the University of Southampton's Faculty of Engineering and Environment Ethics Committee (ERGO Ethics ID:30213) and presented in another PhD thesis (Nazirizadeh, 2018).

Monitoring of disease activity is not part of the scope of this study, and to be eligible, inclusion criteria were defined to recruit participants willing to participate in the study, above 18 years old and without neurological, systemic illnesses or other physical disabilities that may have limited their mobility. Participants were ten healthy adults (female n=6, male n=4; aged 27.2 ± 3.6 years; height: 172.6 ± 9.6cm; body mass: 73 ± 14.7kg; means± standard deviation) recruited via posters on multiple noticeboards around the University of Southampton.

Participants interested in participating in the study signed a consent form following an introduction to the app. Once the consent form was signed, participants were provided with a patient information sheet, smartphone, and smartwatch for seven days.

5.4.2 Data collection

Participants were asked to wear the smartwatch on the non-dominant wrist and keep the smartphone as close to the hip as possible, e.g. in a pocket or using a phone belt provided. Participants did not have to perform specific activities. Still, they were required to maintain a paper-based diary to record moderate to vigorous physical activity times and report instances of devices running out of battery.

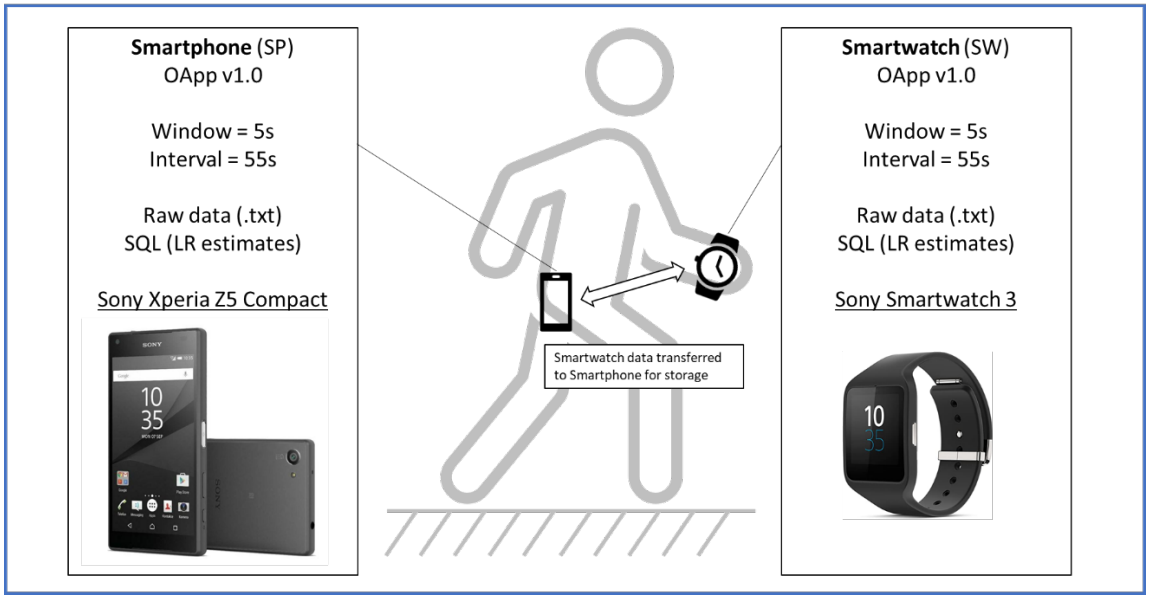


Figure 43: Smartphone & Smartwatch

For this study, the monitoring has been set up to record data samples of 5 seconds (window=5) every 55 seconds (interval=55) to obtain a load value calculated per minute. The samples have been stored on the internal storage of the phone in raw format on text files (as per Study 1). Calculated load rates have been stored in segregated SQL databases and storage folders to avoid

conflicts to create, read, update, and delete (CRUD). The samples were extracted manually from the smartphone at the end of the recruitment period of seven days.

5.4.3 Data analysis

For each participant, we evaluate the performance and accuracy of the app logging in the context of continuous monitoring for the recruitment duration (7 days). The hourly amount of samples collected is compared to the amount expected to be collected, which considering the settings used, aims at obtaining one load rate value per min (i.e. one file per min, 60 files per hour, 1440 files per day). Periods without samples recorded indicate issues but recorded null values suggest that the device is in a static position (e.g., on a table). The integrity of each sample is assessed by comparing the duration collected to the window set (i.e., each sample should contain ~5sec of accelerometer data). A significantly wider window indicates multiple samples are considered corrupted samples as there is no way to weigh each sub-sample.

The smartphone on-board data processing capability is evaluated by comparing the load rates calculated and stored in SQL in-app to those calculated from the raw accelerometer data stored as text files and processed in MATLAB. The amount of physical activity performed depends on each participant and has already been identified in the diaries (Nazirizadeh, 2018). Still, the load rate is also calculated for the smartwatch comparison of trends recorded between devices worn at the wrist and smartphone, using the same unit of measure.

5.4.4 Individual phenotype

Load rate estimates (LRE) are measured using the jerks generated by the accelerometer over time. Continuous plots of raw accelerometer data over time are confusing and unwieldy. So, we define a methodology to extract meaningful and concise characteristics to evaluate an individual´s data. Drawing physical activity characteristics of an individual requires defining features of interest. This group of components can be referred to as individual “phenotype”. Fig 44 illustrates the proportion of time (i.e., the amount of time in percentage) spent at load forces applied represented by LRE (in kg m/s^3) for two individuals. The plot´s left side shows the greatest amount of time spent at the lowest LRE which are periods of low physical activity. The plot´s right side shows the lowest amount of time spent at higher LRE which are periods of high physical activity. The amount of time logged represents the amount of samples recorded. In this context, we propose to use a continuous monitoring score (CMS) which is derived from the correlation formula of r^2 applied to LRE and proportion of time. The CMS gives confidence on the phenotype being reliable, not that data is reliable. A CMS equals 1 means that all LRE variations can be

explained by the variations in proportion of time, which means that all the samples within the period can be explained and correctly captured. So, high CMS indicates that the samples were captured at the expected interval (i.e., 5 sec samples every 55 sec), and the slope gives a good representation. Low CMS means that the samples captured cannot all be explained, so the slope is not a reliable representation of the level of physical activity. The peak is the LRE recorded for the highest proportion of time. The slope of the linear regression indicates the trend of physical activity (measured in kg m/s^3). For instance, a slope of -1 indicates that an individual is more active overall than a slope of -1.5. Intersections between the linear regression and the curve might indicate regular activity (e.g., sitting, running, and walking). However, activity classification is out of the scope of this project.

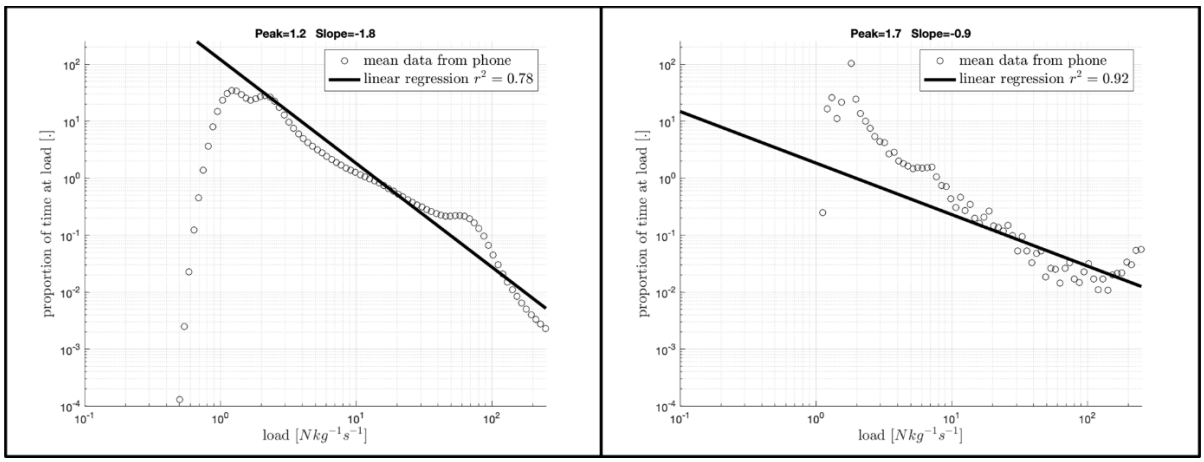


Figure 44: Example of individual phenotypes

5.5 Results

5.5.1 Continuous monitoring

The cohort of 10 participants completed the recruitment period of seven days, and raw and processed data were collected for the smartphone and smartwatches of all patients. Table 7 shows the total amount of readings collected from smartphones (SP) and smartwatches (SW) of participants from 1 to 10.

Participants 1, 7, 8 and 10 generated the most smartphone readings, respectively, with 3.55, 3.01, 2.94 and 3.56 million lines. All other participants generated over 2 million lines except participant 2, which did not record data after the third day. The average number of readings per day is around 300 thousand for all participants, ranging from 262.9 to 395.6 thousand lines for participants 6 and 10. The number of readings is dispersed across the day, with standard

variations ranging from 72 to 125 thousand. On their highest days, participants peaked at around 400 thousand readings.

Participants 2, 7 and 10 generated the most smartwatch readings, respectively, with 3.12, 3.24 and 3.37 million lines. All other participants generated over 2 million lines except participant 8, who did not record smartwatch data on the last day. The average amount of readings per day, as for smartphones, is around 300 thousand for all participants but ranges from 174 to 375 thousand lines for participants 1 and 10. The number of readings is dispersed across the day, with standard variations ranging from 71 to 128 thousand. On their highest days, participants peaked at over 400 thousand readings.

Patient		Recruitment		Amount of Readings (per day)				
ID		Start	End	Total	Average	Min	Max	Std Dev
1	SP	14/08/2017	23/08/2017	3551704	355170.4	111195	506067	125100.6
	SW	16/08/2017	23/08/2017	1218298	174042.6	33062	419287	116060.8
2	SP	15/08/2017	17/08/2017	802634	267544.7	121935	374453	106660.5
	SW	15/08/2017	23/08/2017	3120699	346744.3	192521	422348	94874.6
3	SP	15/08/2017	23/08/2017	2680747	297860.8	104417	418143	90596.42
	SW	15/08/2017	23/08/2017	2534310	281590	137026	426364	91551.94
4	SP	16/08/2017	24/08/2017	2559999	284444.3	153779	379456	88800.38
	SW	16/08/2017	24/08/2017	2000224	222247.1	33660	408746	128685.4
5	SP	15/08/2017	23/08/2017	2578417	286490.8	98447	343736	81175.51
	SW	15/08/2017	23/08/2017	2595041	288337.9	121134	407239	91230.67
6	SP	23/08/2017	31/08/2017	2366498	262944.2	140728	347059	75158.06
	SW	23/08/2017	31/08/2017	2893154	321461.6	207935	429423	83410.82
7	SP	23/08/2017	31/08/2017	3010061	334451.2	113825	471059	89954.76
	SW	23/08/2017	31/08/2017	3247407	360823	138275	428373	103337.4
8	SP	24/08/2017	31/08/2017	2945113	368139.1	192611	444117	72530.45
	SW	24/08/2017	30/08/2017	1873400	267628.6	157656	406521	90786.39
9	SP	25/08/2017	01/09/2017	2229569	278696.1	22711	371446	109016.9
	SW	25/08/2017	01/09/2017	2264227	323461	229608	425580	71708.46
10	SP	23/08/2017	31/08/2017	3561119	395679.9	116904	522637	115339.1
	SW	23/08/2017	31/08/2017	3378534	375392.7	124112	427465	92627.65

SP - Smartphone
SW- Smartwatch

Table 7: Lines of sensor readings per day

Table 8 shows the daily amount of .txt files generated to store the lines of smartphone and smartwatch's raw sensor readings. The average of daily smartphone files collected is consistent across participants and ranges from 804 for patient 4 to 1160 for patient 7. For the smartwatch, the daily amount of files collected for patients 1 to 10 averages 553, 1106, 906, 723, 940, 1047, 1162, 850, 1053 and 1196. The number of smartphone files generated was consistently dispersed, with standard deviations from 269 to 410. Similarly, the standard deviations of smartwatch files range from 235 to 368 for patients 1 to 10. The highest amount of files generated daily is around 1300 across participants for smartphones and smartwatches.

Patient ID		Amount of files (per day)				
		Total	Average	Min	Max	Std Dev
1	SP	9683	968.3	334	1321	373.75
	SW	3874	553.43	105	1332	368.53
2	SP	2423	807.67	465	1273	341.08
	SW	9956	1106.22	614	1347	302.4
3	SP	8400	933.33	233	1326	315.04
	SW	8162	906.89	441	1373	294.76
4	SP	7237	804.11	290	1258	351.18
	SW	6509	723.22	109	1332	419.29
5	SP	9981	1109	398	1387	349.85
	SW	8465	940.56	395	1330	297.79
6	SP	9472	1052.44	560	1403	310.51
	SW	9429	1047.67	678	1400	272.08
7	SP	10440	1160	433	1348	299.86
	SW	10458	1162	445	1378	333.13
8	SP	7251	906.38	531	1292	269.88
	SW	5951	850.14	500	1290	288.52
9	SP	7825	978.13	43	1358	410.5
	SW	7375	1053.57	747	1389	235.16
10	SP	10413	1157	381	1310	284.56
	SW	10769	1196.56	396	1362	295.06

Table 8: Amount of files collected per day

Table 9 shows the average amount of readings per file, around 300 for smartphone and smartwatch participants. The max for smartphones is consistently around 300, with peaks between 1290 and 2180 and a significantly lower minimum value, indicating that several files were generated with substantially less data than expected. The standard deviation for smartwatches is under 5 for all and between 35 and 273 for smartphones, indicating fewer inconsistencies in the files generated from SW than SP. Therefore, the sample data corruptions and inconsistencies observed in the previous study appeared to be lessened.

Patient ID		Amount of Readings (per file)				
		Total	Average	Min	Max	Std Dev
1	SP	3551704	366.8	6	2214	246.39
	SW	1218298	314.48	306	408	2.05
2	SP	802634	331.26	42	1693	204.93
	SW	3120699	313.45	11	324	4.87
3	SP	2680747	319.14	110	1823	192.87
	SW	2534310	310.5	173	316	2.9
4	SP	2559999	353.74	8	2180	242.31
	SW	2000224	307.3	298	607	5.71
5	SP	2578417	258.33	16	2072	84.07
	SW	2595041	306.56	4	312	4.81
6	SP	2366498	249.84	49	862	35.87
	SW	2893154	306.84	297	310	1.07
7	SP	3010061	288.32	26	1374	154.67
	SW	3247407	310.52	164	314	3.58
8	SP	2945113	406.17	65	1894	273.35
	SW	1873400	314.8	221	587	4.19
9	SP	2229569	284.93	5	1290	141.77
	SW	2264227	307.01	192	315	3.63
10	SP	3561119	341.99	9	1539	214.83
	SW	3378534	313.73	243	319	1.61

Table 9: Lines of sensor readings per file

5.5.2 Smartphone & smartwatch comparison

The accelerometer's variations of the smartphone and smartwatches worn over the recruitment period on all axis (x, y and z) for each participant can be seen in Appendix E. Participant 10 (fig 47) captured data continuously for the entire recruitment on both devices. Participant 2 (fig 45) recorded the shortest amount with only a single day of SP data. Periods of missing recording can easily be identified on the plots as periods without data. Most of the other participants managed to capture data correctly whilst recording gaps with at least a day of missing data on either or both devices, except participant 3 (fig 46), where only a few gaps were observed. The periods of sensor variations show device activity, and the flat lines (i.e., constant sensor value) indicate that the device was inactive but still recording. Both types of events can be seen to start and end at the same time between SW & SP devices for all participants.

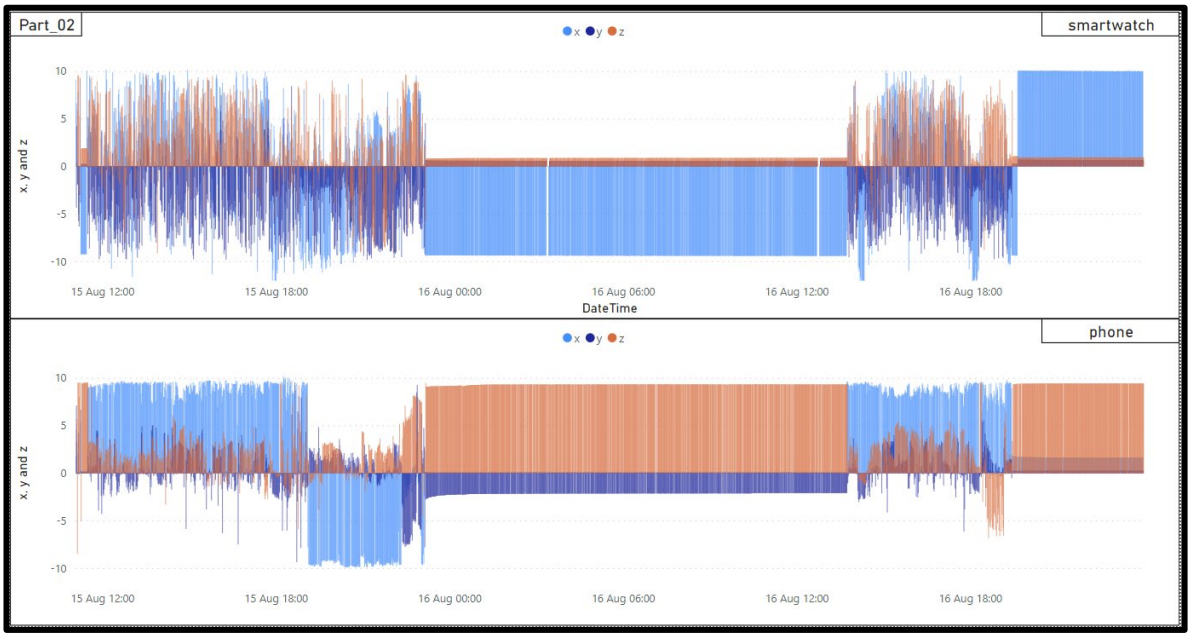


Figure 45: SP & SW raw acceleromter (x, y and z) over time (Participant 2)

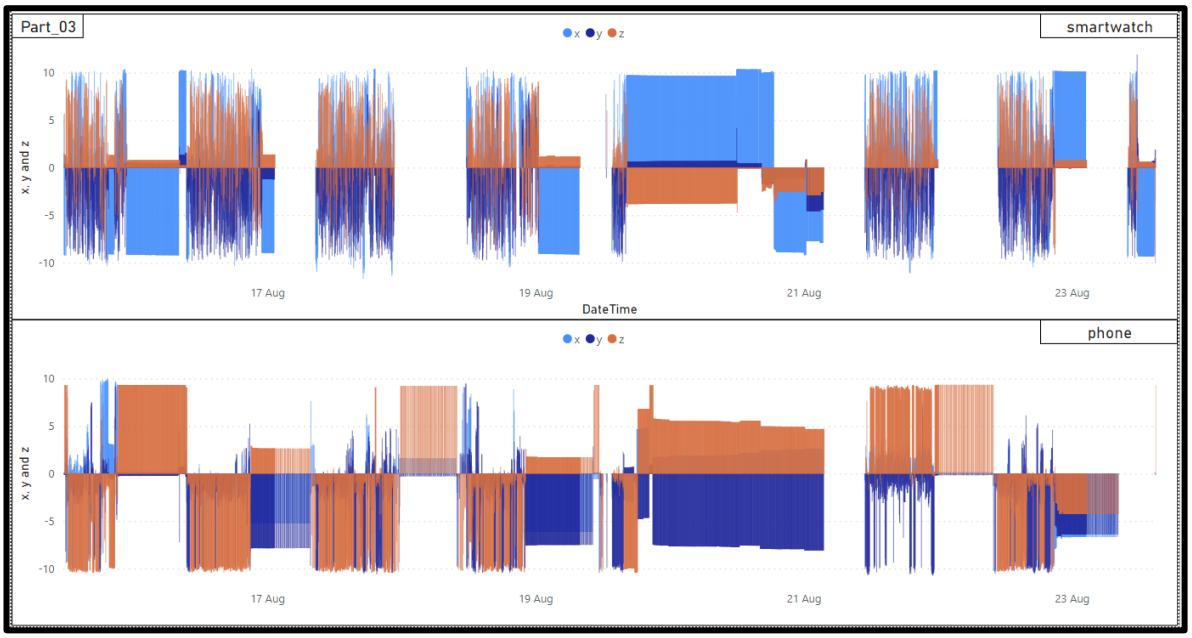


Figure 46: SP & SW raw acceleromter (x, y and z) over time (Participant 3)

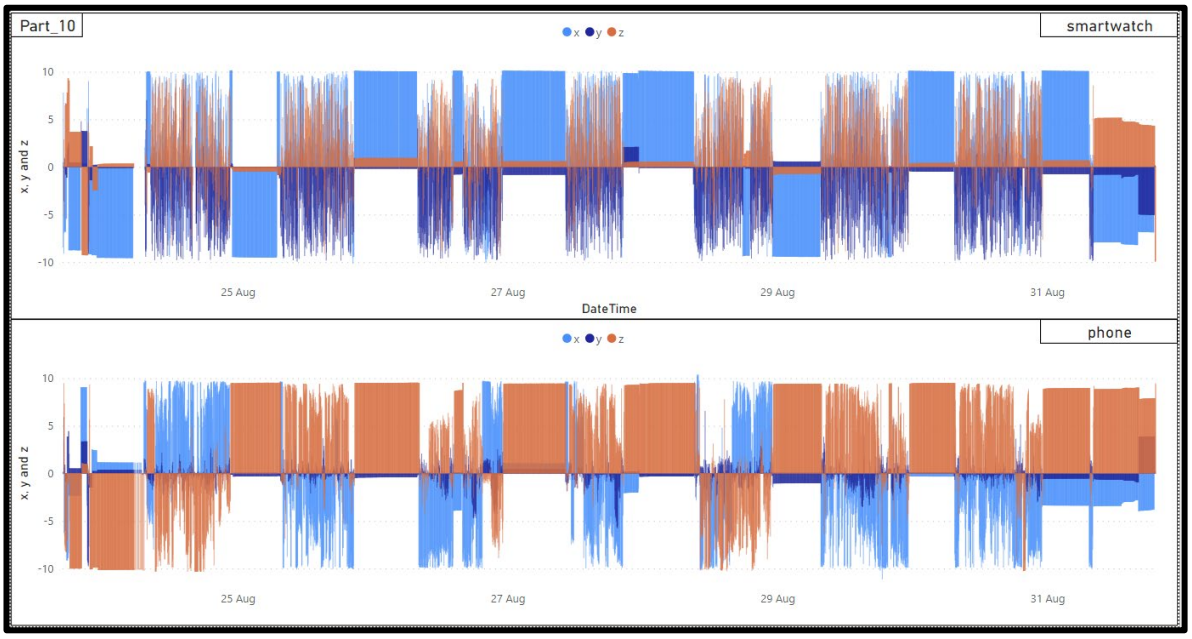


Figure 47: SP & SW raw acceleromter (x, y and z) over time (Participant 10)

As per Android API, raw accelerometer values are recorded between -10 m/s2 and 10 m/s2 for each axis. The plots mean of jerks calculated from each axis over time for both SW and SP can be seen for all participants in Appendix F. Besides showing the periods of inactivity, the value of the load estimation can be seen over time. Participant 2 (fig 48) recorded only a single day but illustrated that spikes of activity can be identified with either device, although reaching higher values on SW (23.45) compared to SP (16.22). Similar spikes can be seen for all patients and consistently getting higher values for SW, over 20, while the highest for SP was observed with participant 7 (fig 49) once at 23.3. The frequency of jerk activity shows that participant 7 was the most frequently active and participant 9 (fig 50) was the least often active.

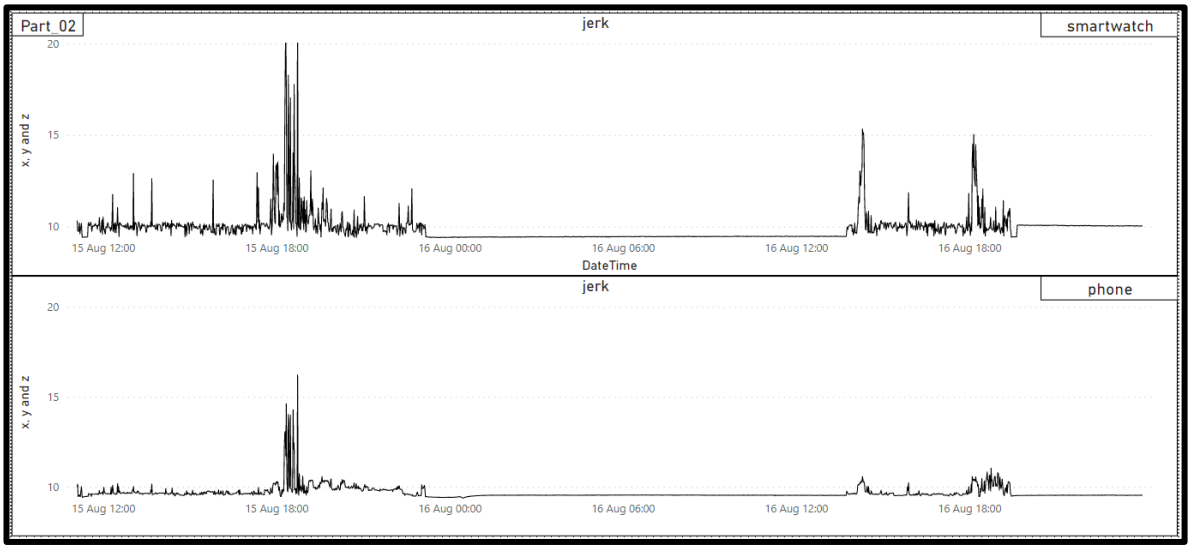


Figure 48: SP & SW jerk (mean) over time (Participant 2)

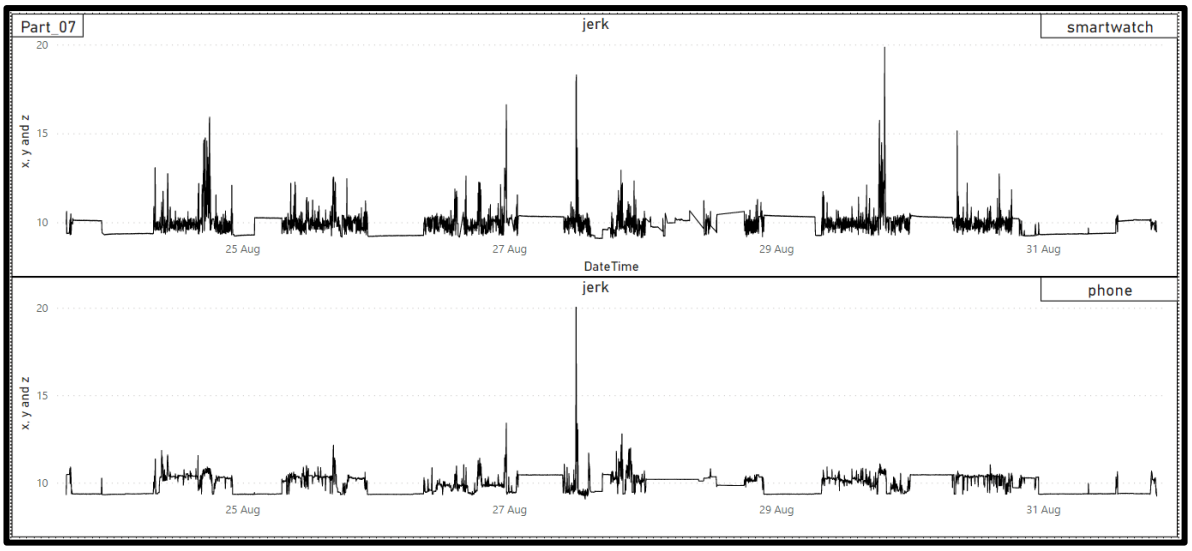


Figure 49: SP & SW jerk (mean) over time (Participant 7)

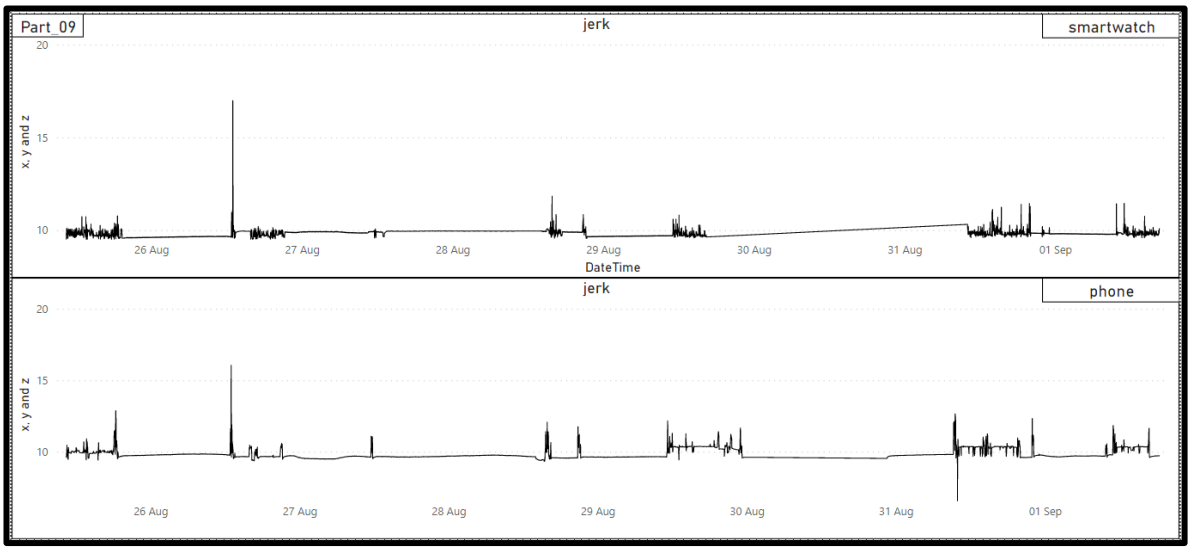


Figure 50: SP & SW jerk (mean) over time (Participant 9)

The phenotypes and characteristics generated for all participants on both SP & SW can be seen for all participants in Appendix G. The CMS permits to verify of the amount of time logged by the device, which is not always perceivable, looking only at the continuous trends. For instance, participant 10 (fig 51) continuously monitored more data on SP (35.6 million lines) with a CMS of 0.95 than on SW (33.7 million lines) with a CMS of 0.77. The linear regression shows a slope of -1.7 on SW compared to -1.6 on SP, indicating more movement recorded on SW than on SP. The average SP peak is above 1.7, which is higher than SW, averaging at 1.2, which means that the most common LRE recorded on SP is higher.

Patient ID		Phenotype		
		Peak (kg m/s3)	Slope	CMS
1	SP	1.8	-1.5	0.89
	SW	1.1	-1.2	0.8
2	SP	1.8	-1.2	0.93
	SW	1.1	-1.5	0.81
3	SP	1.7	-0.9	0.92
	SW	1.2	-1.8	0.78
4	SP	1.8	-1.4	0.92
	SW	1.1	-1.6	0.77
5	SP	1.7	-1.1	0.84
	SW	1.1	-1.3	0.77

Patient ID		Phenotype		
		Peak (kg m/s3)	Slope	CMS
6	SP	1.7	-1.2	0.94
	SW	1.1	-1.4	0.79
7	SP	1.7	-1.2	0.85
	SW	1.2	-1.5	0.83
8	SP	1.8	-1.4	0.92
	SW	1.1	-1.1	0.82
9	SP	1.8	-1	0.87
	SW	1.1	-1.7	0.79
10	SP	1.8	-1.6	0.95
	SW	1.1	-1.7	0.77

Table 10: Phenotype characteristics

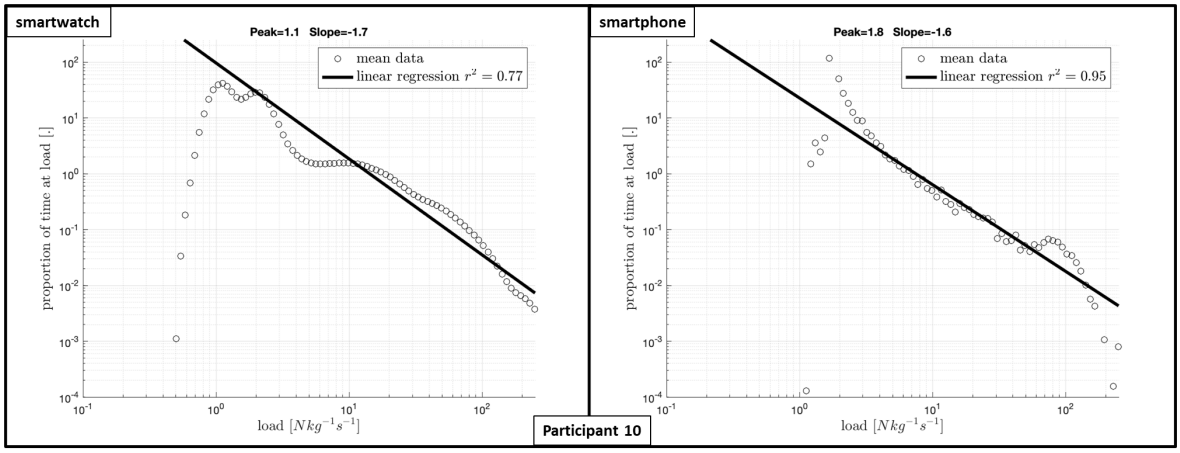


Figure 51: SP & SW Phenotype (Participant 10)

5.6 Discussion

5.6.1 Smartphone monitoring

The sampling for this study was set to capture 5 sec samples of data every 55 sec (equivalent to 5 sec per minute). Yet, more readings were generated compared to study 1, which was set with a window twice larger to capture 10 sec samples every 20 seconds (equivalent to 20 sec per minute). The amount of readings generated per hour is different between patients but consistent across the day, as seen in figure 52. A specific period of inactivity over the day (e.g., at night) might have caused the phone to stop recording data, but the file count, grouped per hour, shows that there is no specific timeframe causing the loss of data. The average amount of files generated per day was around 1000, compared to 225 to 793 for patients from study 1. Each hour, between 50 and 60 data samples were collected for most patients. On most days, no samples have been corrupted, indicating a much higher data collection reliability than for Study 1.

The logging over time is also less scattered, with consistency in the amount of samples generated over time. It is not the same for all patients and ranges between 112 and 432 files per hour, but the trend is consistent without spikes of logging seen. So, the sample recording was not affected by the time of the day or whether the device was on charge. This validates that the impact of the power sample on the interval of samples (every 55 sec here) can be disabled and that the frequency of data sampling can be manipulated to be uniform and unbiased.

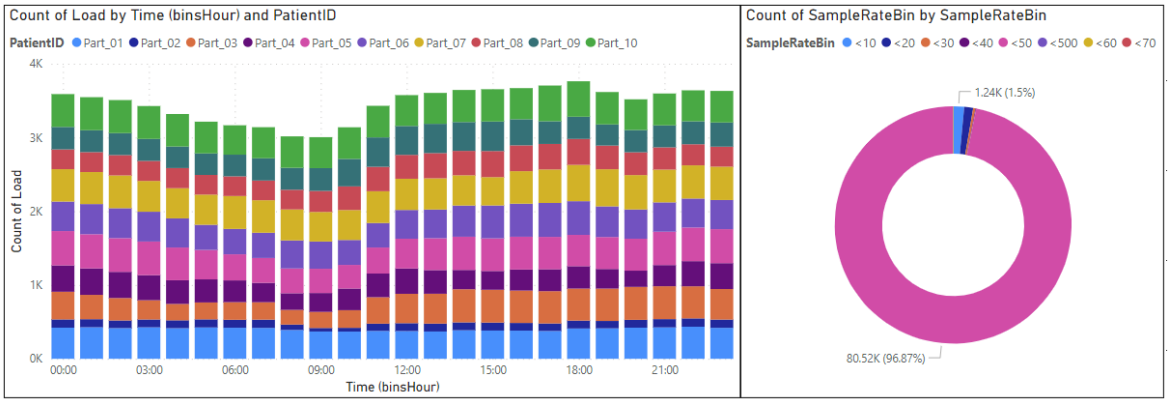


Figure 52: SP sample count and sample rate per hour

Study 1 showed variations in intervals and windows, which can be translated as the count of samples per hour and sample rate for the calculated means. In this study, SP sample rate (<50) when processing the estimates of LR reached 96.87% of the data samples captured. This means that the sample files consistently capture data samples as per the expected window length (of 5 sec here) and validates that the load rates can be estimated using consistent samples. SW sampling is higher than SP due to the mechanism of data sampling by itself, as the samples are transferred from SW to SP via Android APIS. So, SW sample rate (<70) reaches the expected 100%, as seen in fig 53. This sample rate range means that the data processing is reliable and validates that SP onboard processing can be used to estimate load rates.

The estimates stored in SQL match the data calculated for each patient, whether the data were collected through a smartphone or smartwatch. The linear correlation between the samples collected and database entries, neither corrupted, indicates that using SQL database is a reliable way to record and store the data. This approach significantly reduces the amount of data storage and the risks of data corruption related to reading and writing on phone storage. Therefore, we can reliably use the LR estimates stored in SQL rather than raw accelerometer data stored in many txt files.

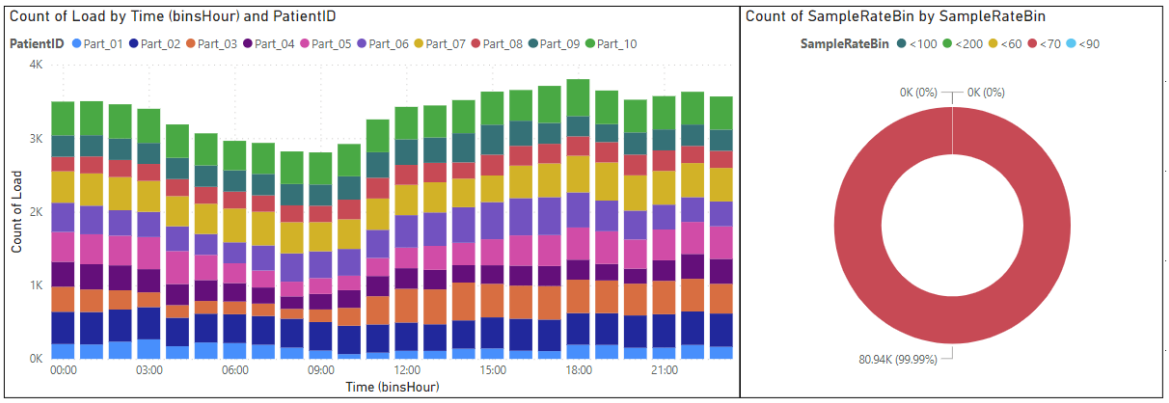


Figure 53: SW sample count and sample rate per hour

5.6.2 Smartphone & smartwatch correlation

When looking at sensor recording on each axis (x, y and z), we can see that the density of sensor events shows more variations on more axis on SW than SP. Fig 54 illustrates an example of daily sensor activity recorded using participant 10 data. The zoom in shows that between 00h00 and 08h00, both SP and SW were at rest (i.e., unused, likely on a table). The constant value recorded is noticeably higher on the x-axis for SW and the z-axis for the SP, corresponding to the force caused by gravity. It is important to note that recordings at constant value denote the inactivity of the device as opposed to the periods of missing data, which indicate missing recordings. After 08h00, we can see that both devices start being used with activity mainly on the x and z axes for the SP while on all axes for the SW.

This trend indicates more movement variations recorded on SW than SP. These movements are expected since arm gestures are generally more frequent than hip gestures and are not associated only to lower limb movements (e.g., arms moving when speaking). Movements associated with arm swing might be beneficial, e.g., to count steps (see Study 3) or for activity classification (Atallah et al., 2011), but introduce noise in the context of jerk estimations. The jerks highlight significant load spikes at 08h30 (likely user waking up) and 18h30 on both devices. The jerks were found to be generally lower on SP than SW, which might be due to the phone being less sensitive and biased to the noise introduced by arm swing or by jolting on the skin (Bouten et al., 1997). It was observed in another study that higher accelerometer readings are found at the waist than chest line (Balogun et al., 1988) (Balogun et al., 1988).

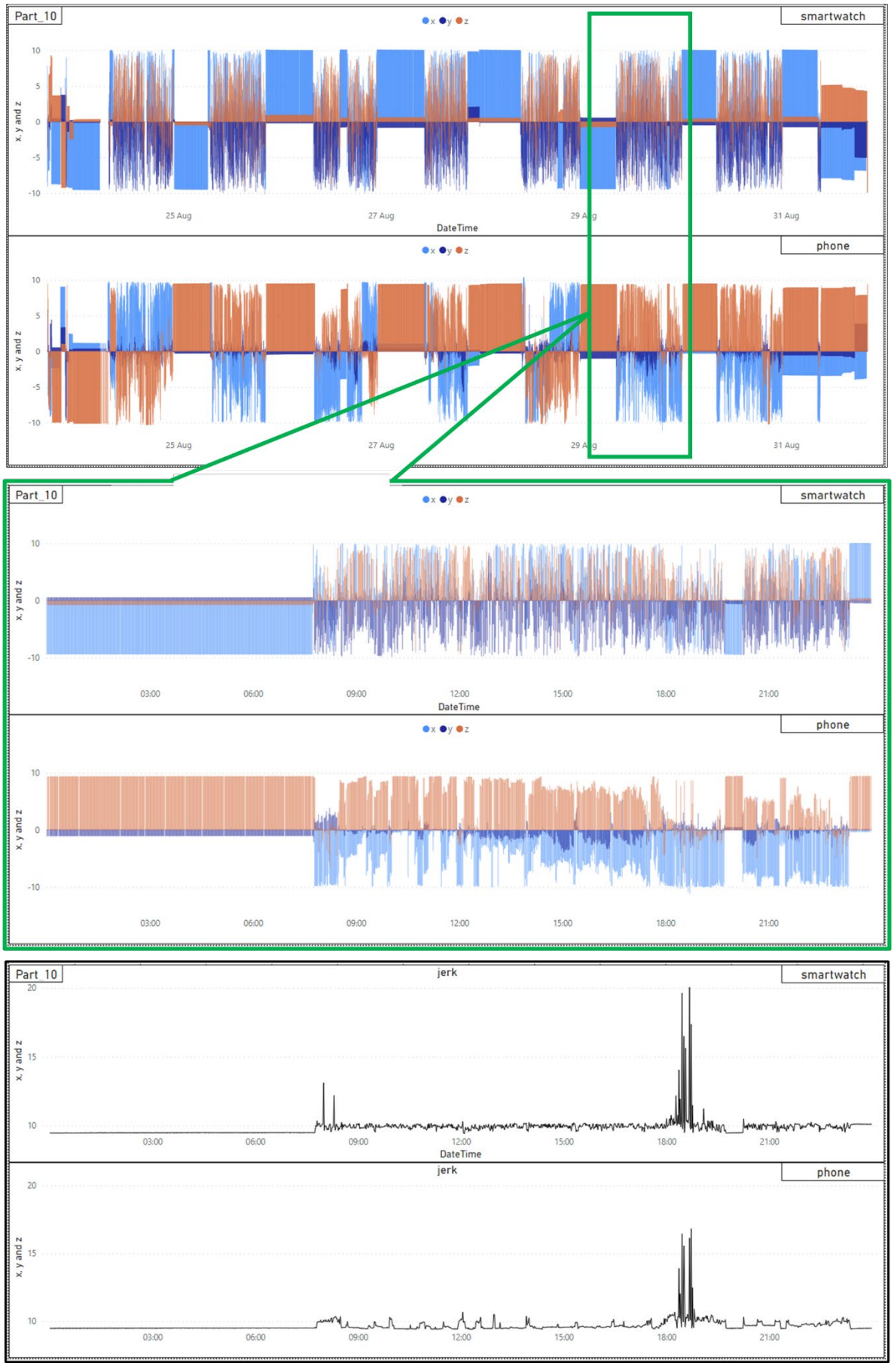


Figure 54: SP & SW raw accelerometer (x, y and z) over time (Participant 10)

Fig 55 shows the relation of load rate mean estimated hourly between SP and SW for all participants, and table 11 lists the Pearson coefficient for each participant. The highest correlation was reached for Participant 2 (0.96), but the amount of samples was limited to a single day. The correlation was above 0.6 for all participants, except the lowest correlation (0.57) for Participant 4. The highest correlations were found for Participants 9 and 10, respectively, reaching 0.83 and 0.86.

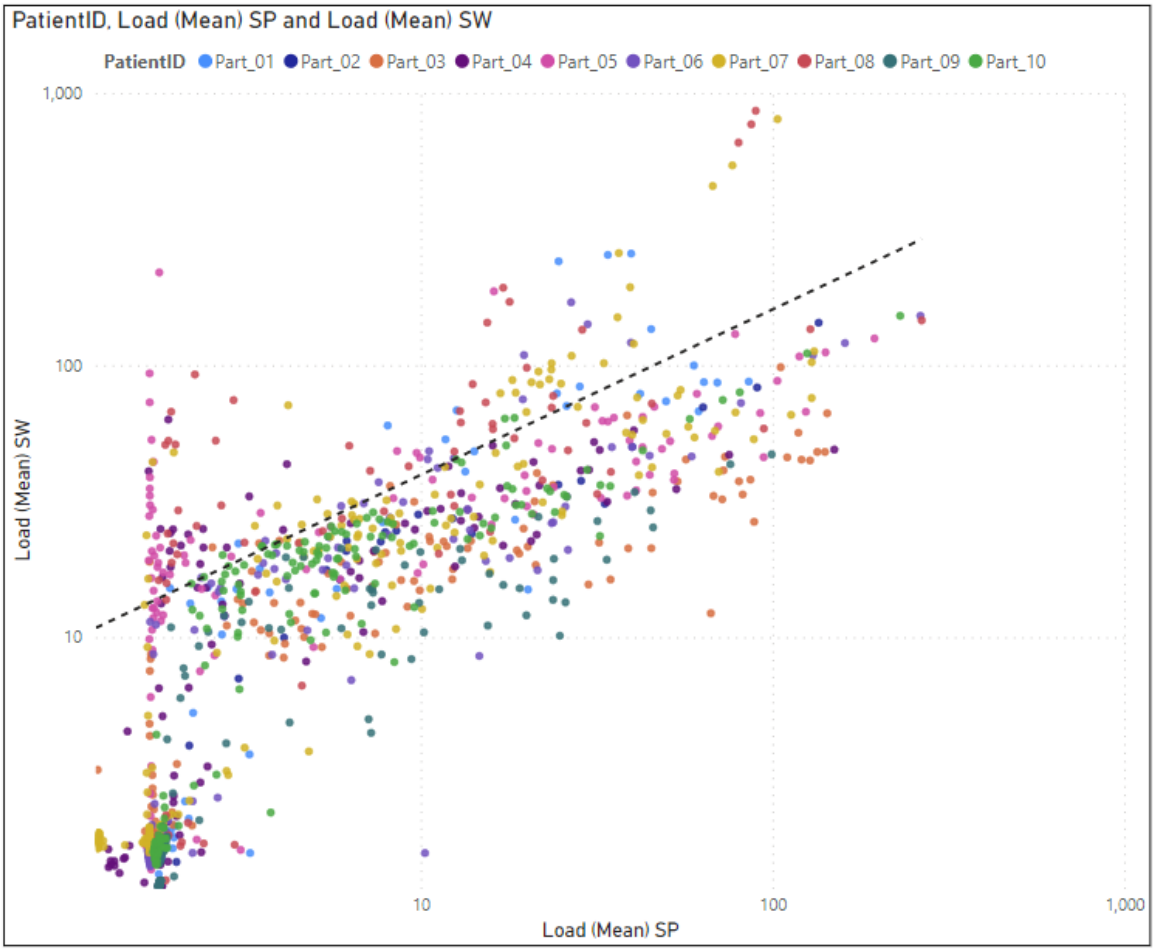


Figure 55: SP&SW LR (mean) correlation

Participant	Correlation	Participant	Correlation
Part_01	0.67	Part_06	0.72
Part_02	0.96	Part_07	0.58
Part_03	0.8	Part_08	0.51
Part_04	0.57	Part_09	0.84
Part_05	0.62	Part_10	0.86

Table 11: SP & SW Correlation

In the context of self-assessments, Smartphones are able to estimate steps as well as traditional pedometer but it also depends on the environment and location of the phone on the body

(Åkerberg et al., 2016). Jerks are estimated from accelerometer values recorded, which are sensitive to the velocity of movements such as arm swing and demonstrated to be higher on the upper limbs (Brambilla et al., 2022). So, it is expected for smartwatches to record higher peaks and values of load rates than smartphones due to their location at the wrist. However, the trends are comparable, which indicates that smartphones can perform as well as a wristband to record physical activity while being worn closer to the body, which is more representative of the load rate on the lower limbs.

Continuous trends can be used to visualize patterns over time, but LRE needs to be quantified to allow comparisons. The phenotypes allow to visualise the amount of LRE over time and provide values that can be compared. The previous study has shown that SP handling affects the data sampling. Here, the CMS provides a representation of the amount of logging recorded providing a single score, here more significant on SP than SW, which means that the amount of samples recorded on SP was closer to the expectations (samples of 5 sec every 55 sec) than on SW. It is expected as Android API's mechanics require the samples recorded on SW to be passed to the SP. The CMS correlates with the amount of files generated and observations on continuous trends, so CMS appears reliable in estimating the device logging. The linear regression slope is very similar for both and ranges from -0.9 to -1.7 on SP compared to SW, from -1.1 to -1.8. . A paired t-test conducted on the values of slope provides a p-value of 0.158 which is not statistically significant. The measured difference could be due to randomness of data and is not significant, indicating that the same motion is recorded on both SP and SW. The peaks recorded on SP were greater or equal to 1.7 for all participants which is consistently greater than peaks recorded on SW, equal to 1.1 for most. A paired t-test conducted on these peaks provides a p-value of 0.0020. The peak value difference indicates that the most common LRE recorded are higher on the hips than on the wrist. Therefore, phenotypes show that similar movements are recorded on both devices but with a different magnitude which is consistent with the location and handling of the devices. These characteristics are consistent with observations of LRE over time while providing a simplified view. So, this methodology will be further used and evaluated in subsequent studies.

5.6.3 Study considerations

This study aimed at evaluating the smartphone capabilities, in data recording and on-board processing, enforcing control over the smartphone settings and disabling power-saving mode. This represents the ideal data recording conditions, yet days of missing samples could be observed. These monitoring issues are likely due to patients using the devices as secondary phones and read/write interaction with the phone storage (raw logs saved as text files), which

triggers battery optimisations and power saving. However, consideration should be made of technological challenges when using smartphones.

Furthermore, participants were asked to keep the smartphone as close to the hip as possible, but the smartphone´s handling varies per user, e.g., carried by hand, in a pocket or in a handbag. Therefore, the handling of the smartphone was biased and did not represent the smartphone´s usage by common users. Therefore, further studies should evaluate the amount of data samples obtained in free living conditions without power-saving restrictions and using the primary smartphone of the participants to provide a more accurate representation of continuous LRE monitoring.

5.7 Conclusion

This study confirms that smartphone sampling can be uniform and unbiased by user's usage conditions by disabling power save mode. Raw accelerometer data can consistently be captured and stored on the phone's memory but require significant storage space. The load rates calculated and stored in SQL database, as opposed to raw data, demonstrate that the smartphone's on-board processing capabilities are reliable, and match raw data processed offline (MATLAB/Python).

Smartwatches are more sensitive to wrist movements, but the overall pattern and load trends measured by smartphones and smartwatches are similar. Smartwatches show higher peaks, but the trends, slope and intersect are comparable, which supports the use of smartphones to evaluate load rates on the lower limbs while being closer to the user's centre of gravity.

Individual phenotypes look at physical activity characteristics (i.e., trend, slope, intersect and peak) and will be further utilised in subsequent studies. Study 3 evaluates the impact of clinical interventions (knee injection) in OA patients' physical activity, compared to wristbands (Fitbit) commercially available, using step count as a standard unit of measure. Study 4 evaluates the trends of load forces that can be measured using the primary smartphone over an extended period.

Chapter 6 OApp Sydney: Continuous monitoring in osteoarthritis clinical trial

This chapter covers the steps undertaken, considering the input from osteoarthritis patients, to compare load force estimates and step count, before and after clinical intervention.

6.1 Introduction

The results of study 2 (chapter 4) have shown that the sampling discrepancies identified in study 1 (chapter 3) can be addressed by manipulating the power save settings of the phone. Moreover, load rate trends obtained from smartphone recordings are comparable to smartwatches worn at the wrist. It represents the ideal usage conditions but not how smartphones are used in free-living conditions. To date, not enough studies have been made on load rate to validate smartphone´s estimates, so we compare estimates of step count from smartphone and Fitbit to use an accepted unit of physical activity (PA). Moreover, it is impossible to correlate disease and physical activity (as per study 1) without significant changes in the patient's condition or treatment. So, OA patients were recruited as part of a clinical trial to validate that the methodology can measure changes in behaviour before and after injection.

6.2 Literature review

6.2.1 Osteoarthritis and clinical treatments

There is no cure for OA, but unlike RA, the disease does not necessarily worsen, and several treatments can help relieve the symptoms (see chapter 2). The damage to tissue and cartilage of OA patients might progress and be monitored using patients-based questionnaires such as the Western Ontario and McMaster Universities Arthritis Index (WOMAC), Knee Injury and Osteoarthritis Outcome Score (KOOS) and Hip disability and osteoarthritis outcome score (HOOS). The use of questionnaires and diaries in the context of PA assessment has already been reviewed, with a positive correlation found using self-reported WOMAC and objectively monitored with Fitbit (Morcos et al., 2020). The key features to consider have been identified as quality and objectivity of PA measures along with cost and limitations (Sylvia et al., 2014). Although questionnaires can add valuable insights, this study does not use them as the focus is identifying the change in PA. The relation between knee loading and pain has been investigated in a cross-sectional study using VAS scales to measure pain intensity (Henriksen et al., 2012). The study identified a positive correlation but only for patients with severe OA. Guidelines for pain

management exist (Bannuru et al., 2019), including prescriptions of paracetamol, NSAIDs and COX-2 inhibitors to manage pain. In most extreme cases, joint surgery might be required to replace the affected joint with a prosthesis. Injections provide an alternative to knee surgery while being stringer than drugs. Injections of corticosteroid injections help reduce inflammation, and injections of hyaluronic acid help support the thinning caused by OA (Ayhan, 2014). Platelet-Rich Plasma (PRP) helps heal injuries and is injected after being mixed with a blood sample. Still, the efficacy of this approach remains the subject of multiple studies (Gato-Calvo et al., 2019) (Paterson et al., 2018).

Supposedly, reducing the pain should remove the patient´s barrier to exercise and improve overall mobility towards following PA guidelines. A systematic review of studies has concluded that only a small amount of OA patients meet the PA guidelines (Wallis et al., 2013). The review included studies measuring moderate to vigorous PA and step counts. A consensus statement in 2022 was released to recommend exercises for OA considering PA to strengthen the bones, prevent falls and reduce the risks of vertebral fractures (Brooke-Wavell et al., 2022). The use of MET was identified by an international cohort (Gates et al., 2017) to standardise the measurements of PA. Metabolic Equivalent of Task (MET) is a unit of energy expenditure defined by the amount of oxygen consumed by an individual at rest. The measure of oxygen at rest is about 3.5 millilitres per kilogram per minute, representing the baseline for comparing activities requiring more oxygen (Jetté et al., 1990). The METs required per activity have extensively been defined in a compendium in 1999 (AINSWORTH et al., 1993) and updated in 2000 (AINSWORTH et al., 2000). The classification describes PA considering major activity groups such as walking, cycling, occupation and home activities. Activities are then assigned a level of intensity which ranges from 0.9 METs (sleeping) to 18 METs (e.g., running) and consider numerous activities such as cycling (8.0 METs) and walking (2.5 METs). Recent years have seen a significant evolution of fitness trackers (Shanes, 2017) routinely used to measure moderate to high level of activities performed outdoor. The commercial expansion of brands such as Fitbit indicates that the broadest audience generally accepts these devices as mean to monitor exercise and physical activity, continuously and outside of specialised fitness environment. Fitness conscious people focus on measuring their performance and improvements while OA patients have to consider that the impact of repetitive motion can lead to damage in muscles, joint cartilage and bones (Milner et al., 2006). The tracking of step count and exercises focuses on repetitive movement with a population already active but does not consider the quality of movement and stress caused on the lower limbs. However, the method can be applied to OA considering that METs-min can be obtained from step count estimates.

6.2.2 Wearable and step count estimates

Wristbands and smartwatches typically use embedded sensors that allow tracking of energy expenditure (calories) and heart rate and are positioned at the wrist are ideal for monitoring arm swing and defining patterns, e.g. activity recognition (Attila Reiss, 2013) and step counts. These devices encourage a competitive mindset, and pedometers have shown effects on diabetes patients (Delfien Van Dyck, 2013). This effect has also been demonstrated in RA patients through a significant increase in physical activity and subsequent decrease in fatigue in a population of 96 participants after using a pedometer for a week (Katz PP, 2015). The level of participation in activities such as walking can be increased through structured classes, as shown in a study with eighteen older adults (Tudor-Locke et al., 2002). Fitbit´s daily target is set by default to 10000 steps (ROSENBAUM, 2019) which is not typically achieved in daily activities, as identified in a literature review (Choi et al., 2007). The search included studies published between 1982 and 2006 that used pedometers and step counters. The daily count can be improved through sports and home activities, but the results found a deficit of 4000 to 6000 steps when performing only routine activities. So, expecting OA patients to perform 10 000 steps routinely is not suitable and a preliminary study identified the step count range to classify the level of PA (Tudor-Locke and Bassett, 2004). In this context, the daily step count considers individuals as active ≥10 000 and highly active ≥12500 steps/day. Under 10 000 steps/day is considered somewhat active for 7500–9999, low active at 5000–7499 and sedentary under 5000 steps/day. A literature review that covered 1594 articles was further commissioned in 2010 by the Public Health Agency of Canada (PHAC) and concluded that using step count estimate is a good baseline for PA guidelines communicated to the general public (Tudor-Locke et al., 2011).

Besides acceptance from the public measured by the commercial success, wrist-worn devices wristbands have been validated amongst other commercially available devices, PA trackers (Wahl et al., 2017) to provide reliable step count estimates. Fitbit was also validated against research-grade devices such as Actigraph GT3X+ in various studies found in a library of publications that have used Fitbit devices in research (Fitabase, 2022). The step count recording was validated with a community of 32 older people over seven days (Paul et al., 2015). Assessment of active minutes in free-living environment was also validated, with 53 participants wearing the device for seven days (Brewer et al., 2017). The usage is also acceptable for a more extended timeframe, as shown in a study with 42 female breast cancer patients (Hartman et al., 2018). The findings demonstrated that over 12 weeks, the Fitbit trends allowed us to visualize changes in PA habits triggered by a research intervention. Therefore, Fitbit provides a reliable way to track PA although a systematic review of 67 studies identified caveat such as overestimating step count in free living conditions (Feehan et al., 2018).

Step counters have become popular but the principle for counting steps is not new and can be achieved in several ways, as identified in a literature review of step count measurement considerations (Bassett et al., 2017). The devices can be worn at the waist, pocket, thigh or foot and the most accurate position is at the ankle. A basic method is to identify the steps measuring the vertical variations of ups and downs at the waist, which can be achieved using a spring levered or, more recently, with accelerometers. Devices worn at the wrist consider the movements related to arm swing, and algorithms have been developed for PA classification (Zhang et al., 2012) as well as optimising step count estimates. Smartphones´ ability to estimate step count has been measured, and multiple algorithms have been developed to handle the noise introduced by free usage and device handling. The general principle is to capture accelerometer data and apply noise filtering to identify peaks. A study used a low pass filter (Myo et al., 2018). It was also proposed to estimate steps by flattening insignificant acceleration changes using a discrete Kalman filter (Kinh et al.) on iPhone accelerometers. Walk detection and step count might filter the angular movement using Fast Fourier Transform, as seen in studies using unconstrained Android devices (Kang et al., 2018, Dirican and Aksoy, 2017). Adaptive threshold calculation for magnitude and temporal filtering following the peak extraction was used to deal with the dynamics of smartphones in various device handling, such as texting and calling (Lee et al., 2015). Another proposed technique adaptively tunes the filters and thresholds without the need for pre-sets classifier (Khedr and El-Sheimy, 2017). Android provides a set of sensor APIs and, in more recent releases, improved the support for fitness applications by releasing the ability to record step count (Google, 2022a). This project does not focus on the algorithms provided by Android, but step count is a unit typically used by most fitness trackers as an intuitive measure and widely advertised by Fitbit. The most significant benefit in using the step count API is preventing the risks of tracking termination due to power save identified in Study 2.

6.3 Rationale

The correlation between disease and physical activity requires continuously monitoring a patient with significant symptoms for an extended amount of time. So, we measure smartphone monitoring with patients before and after a clinical intervention. The changes in behaviour triggered by injection emulate disease activity changes usually seen in an extended timeframe.

Load rate estimation is not a standardised way to measure physical activity, so we compare the step count measured by smartphones with Fitbit estimates to validate the smartphone´s ability to perceive significant variations.

6.4 Aim & objectives

6.4.1 Aim

This study aims to evaluate the variations of physical activity, load forces and step count, measured by smartphone in a clinical trial that introduces a noticeable medication change through knee injection of OA patients.

6.4.2 Objectives

The objectives of this study are:

1) Evaluate the ability to continuously capture load forces measured by smartphones in free usage conditions (i.e., without power save mode restrictions or handling restrictions)
2) Compare the amount of step counts measured by smartphone and Fitbit
3) Assess the patient´s behavioural changes before and after the injection

6.4.3 Hypothesis

If a clinical intervention introduces a change in physical activity behaviour, measuring the load forces through a smartphone's continuous monitoring could identify and assess this change.

6.5 Methodology

6.5.1 Patient recruitment and study procedure

This study is a pilot, observational and quantitative study, focusing on comparing the trends and correlation of physical activity recorded via smartphones and Fitbit. Each participant wore a Fitbit Flex 2 wristband and used a Samsung Galaxy A5 preloaded with OApp™. Samsung devices (12 phones) were loaned in rotation to cover the broader amount of participants.

The protocol and additional documentation, including consent forms, have been reviewed and approved by Northern Sydney Local Health District HREC (Reference number: LNR/17/HAWKE/370).

Patients recruited were sixty-six adults, part of the RESTORE trial (ANZCTR registration - ACTRN12617000853347), split into placebo and intervention groups based on their consecutive entry into the study. Inclusion criteria were defined to recruit patients diagnosed with Osteophytes on x-ray, willing to participate in this study and above 50 years old. The clinical trial

requires patients explicitly with a minimum pain score of 4 on an 11-point numeric rating scale for the last week and knee pain on most days in the previous month.

The clinical team first screened participants, giving verbal descriptions and information on the trial over the phone before sending the Participant Information Sheet and Consent form by postal mail. The suitable participants then underwent X-ray, blood tests; and physical screening before being invited to join the trial.

Once recruited, patients received the loaned Fitbit wristband and smartphone preloaded OApp™ for seven days. Fitbit and smartphone were then returned, and the clinician collected data. Patients repeated this procedure for an additional seven days following a RESTORE trial's standard timeframe of two months.

6.5.2 Data collection

For this study, the monitoring aimed at recording hourly step count and data samples of 5 seconds (window=5) every 15 seconds (interval=15). Step count and LRE have been stored in segregated SQL databases to avoid CRUD conflicts and uploaded to Azure table storage at the end of each day. Fitbit data are synched online through the Fitbit app that was preloaded on the phone. Fitbit daily step counts have been exported as an excel spreadsheet through Fitbit's online portal.

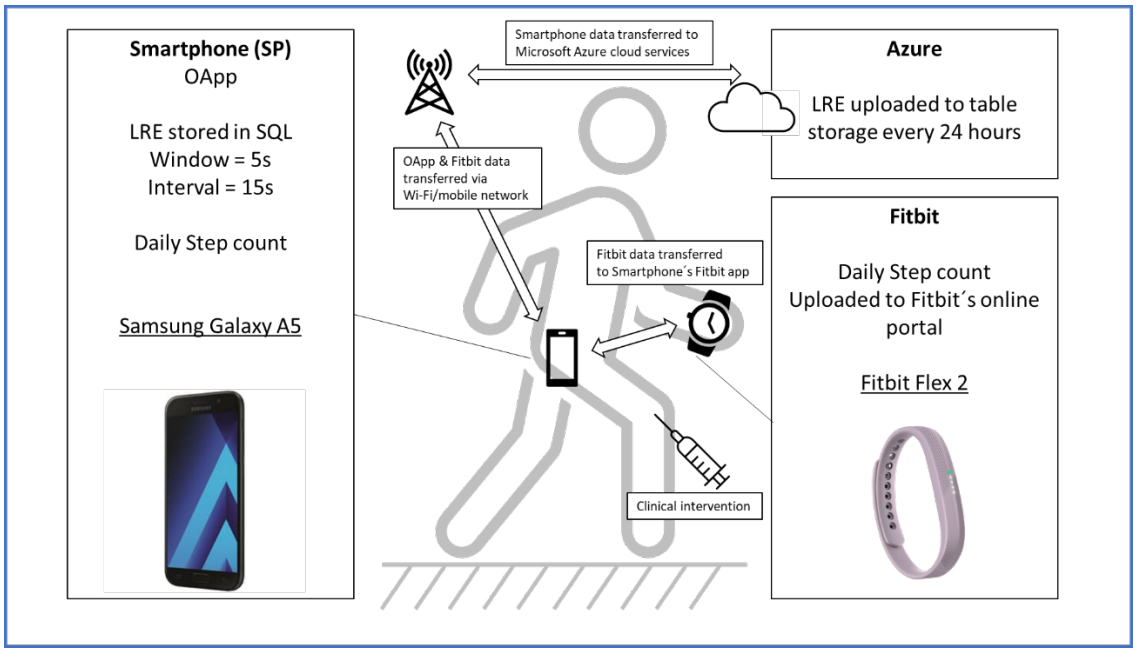


Figure 56: Smartphone & Fitbit

6.5.3 Data analysis

The previous study (study 2) showed that the sampling could be uniform with power saving disabled. This study aims to emulate conditions of usage closer to real life. So, smartphones were set with their default settings (i.e., power saving mode enabled) which causes variations in the data capture interval preventing comparisons. As per previous study, the sampling can be manipulated but the aim is to emulate free living conditions. Doze mode might extend the sampling interval up to 5min (instead of 55sec). So missing LRE samples, for periods of interval longer than 5 min, were assumed to be due to power saving and replaced with imputed samples using the surrounding values, as illustrated in fig 57. The LRE samples were then interpolated into fragments of 5 min to obtain uniform sampling across all patients.

For each participant, the steps recorded by the phone are compared with Fitbit's daily step count as a baseline of physical activity trends. As for study 2, we draw individual phenotypes to represent the proportion of time at load. The continuous monitoring score (CMS) indicates reliability on the amount of samples recorded. The peak represents the load with the highest proportion of time, and the linear regression slope illustrates the overall trend of physical activity.

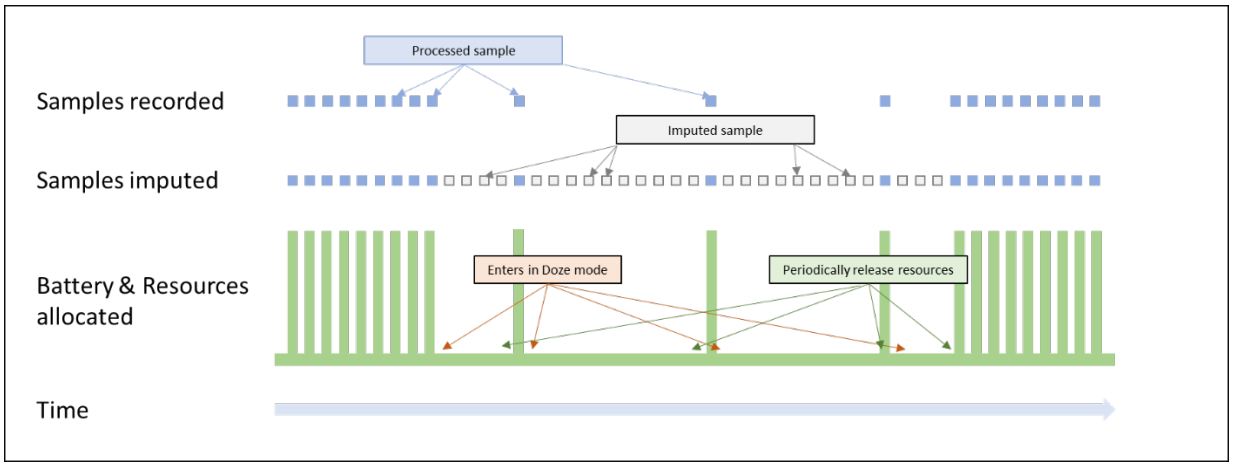


Figure 57: Data Sampling

6.6 Results

6.6.1 Overall participation

Table 12 summarises the amount of data collected over both sessions. Most of the 66 OA patients completed the recruitment period except patient 2, who had technical issues and 8, who dropped out of the study. The data collection of smartphones and Fitbit for both recruitment periods were completed by 15 patients (23%). Most patients (60%) completed smartphone data collection at baseline and after 2 months except 12 patients (17%) who couldn't attend the second recruitment

period. A total of 146892 LRE were recorded for both periods across all participants. Patients generated an average of 2260 LRE, with patient 3 generating the most (8350 estimates) and patient 63 generating the least (98 estimates). Fitbit data were recorded successfully on both sessions for 19 patients (29%), but 39 patients (60%) recorded less than ten days overall due to technical issues or not being able to attend the second session, and the remaining eight patients (12%) completed only the first session.

ID	LRE Days	Fitbit Days	ID	LRE Days	Fitbit Days	ID	LRE Days	Fitbit Days
2	2	1	24	15	15	46	14	14
3	15	14	25	14	7	47	14	14
4	14	14	26	14	8	48	14	14
5	14	1	27	14	5	49	14	
6	15	14	28	14	2	50	8	7
7	15	15	29	10	7	51	14	1
8	1		30	15		52	14	7
9	14	13	31	8	6	53	8	7
10	14	14	32	14	2	54	15	14
11	15	13	33	14	7	55	14	7
12	8	7	34	8	1	56	14	14
13	14	11	35	14	9	57	14	10
14	8	7	36	14	14	58	1	
15	15	7	37	14		59	7	7
16	14		38	14		60	14	6
17	7	7	39	12		61	14	7
18	14	14	40	14		62	14	13
19	15	8	41	7	2	63	1	1
20	8	7	42	14		64	14	14
21	14	14	43	14		65	14	14
22	14	13	44	14	11	66	7	
23	14	7	45	14	14			

Table 12: Summary samples (SP LR estimates and Fitbit) across patients

6.6.2 LRE Monitoring

Smartphones were used with their default power save setting (i.e., disabled) which follows the Android pattern of battery optimisation. As expected, the recording of samples follows the patterns of usage of the phone, which experience LRE much lower at night than during the day. Fig 58 shows that LRE is at least four times lower between 00h00 and 04h00 than during the rest of the day and gradually reduces from 19h00 to 23h59.

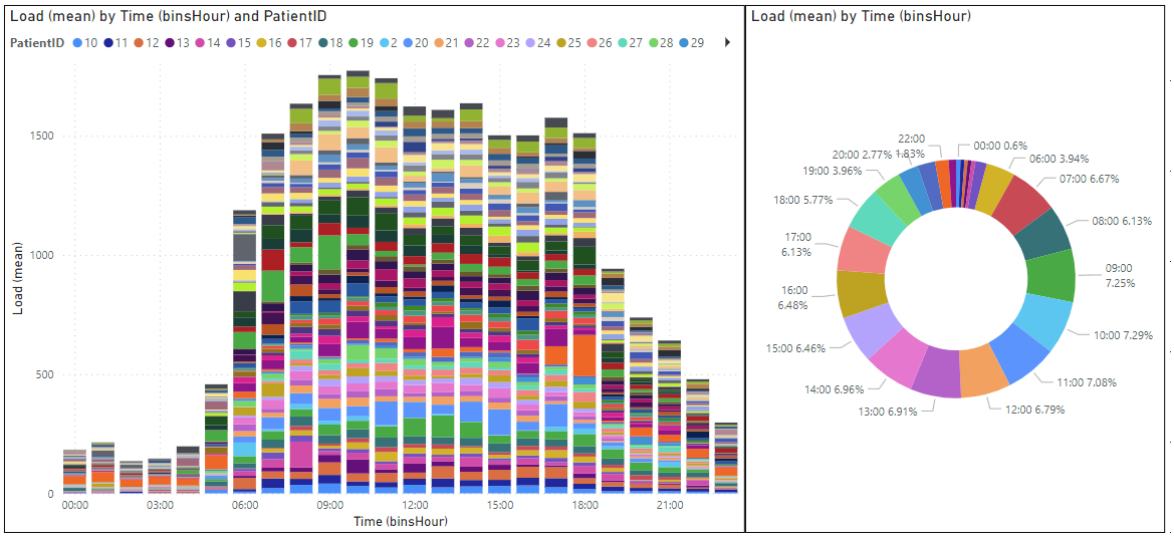


Figure 58: LRE (mean) per hour

As a result, the interval between samples is not uniform over time and less samples are generated between 00h00 and 05h00 than the rest of the day, as seen in fig 59. A third of the sample (29.9%) follows the window setting with an expected sample rate of 50. However, most samples are smaller than expected, with 55.75% of samples generated with a rate lower than 10.

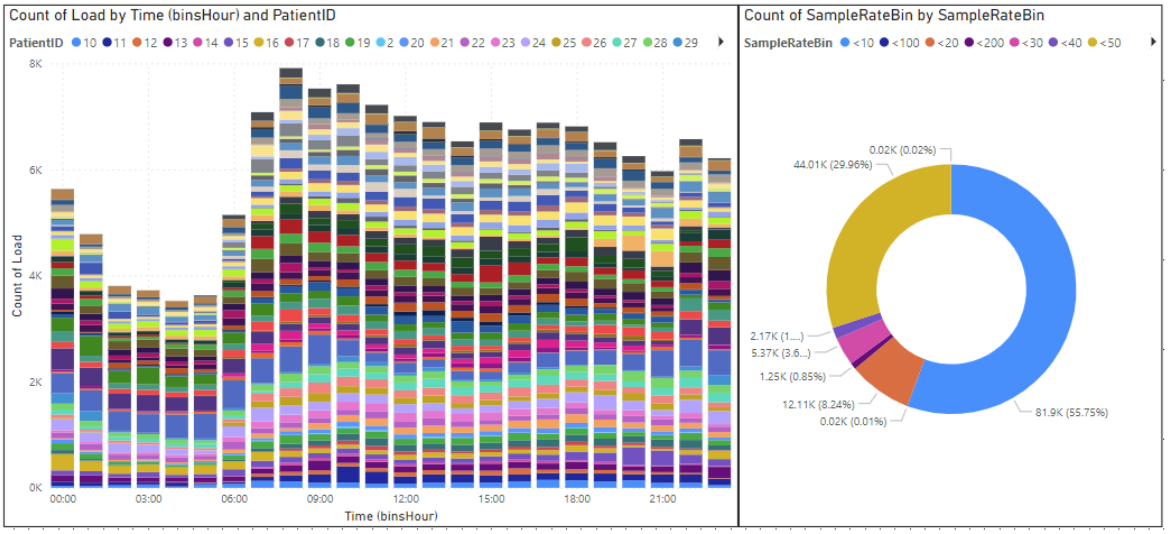


Figure 59: Count of LRE per hour

The missing samples introduce a bias to the mean LRE from raw accelerometer data. Therefore, samples are imputed based on the surrounding values and interpolated, as seen in fig 60, which shows the hourly LRE from raw and interpolated data. As expected, the trend of interpolated is like raw data and shows lower estimates between 00h00 and 04h00. However, looking only at raw data overestimates the mean of LRE. Interpolated LRE mean is also much lower as considers periods of inactivity which are frequent.

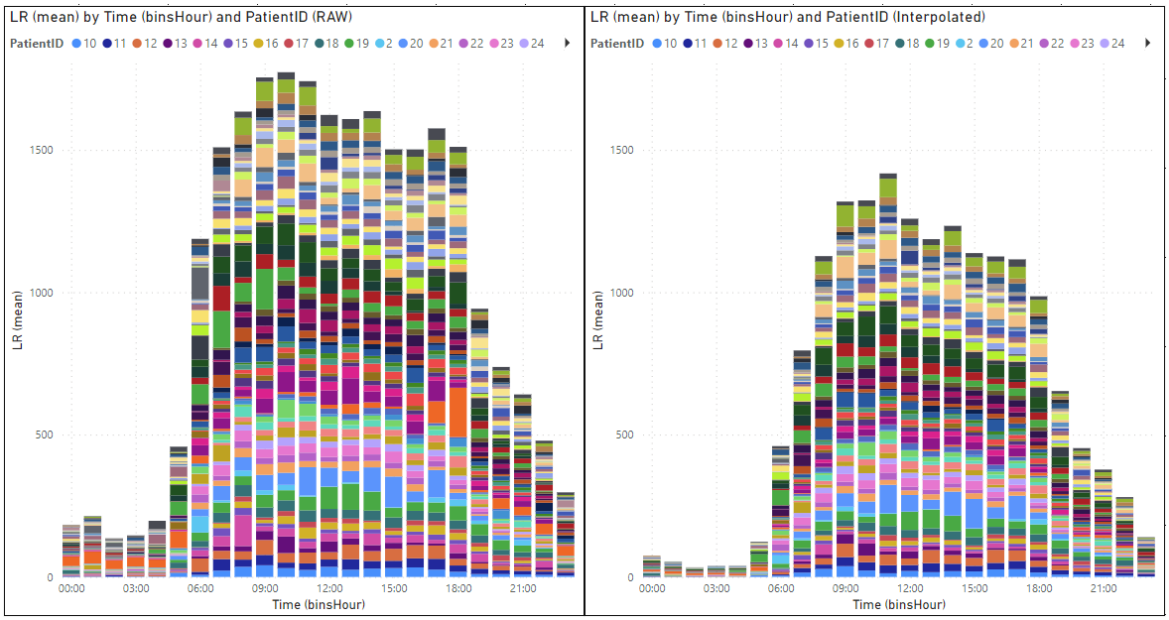
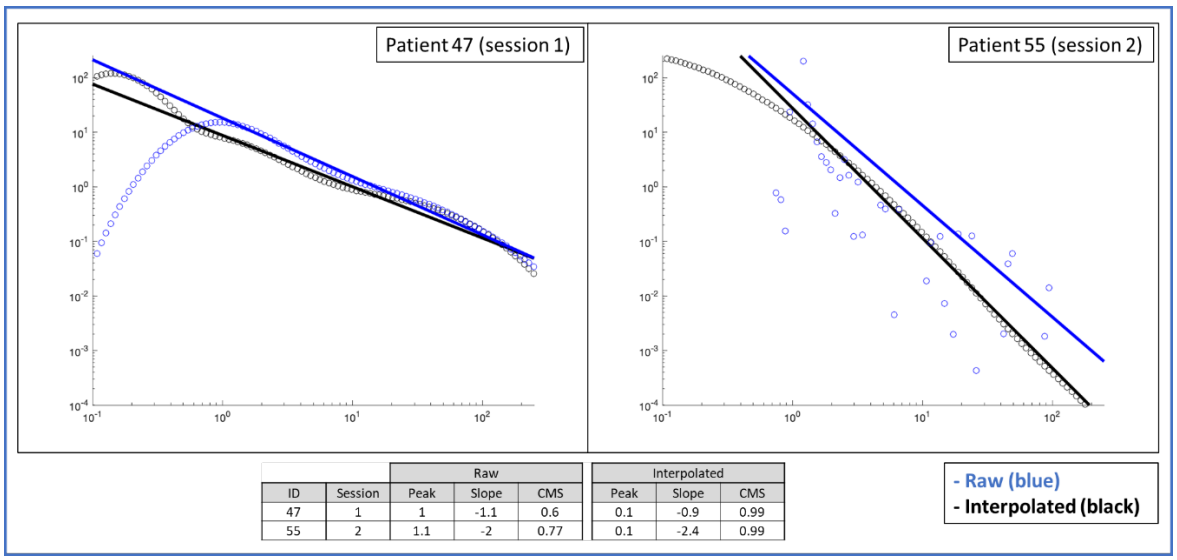


Figure 60: LRE (mean) per hour (raw and interpolated)

Fig 61 illustrates an example of phenotype for the most significant between-person variation of the most and least active sessions, respectively, of patient 47 at baseline (with a slope of 0.90) and patient 55 after injection (slope of -2.40).



		Raw			Interpolated		
ID	Session	Peak	Slope	CMS	Peak	Slope	CMS
47	1	1	-1.1	0.6	0.1	-0.9	0.99
55	2	1.1	-2	0.77	0.1	-2.4	0.99

Figure 61: Raw & Interpolated Phenotype (Patients 47 and 55)

Fig 85 shows the phenotype characteristics from raw data on both sessions for all participants. The CMS across patients and sessions ranges from 0.35 for the second session of patient 12 to 0.92 for the second session of patient 3, which means that not all samples are accurate LRE representations. The slope should consider the interpolated data in combination with the CMS. Most patients had a slope under -1.4, indicating a low amount of physical activity. This is expected for a cohort of OA patients but could also suggest that the smartphone was not used if the CMS is

low. The peak ranges from 0.6 during the first session of patient 36 to 1.3 in the second session of patient 62. However, peak 1.3 occur with CMS lower than 0.37, so the most common LRE recorded is similar for this cohort across patients and sessions. For most patients, the peak of raw data is 1.1, and the interpolated average peak is 0.2 because of imputing missing samples, which generally occurs when the phone is inactive.

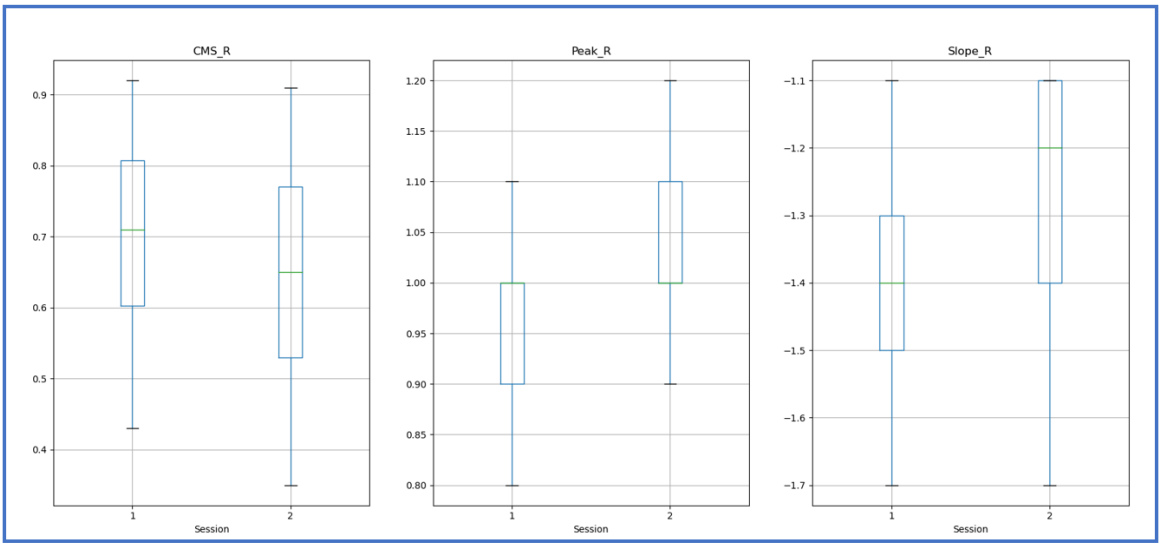


Figure 62: Phenotype characteristics

6.6.3 Step count monitoring

To evaluate the correlation, we consider the patient samples with at least five days of recording more than one step on both smartphone and Fitbit, which reaches 0.78, as seen in fig 86 (left). When looking at the correlation for days with a minimum of 1000 steps completed on both devices, the overall correlation reaches 0.86, as seen in fig 63 (right). The plot shows a relation of proportionality with a 1:1 ratio for most patients (56%) as seen in table 13. However, a proportion twice higher can be seen, and patients 14 and 19, respectively, recorded three and four times more steps with Fitbit than SP.

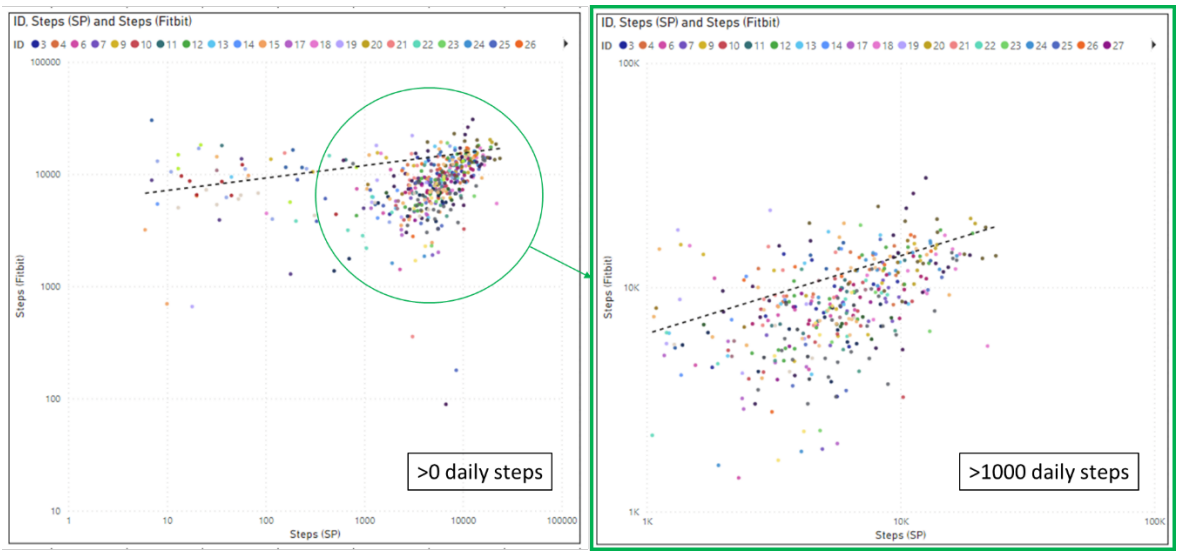


Figure 63: Daily step count correlation

ID	Days	Steps (SP)	Steps (Fitbit)	Correlation	Ratio
3	14	72813	121052	0.18	2
4	14	98021	164367	0.45	2
6	13	52128	92245	0.31	2
7	15	105740	121353	0.56	1
9	11	69040	117534	0.09	2
10	14	78736	94138	0.25	1
11	11	64785	107605	0.79	2
12	7	88324	102558	0.57	1
13	10	55545	101200	0.44	2
14	6	23327	60180	0.85	3
17	7	21153	27859	-0.57	1
18	13	161166	152327	-0.14	1
19	7	17755	73057	-0.08	4
20	7	93429	101430	0.94	1
21	12	50447	120692	0.27	2
22	9	46097	65276	0.84	1
23	7	49982	60527	0.07	1
24	15	135719	169390	0.76	1
25	6	49797	49898	0.59	1
26	8	54934	95020	0.69	2

ID	Days	Steps (SP)	Steps (Fitbit)	Correlation	Ratio
27	5	22400	35534	0.5	2
31	6	40374	67416	0.9	2
33	7	41587	50329	-0.23	1
35	5	29905	60375	0.52	2
36	12	60512	74824	0.39	1
44	9	82485	107790	0.3	1
45	14	110277	131578	0.71	1
46	14	110751	178418	0.44	2
47	14	184170	214177	-0.01	1
48	13	60354	103038	0.7	2
52	7	42972	47843	0.08	1
53	7	30984	45183	0.46	1
54	13	71161	136002	0.29	2
56	13	73516	92322	0.52	1
59	7	31884	39770	0.84	1
60	5	22180	38532	0.76	2
62	5	19194	28711	0.38	1
64	14	127039	102484	0.83	1
65	14	90785	128873	0.83	1

Table 13: Correlation smartphone and Fitbit (daily step count)

6.6.4 Monitoring before & after injection

Figures 64 to 67 show the trends at baseline (left) and after 2 months (right) of daily Fitbit and SP step count and accumulated load rates for patients that participated for more than 9 days overall and recorded more than 1k daily steps. Most days, the smartphone´s count of steps follows the same trend as Fitbit. However, no substantial behavioural changes can be observed on the daily step counts before and after injection. Fig 68 shows the distribution of LRE for each participant before and after clinical injection. As for step count, no perceivable changes in LRE behaviour can be observed at baseline and two months after injection.

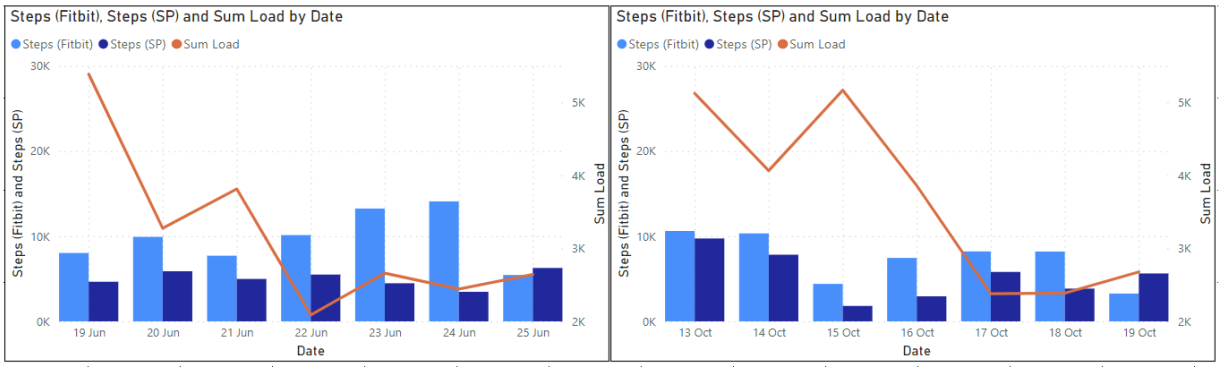


Figure 64: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 3)

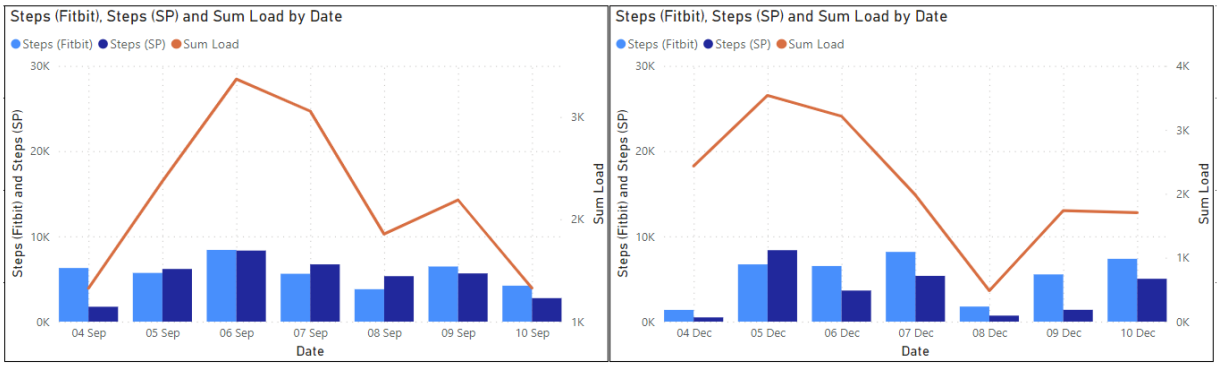


Figure 65: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 36)

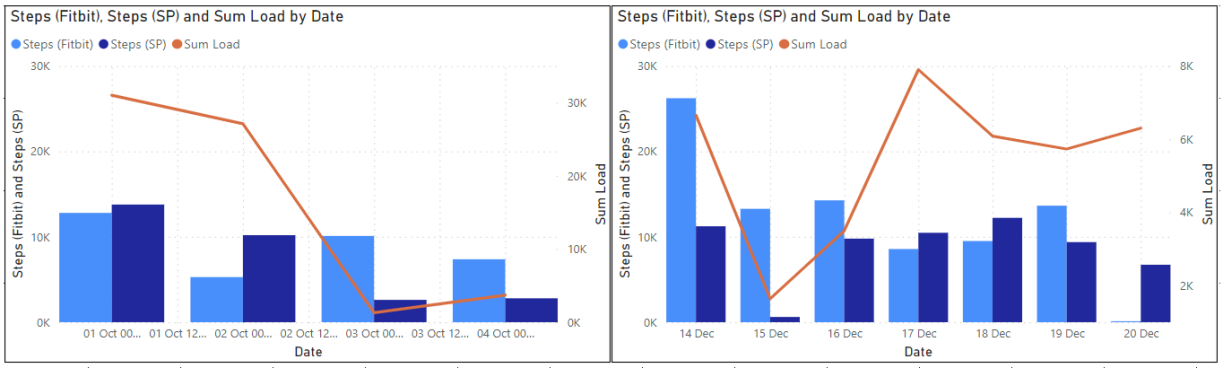


Figure 66: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 44)

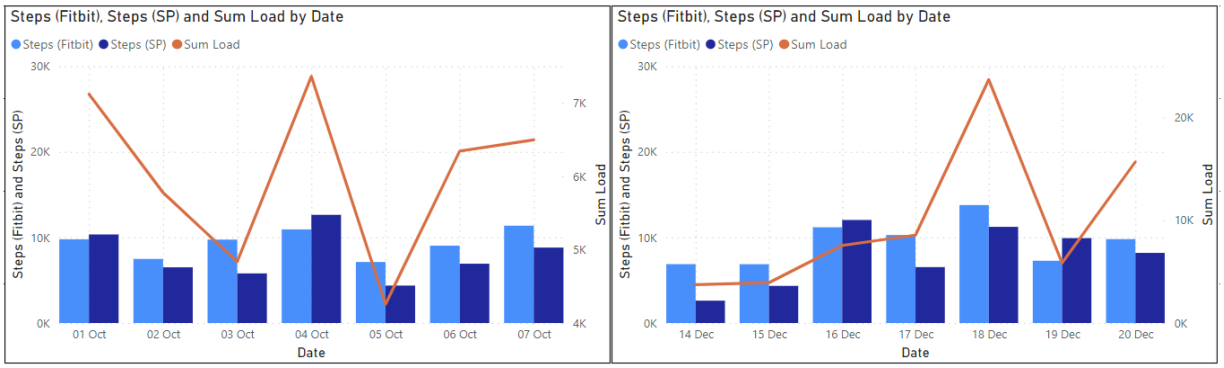


Figure 67: Sum of LR and Step count (Fitbit & SP) at baseline and after 2 months (Patient 45)

Fig 91 shows the distribution of step count recorded across the cohort of patients, before and after injection, through Fitbit and smartphones. A paired t-test conducted on the step count recorded via Fitbit provides a p-value of 0.391 which is not statistically significant. Similarly, A paired t-test conducted on the step count recorded via SP provides a p-value of 0.574 which is not significant either.

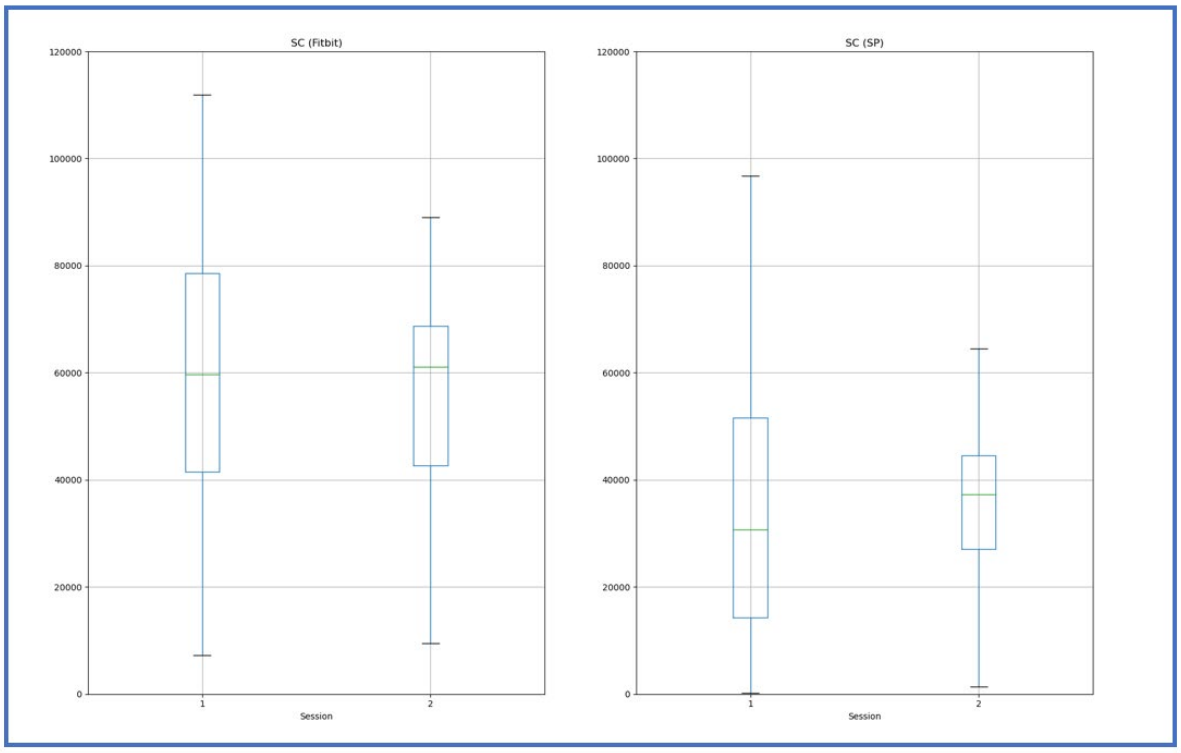


Figure 68: Boxplot of Step count across all patients, before and after injection

Fig 69 shows the distribution of LRE for each participant before and after clinical injection. As for step count, no perceivable changes in LRE behaviour can be observed at baseline and two months after injection. A paired t-test conducted on LRE recorded provides a p-value of 0.362 which is not significant.

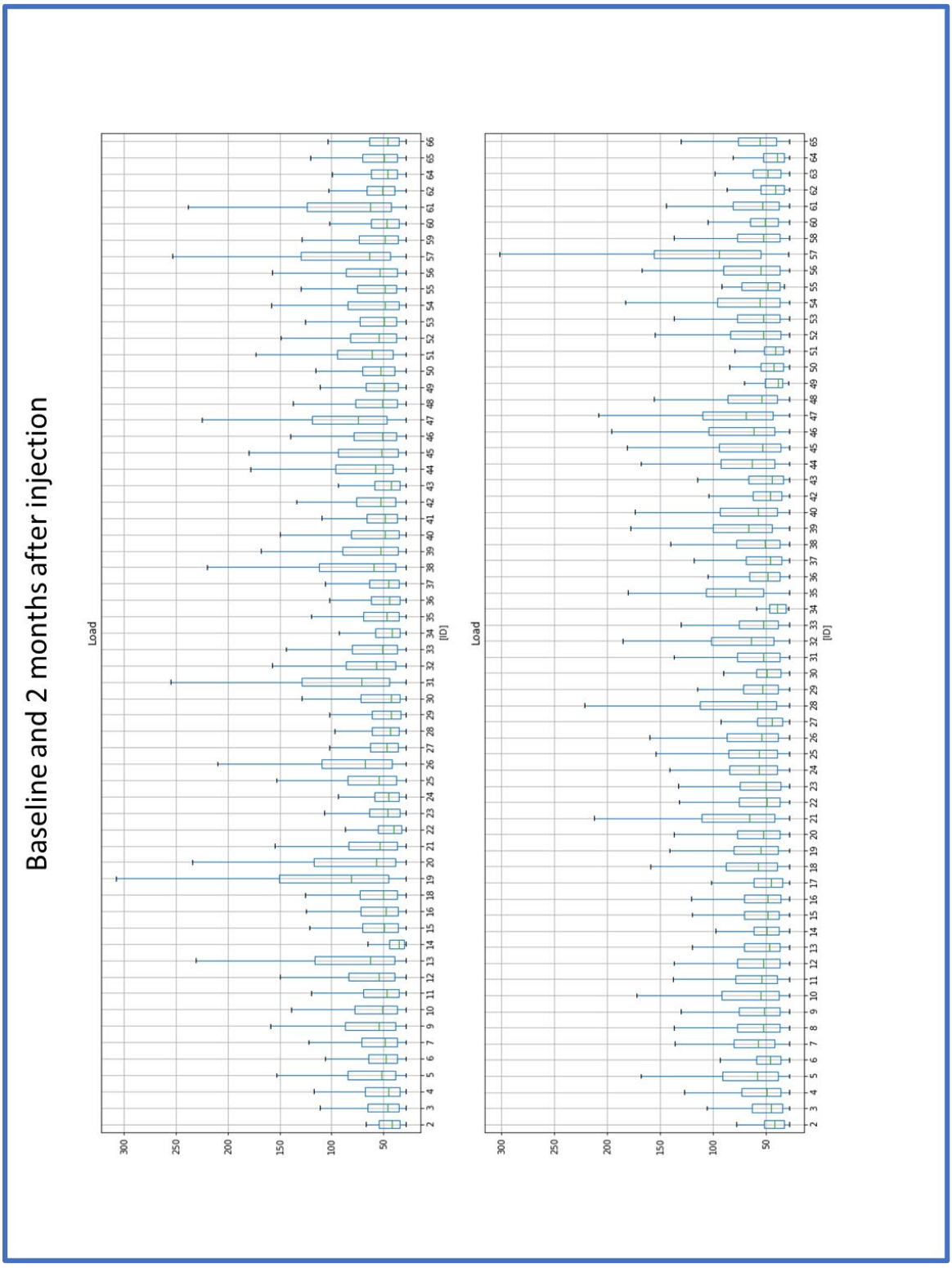


Figure 69: Boxplot LRE for all patients, before and after injection

6.7 Discussion

6.7.1 Continuous monitoring

The sampling for this study was set to capture 5 sec samples (window=5) every 15 seconds (interval=15), equivalent to 20 sec per minute. As expected, Android´s power saving affects the ability to record uniform sampling of accelerometer data. The mean LRE and amount of samples recorded at night are significantly lower than during the rest of the day. Users are typically sleeping at night and so this appears to confirm that the overall sampling is affected by the user's interaction with the device, as seen in study 1. Furthermore, the proportion of samples recorded at the correct window size was much smaller than in Study 2. The sample rate for 55.75% of samples was five times smaller than expected, which means that estimates were generated for smaller window sizes because of the power-saving algorithm shortening the data recording over a smaller window. The sample window size was smaller than the setting (under 50, 70% of the time) but there is no correlation with the value of LRE, as seen on fig 70. The distribution of samples is not random, so the length of sample rate is not random. Therefore, smartphone samples are not corrupted by the power save but bias the ability to estimate the individual mean of physical activity.

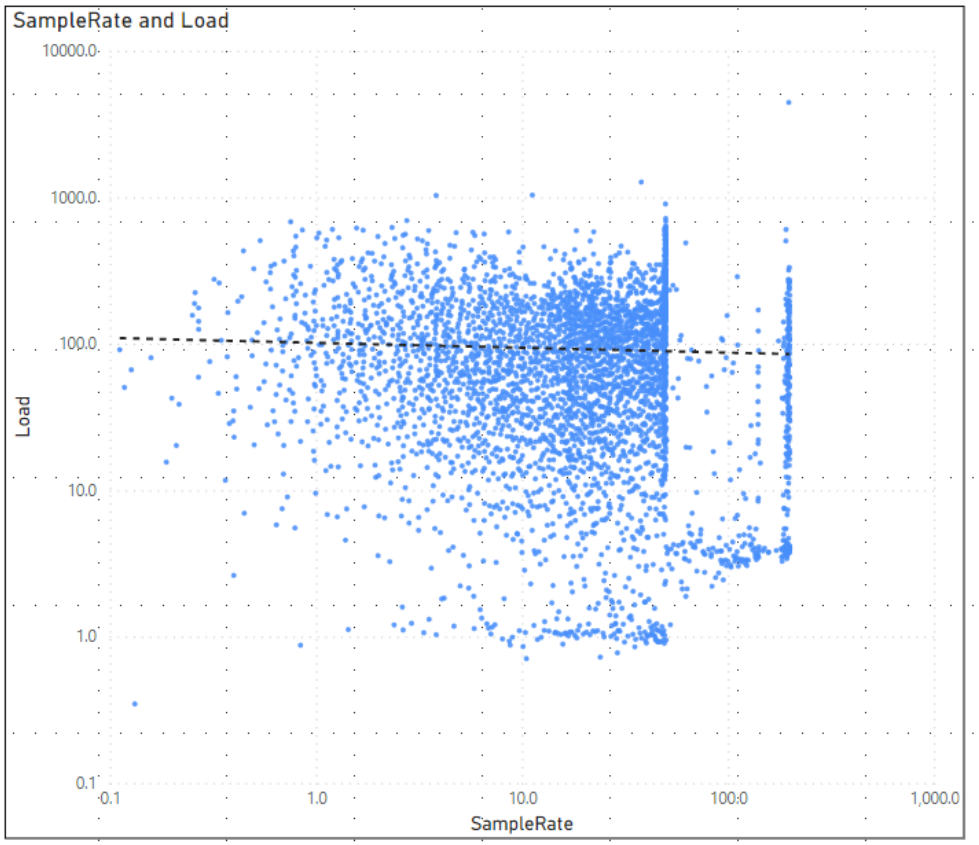


Figure 70: LRE and Sample rate

6.7.2 Step count comparison

Smartphone steps are recorded using Android´s API and estimated using Google´s proprietary algorithms, but APIs are not subject to the same restrictions as third-party apps (here OApp™). So, the amount of steps recorded by smartphones can be considered a reliable representation of smartphone estimates. Data recording challenges were also found with Fitbit, and not all patients recorded data as expected. Fitbit uses a smartphone app to retrieve the daily step count but manufactures the smartwatches and owns proprietary algorithms to estimate daily step counts. Therefore, Android power save doesn't appear to affect the ability to record step counts on a smartphone or Fitbit HR. The performance and accuracy to evaluate step counts, measuring the peak to peak of arm swing, is affected by the type of device and incurs false positive (Åkerberg et al., 2016). So, it is expected to see the correlation increase, when over 1000 steps have been measured, as the proportion of false positive decreases with higher amounts of steps. The Smartphone´s estimate of step count appears to be proportional to Fitbit estimates with a 1:1 ratio for 56% of participants. The remaining 44% observed ratios up to 4 times lower on smartphone which indicates that the phone may have not been used when the Fitbit was worn. Negative correlations, on the other hand, indicate more steps recorded by smartphone than by Fitbit, potentially due to Fitbit not being worn. . The impact on samples is expected because of power save. However, the step count results also show that smartphones were not carried as much, in most cases, as Fitbit devices. These results are consistent with Study 1 and the study protocol, as the phones were used as secondary, which caused limited interactions with smartphones and in carrying them.

6.7.3 Behaviour changes before/after injection

Fitbit doesn't allow access to raw accelerometer data, so it is impossible to compare LRE. The amount of steps recorded by a smartphone is generally lower than by Fitbit. It is expected considering the arm swing movement associated with devices worn at the wrist, as observed in the previous study. However, the step count trend of smartphones and Fitbit are similar at baseline for patients 7, 10, 11, 22, 24, 43 and 45, which confirms that the smartphone's recording is comparable to Fitbit. For patients 22 and 24, the trends after two months are also similar while different for the other patients. The position and usage of the phone explain these discrepancies. The step count reported by the phone is very low, which indicates that it was left in a static location (e.g., on a desk) instead of being carried close to the body. The daily amount of step count between participants is different, which indicates different levels of physical activity.

The aggregated LRE, however, does not always follow the same trend as step count and lower aggregated LRE can be seen on days with higher daily step count. The amount of force a patient generates is not only to perform steps. So, evaluating LRE beyond daily aggregates allows for gaining further monitoring insight. The CMS provides a measure of confidence in the amount of time logged by the device (raw data). The variations observed indicate that the sampling reliability was not the same across patients. As expected, CMS values are lower when power save is enabled, with an average of 0.67. Interpolated CMS is closer to 1 as it is obtained by imputing missing samples, so we consider only raw CMS. Low CMS affect the reliability of features extraction from phenotypes as a mean of comparison between patients and changes following injection. The slopes recorded are biased towards the most common LRE values recorded. The peak recorded were low for most, indicating that the smartphones were not used most of the time. As for CMS, the interpolated peak is biased towards low imputed values and can be ignored.

Knee injection is a substantial procedure for OA patients. The monitoring of behavioural and physical activity using wearable presents technical challenges but appears feasible. However, none of the patients showed drastic changes in physical activity and not enough variations were observed on the trends of step count and LRE variations to establish a direct correlation between knee injection and physical activity. Knee injections are typically scheduled shortly after being recommended and so increasing the monitoring timeframe is practically challenging as it might prevent access to care. It should also be noted that placebos were used in this trial, which increases the complexity as changes might not be expected at the first place. Nevertheless, the distribution of LRE intensity highlight outliers significantly outside the range of LRE. These spikes of LRE represent jerks recorded beyond the mean of individual physical activity, which might provide further insights into OA symptoms and conditions.

6.8 Conclusion

This study confirms that using smartphones as a secondary phone affects the sampling performance due to the smartphone’s power saving. However, the number of steps recorded shows that the sensors remain active when processed by the OS. The trends of step count recorded by the phone are like Fitbit, which validates the smartphone’s tracking ability.

Fitbit consistently recorded more steps due to the sensor's location on the body (i.e., dominant and non-dominant wrist) and sensitivity to upper limb movement (e.g., wider arm swing, hair stroke). Smartphones are not always close to the body (i.e., kept on a desk), so they are likely to record fewer steps than devices worn at the wrist, as seen in study 2.

Wristbands are worn at a static position on the wrist (dominant or non-dominant), which is optimised to estimate arm swing patterns and step counts. Smartphones can be held and carried at multiple locations (e.g. pocket, hand) but are typically carried closer to the centre of gravity of the human body. Therefore in the context of patient monitoring, smartphones provide a better representation of the impact loading on the lower limbs and joints, including bursts of a sudden jerks. As for study 1, the amount of recorded samples and timelines are insufficient to draw a correlation between disease and physical activity. As a result, study 4 will recruit participants willing to install the app on their phones.

Chapter 7 RApp2: Long-term monitoring

This chapter covers the steps undertaken to assess load force estimates, considering the input from rheumatoid arthritis patients, over 3 months.

7.1 Introduction

The results of Study 1 raised the challenges associated with smartphone sampling, and study 2 confirmed that these can be addressed through power save settings. Study 3 observed a low sampling performance, but the capability to continuously monitor the sensors of the phone remains, with comparable step count trends as Fitbit trackers. Smartphone power-saving is triggered by a long period of user inactivity. Therefore, smartphones used as secondary devices do not provide an accurate representation of free usage and cause a significant amount of data losses. So, for this study, we evaluate LRE that can be obtained in real-life use with RA patients recruited, as for Study 1, but asked to load RApp™ on their smartphone (i.e. primary device) and for a more extended recruitment period of 3 months (6 weeks in Study 1).

7.2 Aim & objectives

7.2.1 Aim

This study aims to assess whether smartphone's continuous monitoring over an extended period can be used to evaluate the variations of physical activity and load forces.

For this study, two sets of samples were collected:

1) RA patients recruited for three months
2) Researcher data over five years

7.2.2 Objectives

The objectives of this study are:

1) Evaluate the ability to continuously capture load forces by a smartphone used in free-living conditions, as the primary device and with default power save mode optimisation
2) Assess the changes in physical activity, and load forces over an extended period

7.2.3 Hypothesis

If load forces and step counts are continuously monitored on a patient's primary smartphone, in free living conditions and for an extended period, then physical activity trends and patterns can be obtained, that can be useful to assess correlations with disease activity and symptoms.

7.3 Methodology

7.3.1 Patient recruitment and study procedure

This study is observational and quantitative and compares trends and correlations of self-assessed disease activity and load rates recorded via smartphones. The protocol and additional documentation, including consent forms, have been reviewed and approved by London Stanmore's ethics committee (REC reference 16/LO/0182). To be eligible, inclusion criteria were defined to recruit any participants diagnosed with rheumatoid arthritis, willing to participate in this study and above 18 years old. Participants were recruited from the database of patients of the University Hospital Southampton NHS Foundation Trust.

Due to COVID restrictions, patients were invited to review the study details by accessing the Participant Information sheet from a QR code advertised on a poster at the hospital. Patients interested could then access an online consent form and receive an email confirmation, once consented, with the details to download RApp™ from Google Play directly on their phone. Each participant used their smartphone (Android based) and installed RApp™, which should address the impact of power saving and quantify the amount of tracking obtained in real-life monitoring. The updated user interface includes more joints (DAS) and only the RAPID3 questionnaire provided, which is not designed for weekly use, nor validated but fits the purpose of self-assessment by the recruited patients. No specific considerations are required on the medication, as this is a pilot study, but the recruitment period has been extended to 3 months.

Obtaining ethics to recruit patients within the NHS requires obtaining approvals from the relevant committee, so recruiting patients nationally (or across countries as for study 3) requires obtaining all the relevant approval. Similarly, obtaining data from participants for longer than three months requires unrealistic commitment on their end, considering that this study is observational and does not provide intervention or improvement of their medical condition. Hence, the second part of this study considers control data generated over the years by the researcher. RApp™ was installed on personal smartphones, which changed over the years to control the protocol's correct functioning and technical aspects.

7.3.2 Data collection and analysis

For this study, the monitoring aimed at recording hourly step count and data samples of 5 seconds (window=5) every 15 seconds (interval=15). Self-assessments, step count and calculated load rates have been stored in the smartphone´s segregated SQL databases and asynchronously uploaded to Azure table storage whenever a cellular or Wi-Fi connection was available. The samples were then exported.

For each participant, we evaluate smartphone performance to continuously monitor LRE for the recruitment period (3 months). As for previous studies, we assess the amount and integrity of recorded samples. The phenotype and characteristics introduced in Study 2 are further used to evaluate the proportion of LRE over time.

7.4 Results

7.4.1 Patient´s continuous monitoring (3 months)

Due to COVID restrictions and associated safety preventions, a limited cohort of three RA patients (two males and one female) was recruited at the hospital and consented remotely via online form. All patients could download and install RApp™ correctly on their phones, as summarised in Table 14. Patients were invited to use RApp™ for over three months if they wished. Patients p73 and p74 correctly completed the recruitment period and kept using the app for six months and three months. Patient p75 had smartphone settings issues and completed a partial timeframe of 2 months.

Patient		Recruitment		Amount of Mean LR (per day)				
ID	F/M	Start	End	Total	Average	Min	Max	Std Dev
p73	M	11/05/2021	05/11/2021	69897	397.14	13	984	203.4
p74	M	25/05/2021	24/09/2021	29339	916.84	13	2455	695.85
p75	F	25/05/2021	19/07/2021	14665	977.67	110	2134	610.43

Table 14: Patient recruitment summary

Fig 71 shows the trend of LREs for each patient. The density of the plot indicates that patient p73 consistently recorded LRE over the recruitment period. Patient p74 recorded LRE between 25/05 and 02/06 and a few random days in June. The peaks in August and September are related to days of the RApp™ periodically restarting the sampling before being uninstalled on 24/09. Patient p75 recorded an overall of 11 days which can be seen as distinctive periods with density. The app did not run correctly for the rest of the time until being uninstalled on 19/07.

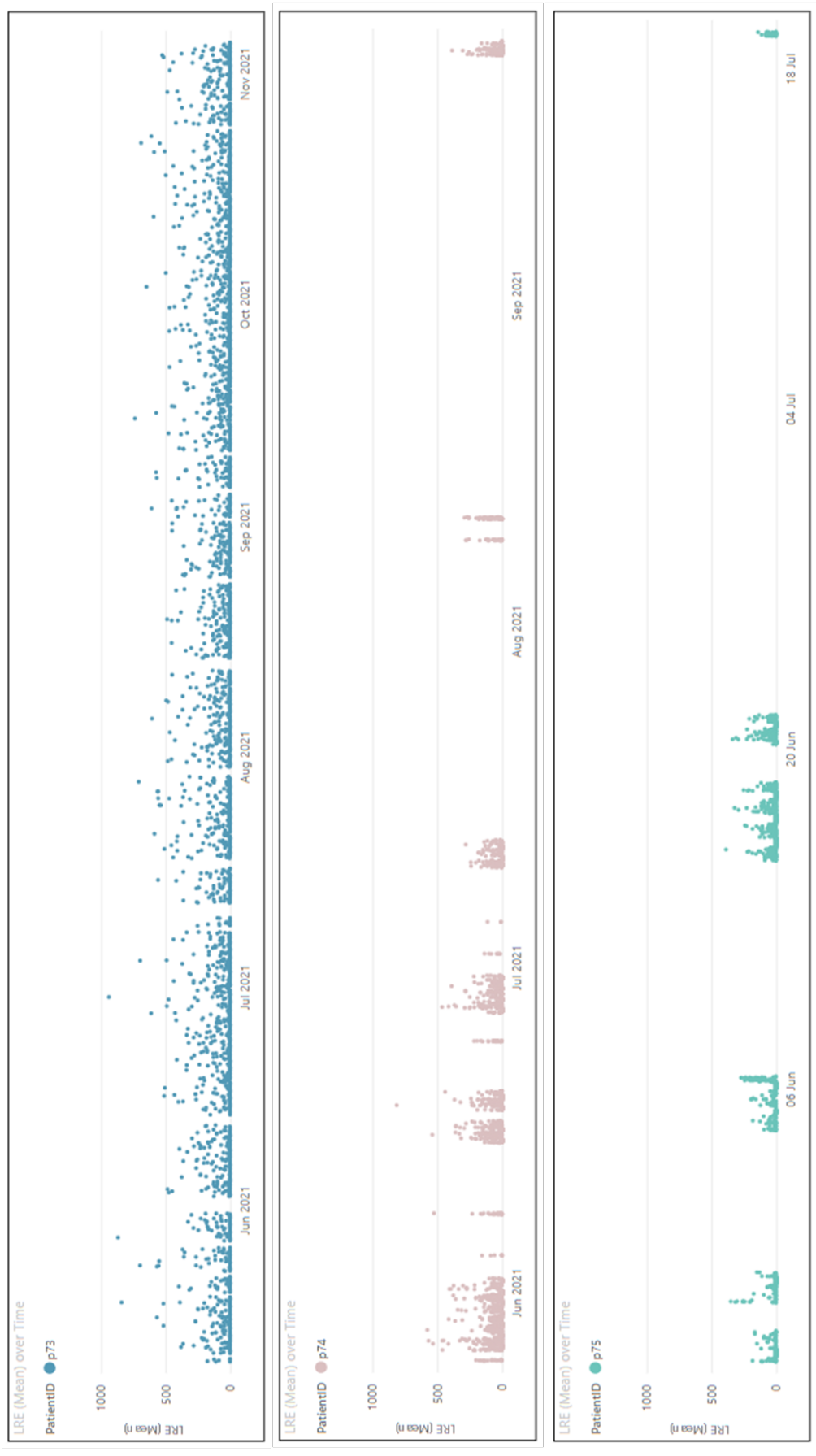


Figure 71: Smartphone's LRE (Mean) over time

Fig 72 shows that the amount of sample recorded over time is higher in the morning than the rest of the day. This is expected as smartphones are generally put on charge overnight, which means that the sample distribution is not uniform. On the other hand, most of the samples follow the expected rate of 50 (70.5%) or slightly greater at 60 (26.5%).

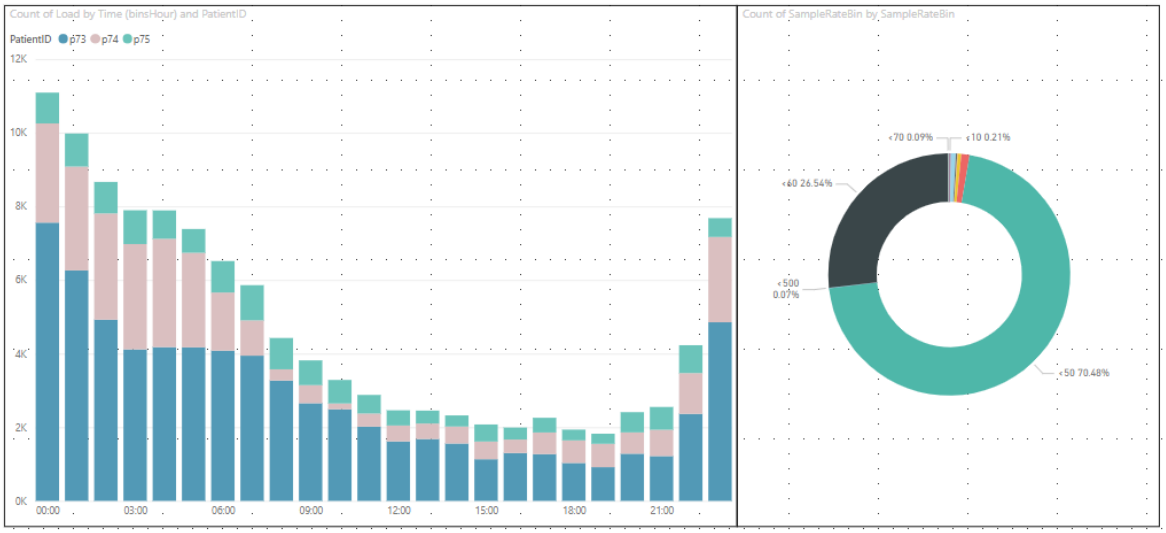


Figure 72: SP sample count and sample rate per hour

Fig 73 shows the mean of LRE evaluated from raw and interpolated data. Interpolated data consider the whole period that each patient had the app installed. As a result of missing days, the overall LRE mean is significantly lower than with raw data.

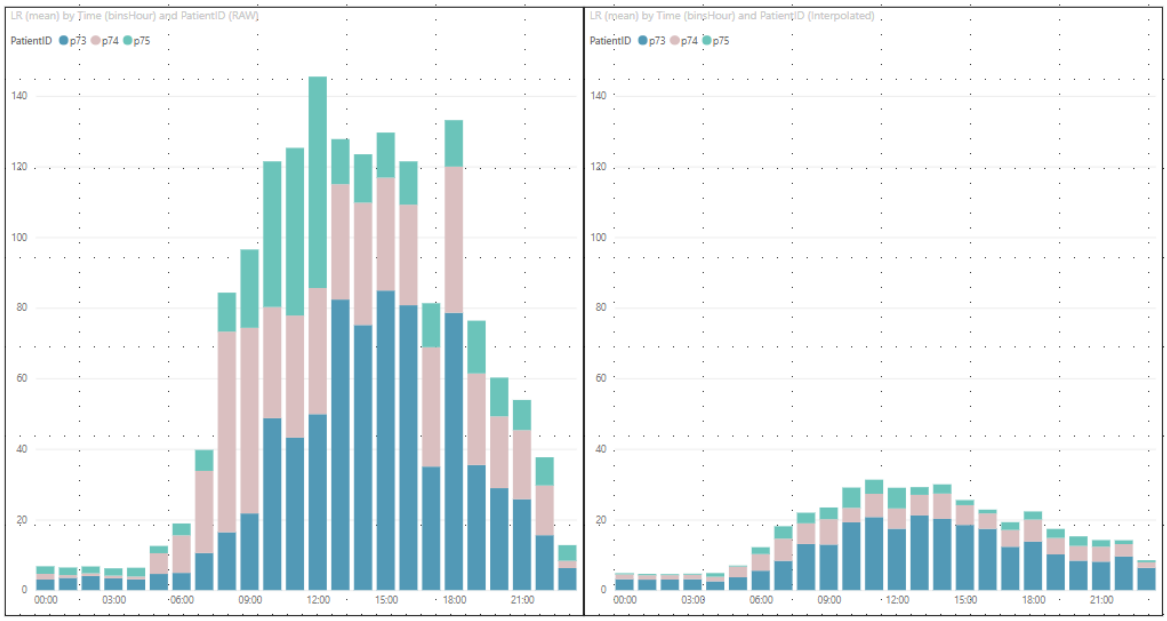


Figure 73: LR (mean) estimates per hour (raw and interpolated)

Fig 75 to 77 and Table 15 illustrates the phenotype and characteristics for patients p73, p74 and p75. The CMS of p74 is the highest, which means that most samples can be explained. Although p74 has the highest interpolated slope of -1.2, the LRE peak of 0.7 means that the phone recorded less movement than p73 and p75, with a peak of 2.

The boxplot (Fig 74) shows a low distribution of movement for p74 compared to p73 and p75, which recorded a wider range and higher maximum of LRE. Outliers of p73 reach LRE burst 800 times higher than its max LRE and at least 100 times higher for p74 and p75.

	Raw			Interpolated		
ID	Peak	Slope	CMS	Peak	Slope	CMS
p73	2	-1.5	0.76	0.1	-1.3	1
p74	0.7	-1.5	0.98	0.1	-1.2	0.98
p75	2	-1.6	0.7	0.1	-1.6	1

Table 15: Phenotype characteristics

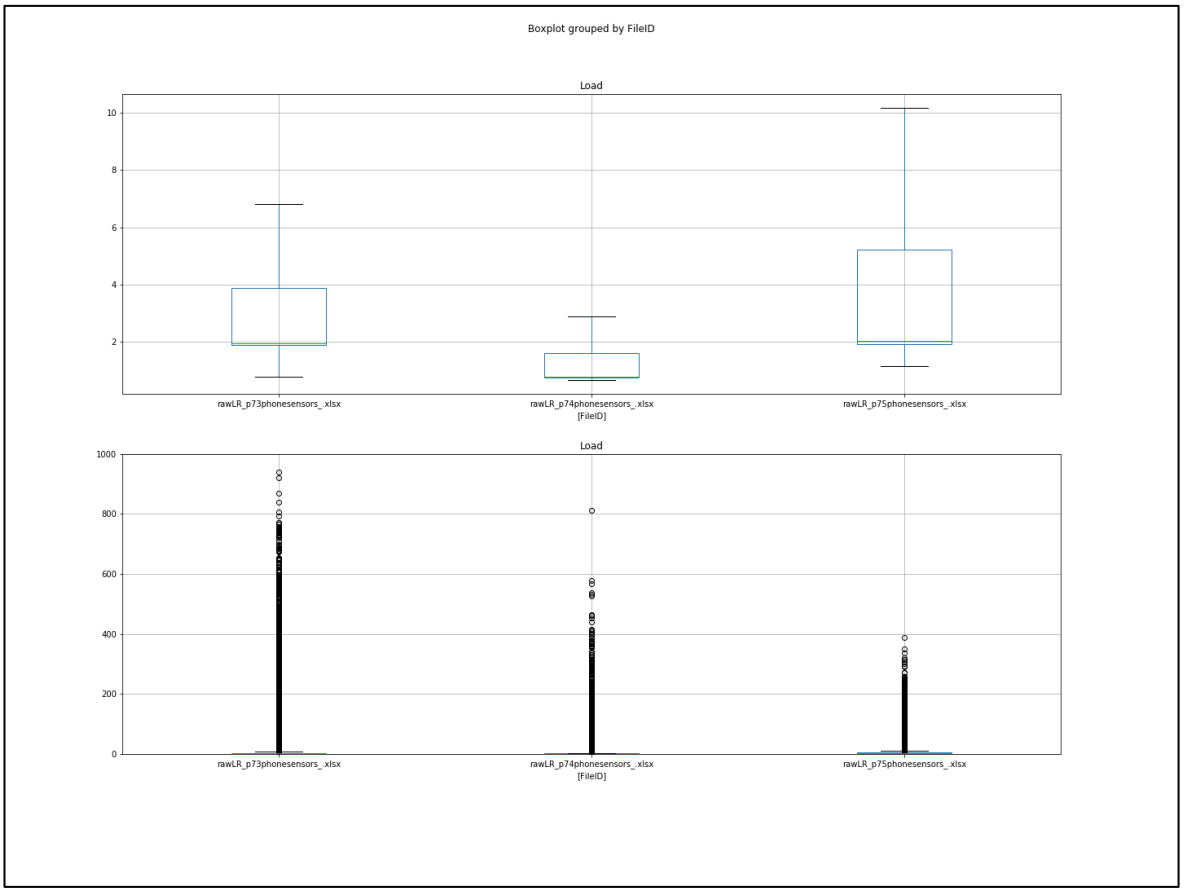


Figure 74: Boxplot (with and without outliers)

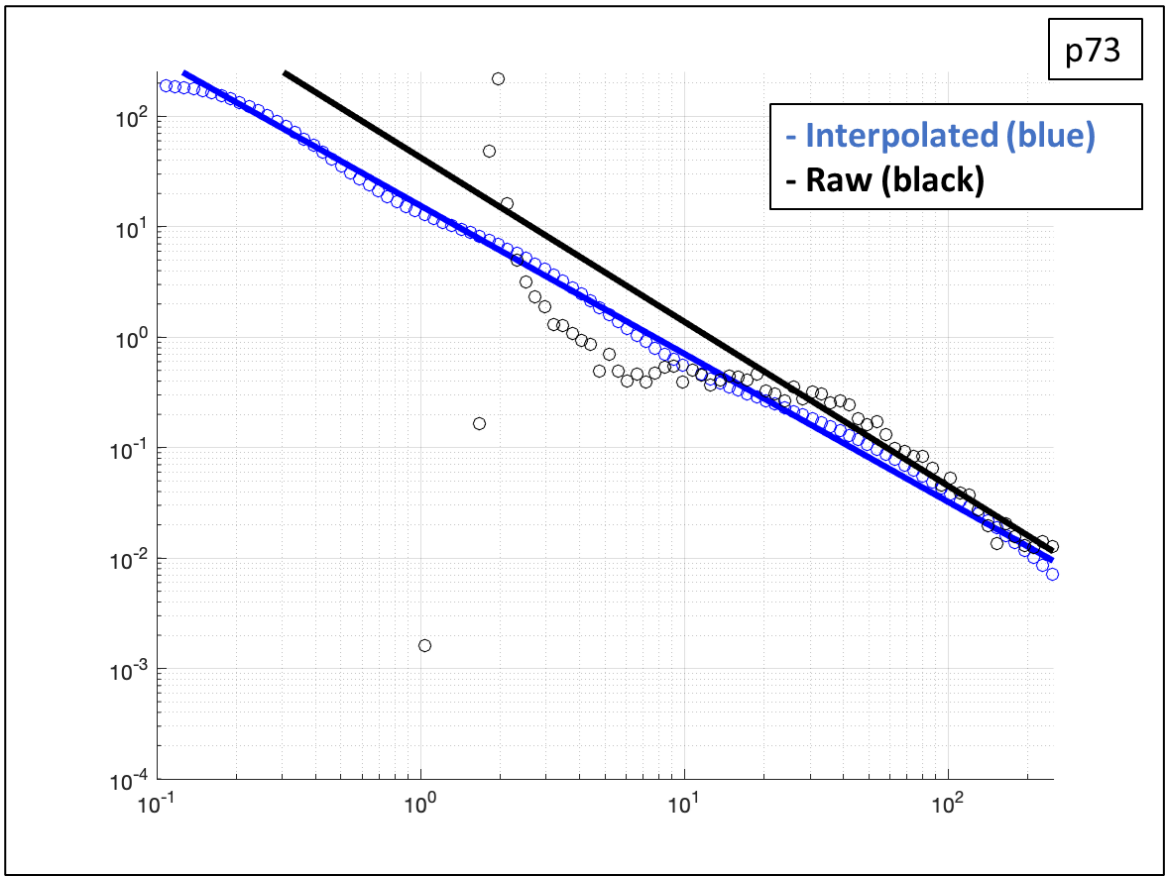


Figure 75: Phenotype (patient p73)

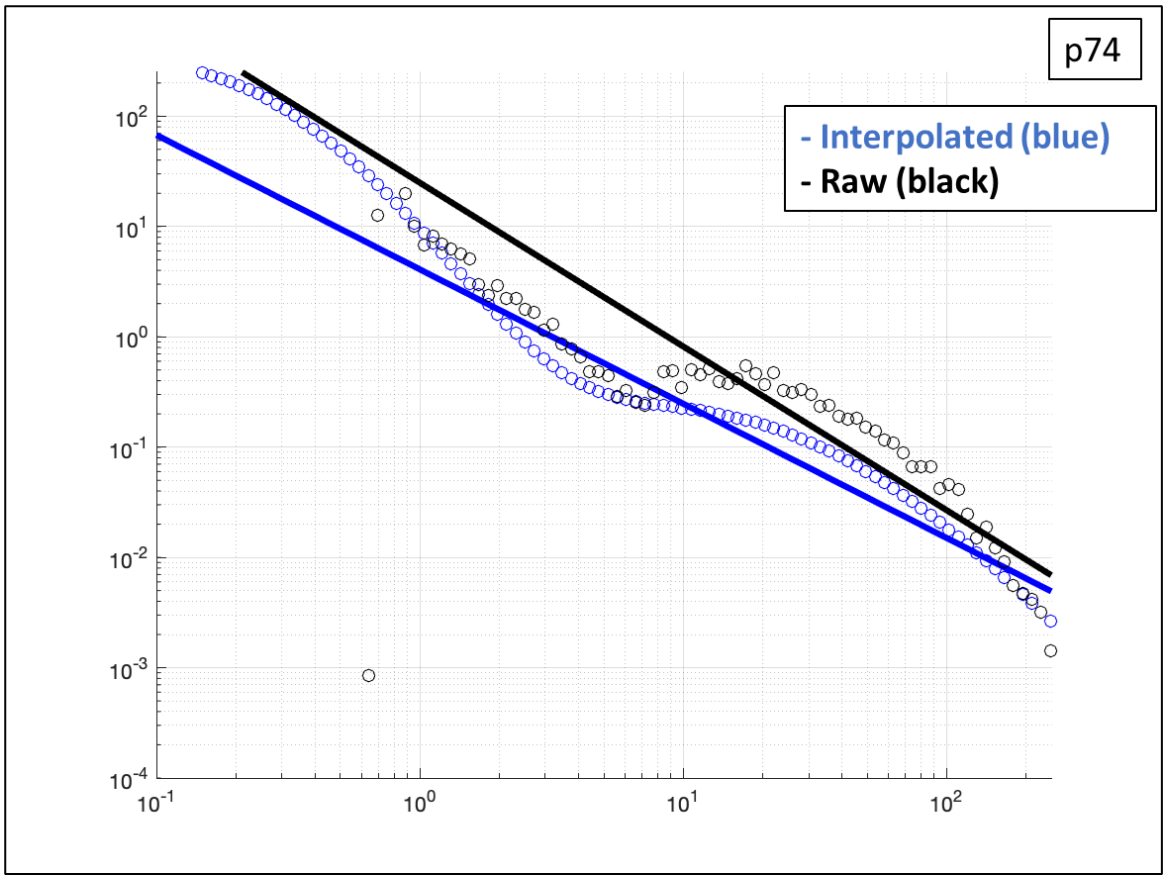


Figure 76: Phenotype (patient p74)

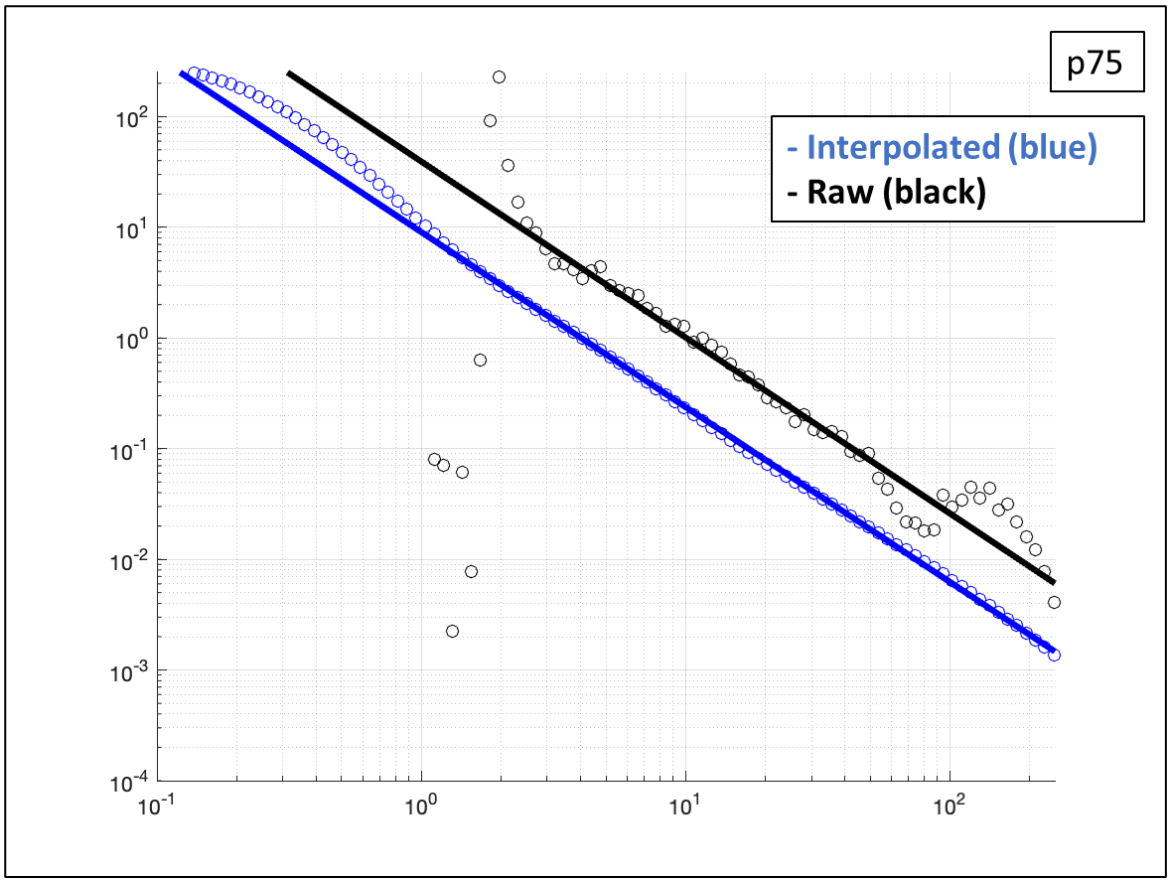


Figure 77: Phenotype (patient p75)

7.4.2 Continuous monitoring over 5 years

This project's development and technical follow-up required a control and troubleshooting mechanism. So, RApp™ has been running on the primary device of the researcher between 2018 and 2022. Table 16 summarises the period each smartphone was for, along the quantity of samples gathered. Early testing was done using HTC smartphones, but the first primary phone used was a Samsung S7 from January to March 2018, referring to the datasets s72901, s7110218 and s7040318. A gap in data occurred until using a Samsung S9+ from September 2019 to May 2020, corresponding to datasets s9jimmymay and s9jimmymay2020 in the plots. Another gap occurred until using an Xperia 1i from September to May 2022, corresponding to Xperia 1i and p72UK. The last dataset was captured using a Samsung 22+ from April to December 2022, referred to as p78. Fig 78 shows the timeline of samples recorded over time with each smartphone.

Patient		Recruitment		Amount of Mean LR (per day)				
ID	Model	Start	End	Total	Average	Min	Max	Std Dev
s72901	Samsung S7	28/01/2018	11/02/2018	29990	1999.33	69	3057	800.73
s7110218	Samsung S7	11/02/2018	04/03/2018	28580	1299.09	259	2710	598.6
s7040318	Samsung S7	04/03/2018	07/04/2018	74170	2119.14	491	4119	869.5
s9jimmy	Samsung S9+	09/09/2018	28/09/2019	192668	500.44	11	2162	353.61
s9jimmymay2020	Samsung S9+	04/12/2019	17/05/2020	115730	697.17	125	1952	345.07
Xperia1i	Sony Xperia 1i	05/09/2020	21/11/2020	68755	881.47	223	1808	445.1
p72UK	Sony Xperia 1i	02/04/2021	05/11/2021	213603	998.14	3	3025	633.03
p78	Samsung S22+	24/04/2022	23/11/2022	184388	861.63	151	1905	342.57

Table 16: Smartphone devices summary

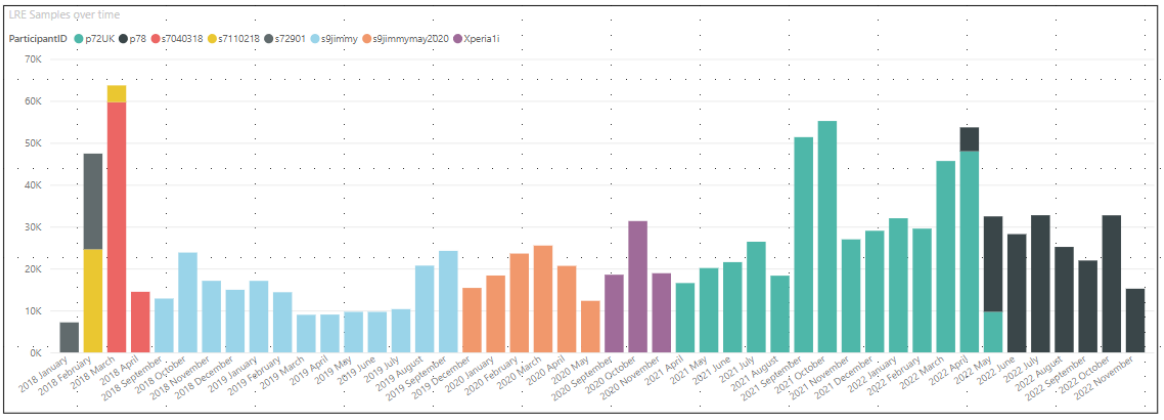


Figure 78: Smartphone devices and samples recorded over time

Fig 79 and 80 show the yearly LRE (mean) over time across all devices (top plot). 2022 shows a higher overall LRE magnitude than the previous years. 2019 and 2020 shows the lowest magnitude, which matches the timeline of working from home and Covid restrictions.

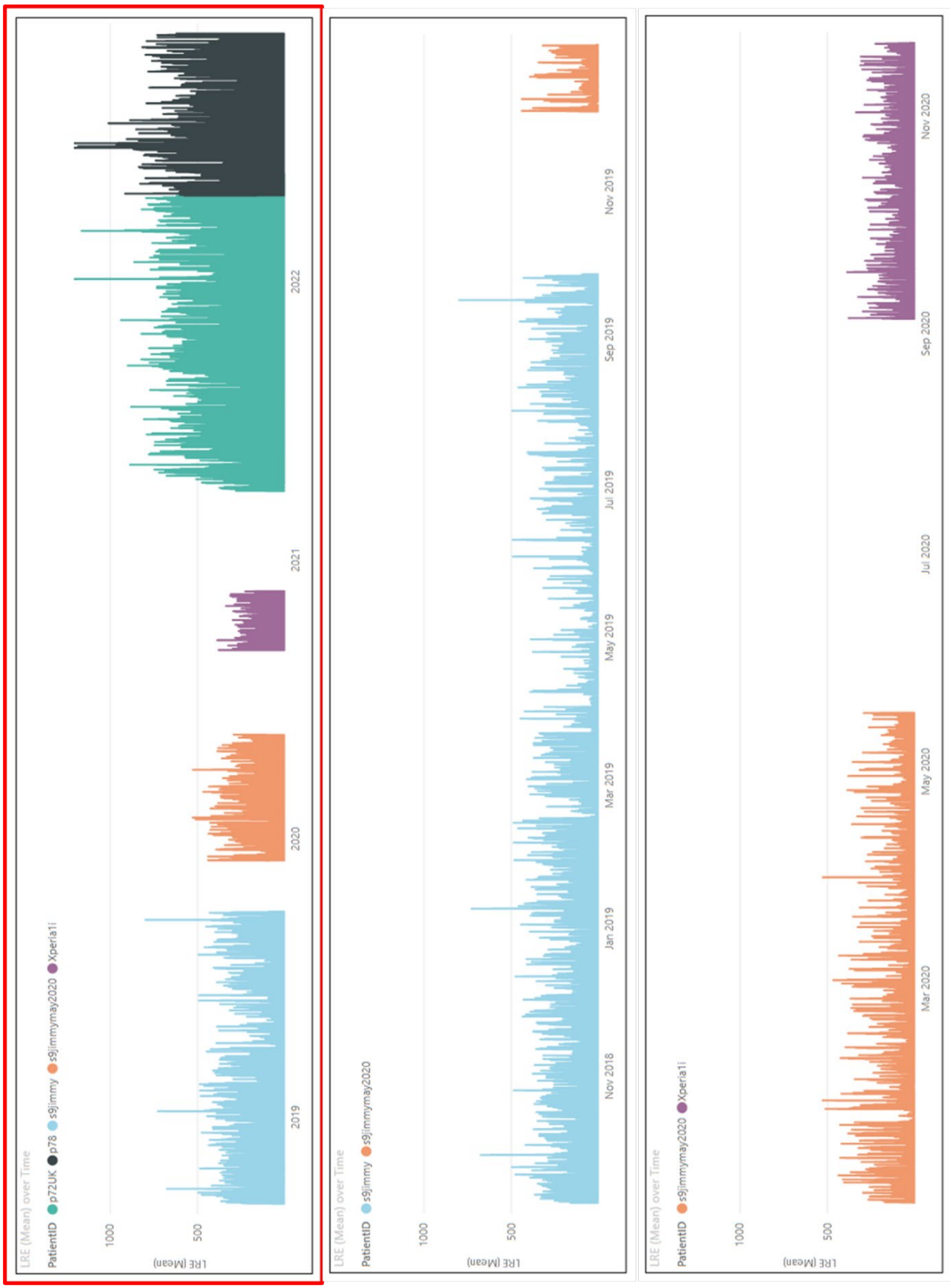


Figure 79: Smartphone's LRE (Mean) over time (2018-2020)

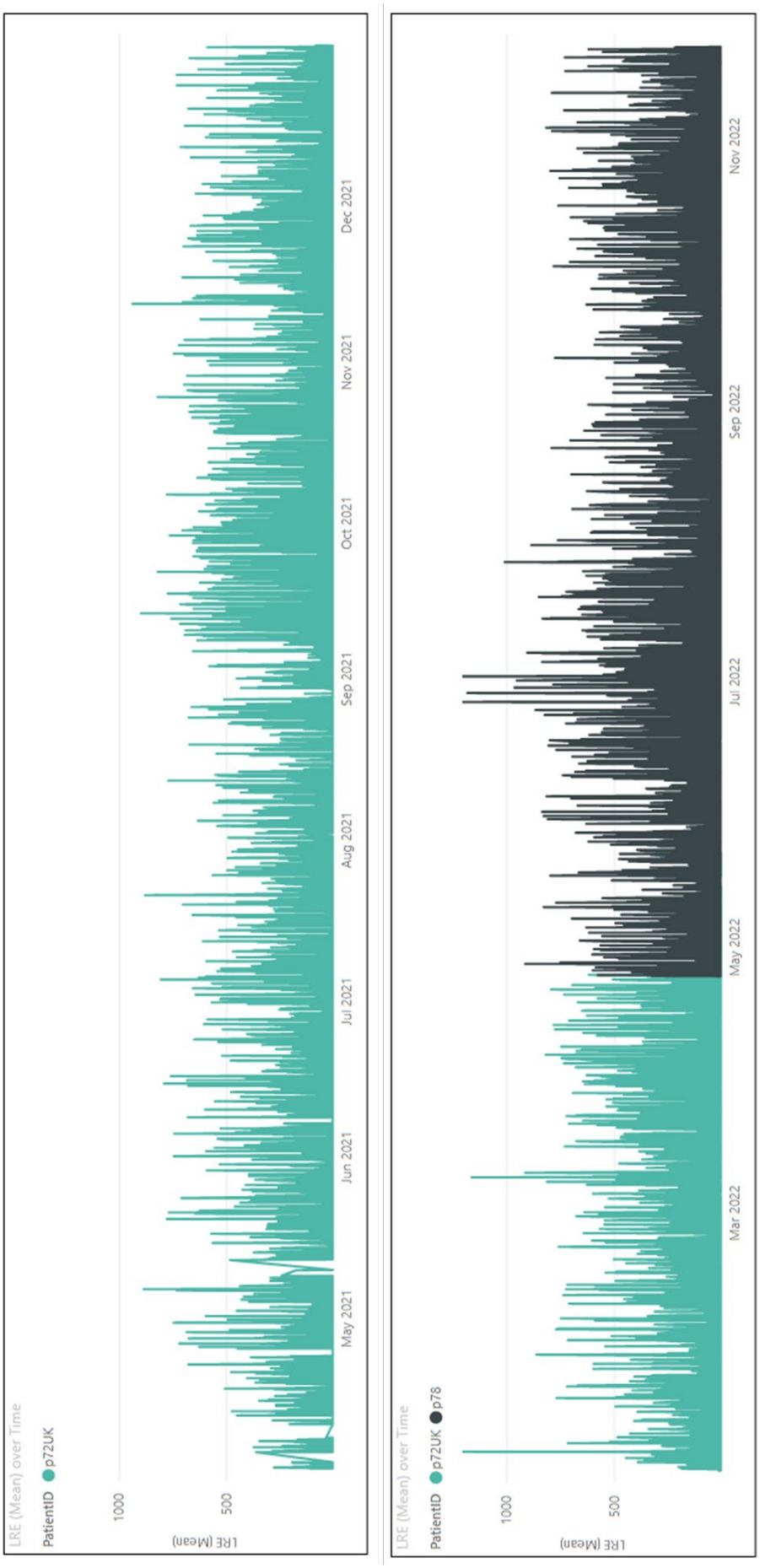


Figure 80: Smartphone's LRE (Mean) over time (2021-2022)

The smartphones were used as primary device with default setting. Fig 81 shows that more amounts were captured across the devices between 23h00 and 08h00. This is because the devices were charging, which disables Android power saving. However, the sample rate remained unaffected, so all samples collected were captured with the expected window size. Fig 82 shows the LRE recorded over time which, as expected, is higher during the daytime than at night. Fig 83 shows that interpolated samples flatten the overall magnitude of LRE.

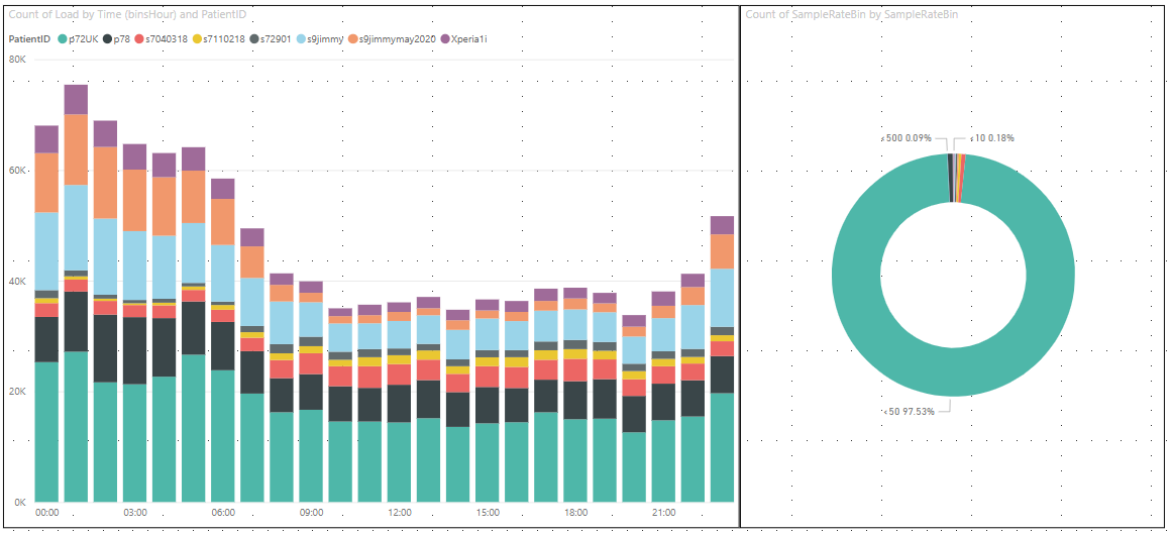


Figure 81: Count of LRE per hour

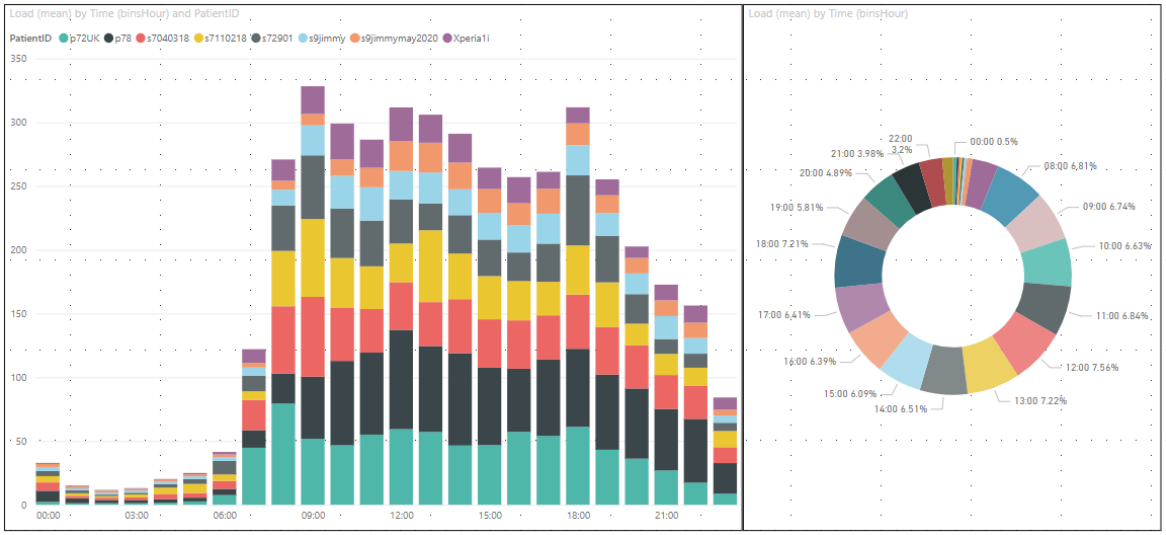


Figure 82: LRE (mean) per hour

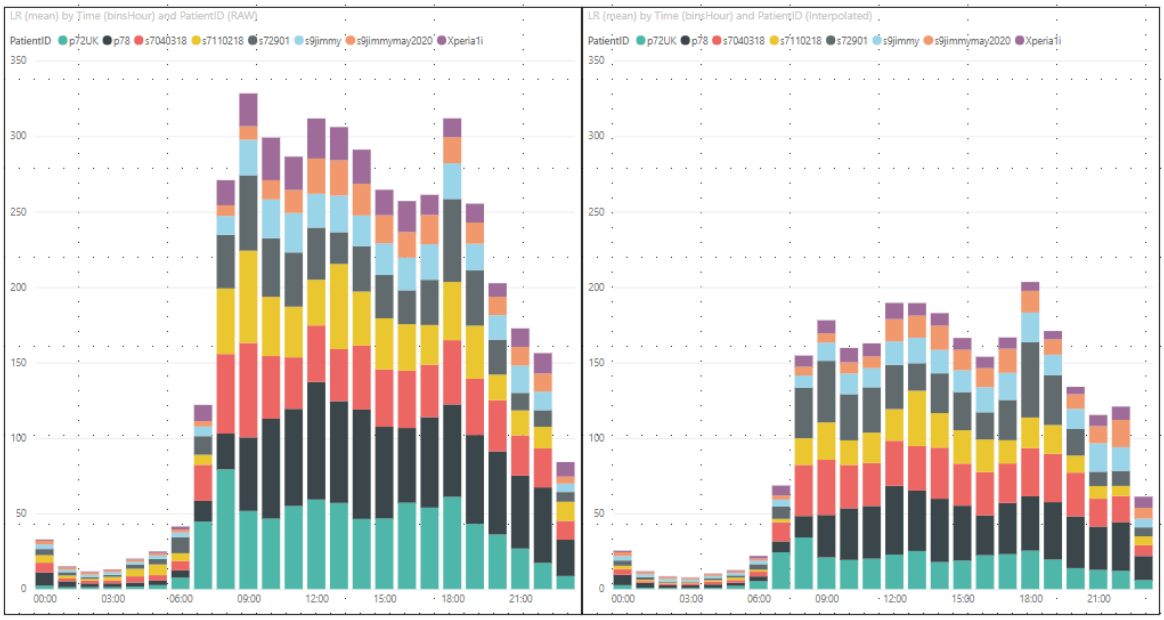


Figure 83: LRE (mean) per hour (raw and interpolated)

Fig 84 shows the yearly phenotype between 2018 and 2022. The CMS was the highest in 2022 and lowest in 2019, consistent with the observations in figs 74 to 78. The peak remains the same over the years but the slope changes. Specifically, 2019 and 2020 were the least active years corresponding to Covid restrictions and lockdown. 2022 shows a similar level to 2018, which is consistent with the timeframe pre and post-Covid.

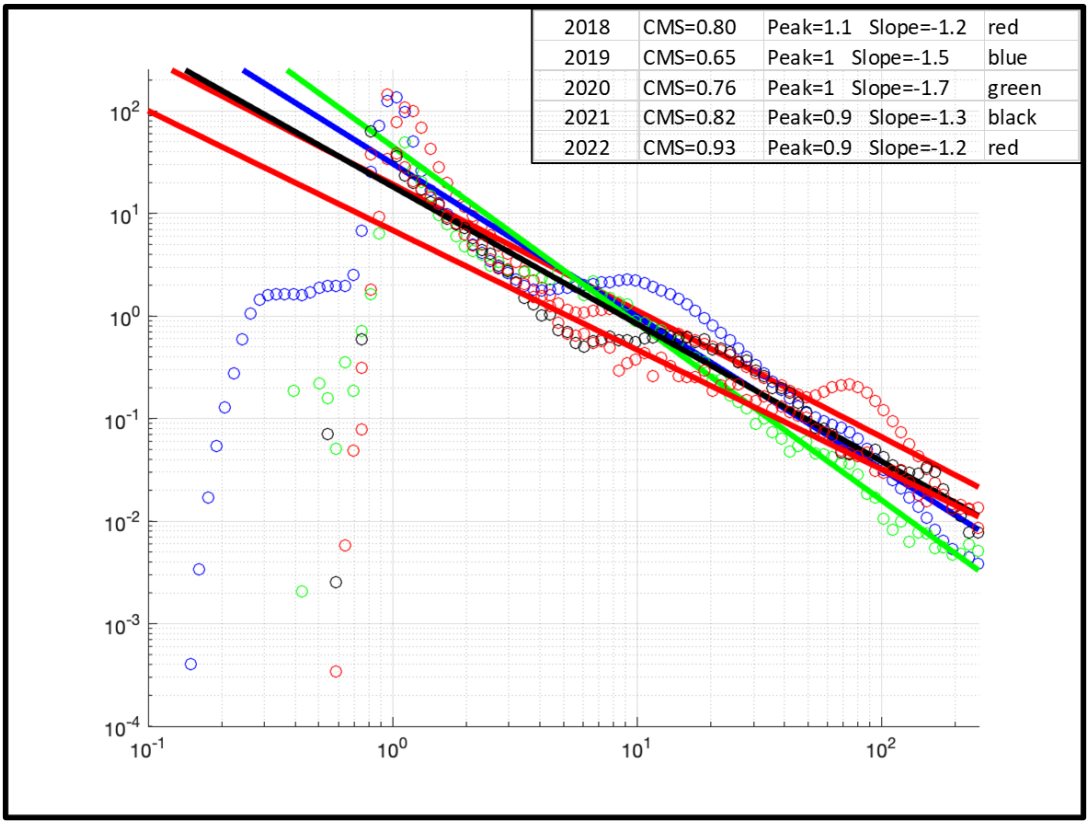


Figure 84: Year on Year Phenotype (2018 to 2022)

Fig 85 shows that the LRE (mean) is consistent across the years but with a higher maximum LRE than the following years, indicating that more dynamic activities were performed. The outliers reach higher values in 2021 and 2022 than in 2018. This might indicate a higher burst of jerk occurring post than pre-Covid.

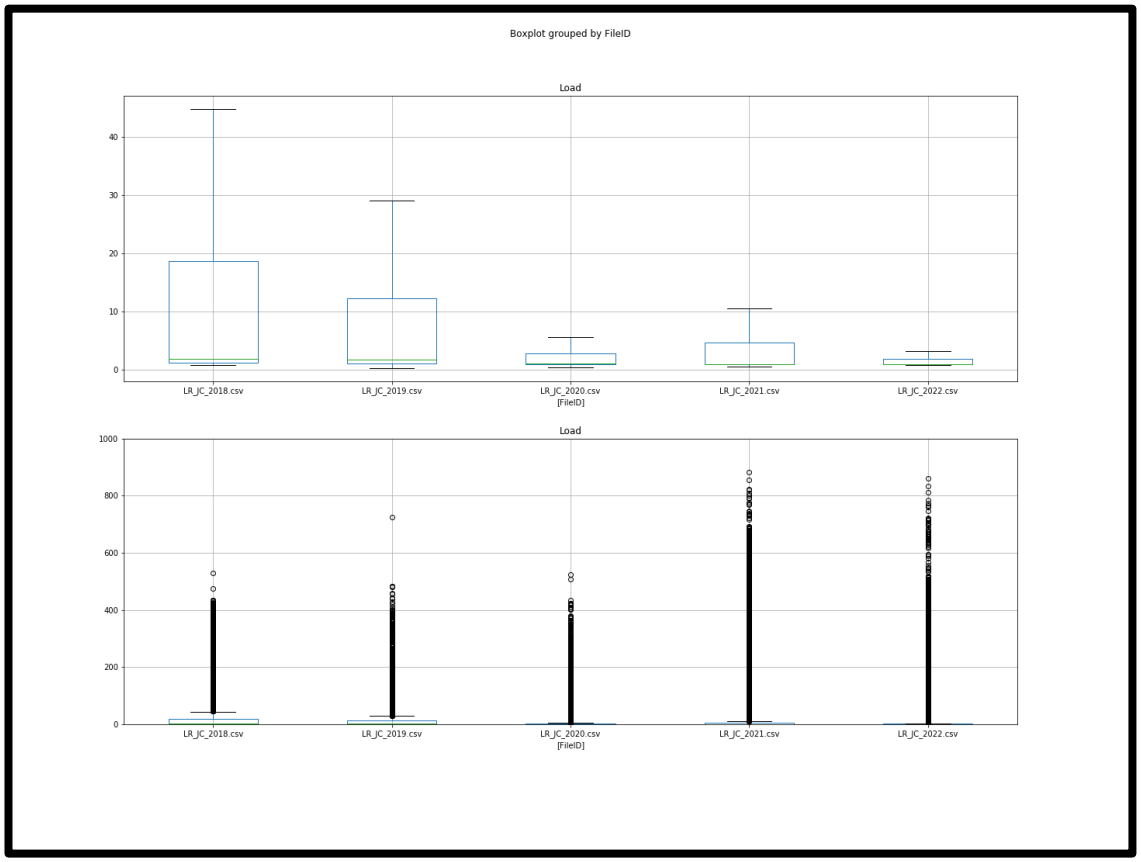


Figure 85: Boxplot Year on Year (2018 to 2022)

Table 17 shows the monthly phenotype characteristics over the years. Most months had CMS greater than 0.7. So, the peak and slope values can be compared for most of the timeline except April 2018, which has a CMS of 0.33. The peak is consistently at one between 2018 and May 2020. A lower peak of 0.9 can be seen from May 2021 to April 2022. The second half of 2022 (May to November) recorded the highest movement, reaching 2.1 peaks. The slope shows a higher level of physical activity in early 2018 than in later 2022. These variations of slope and LRE recorded for the highest proportion of time are consistent with the timeline pre, post and during Covid.

2018/01	CMS=0.73	Peak=1.2 Slope=-1.2
2018/04	CMS=0.33	Peak=1.1 Slope=-1.1
2018/09	CMS=0.49	Peak=1 Slope=-1.3
2018/10	CMS=0.66	Peak=1 Slope=-1.4
2018/11	CMS=0.58	Peak=1 Slope=-1.5
2018/12	CMS=0.54	Peak=1 Slope=-1.6
2019/01	CMS=0.54	Peak=1 Slope=-1.5
2019/02	CMS=0.77	Peak=1 Slope=-1.5
2019/03	CMS=0.85	Peak=1 Slope=-1.3
2019/04	CMS=0.86	Peak=1 Slope=-1.6
2019/05	CMS=0.90	Peak=1 Slope=-1.9
2019/06	CMS=0.84	Peak=1 Slope=-1.5
2019/07	CMS=0.89	Peak=1 Slope=-1.6
2019/08	CMS=0.82	Peak=1 Slope=-1.5
2019/09	CMS=0.75	Peak=1 Slope=-1.4
2019/12	CMS=0.80	Peak=1 Slope=-1.6
2020/01	CMS=0.79	Peak=1 Slope=-1.6
2020/02	CMS=0.66	Peak=1 Slope=-1.8
2020/03	CMS=0.71	Peak=1 Slope=-1.8
2020/04	CMS=0.78	Peak=1 Slope=-1.9
2020/05	CMS=0.71	Peak=1 Slope=-1.8
2020/09	CMS=0.89	Peak=0.5 Slope=-1.3
2020/10	CMS=0.97	Peak=0.4 Slope=-1.5
2020/11	CMS=0.97	Peak=0.4 Slope=-1.6
2021/05	CMS=0.90	Peak=0.9 Slope=-1.4
2021/06	CMS=0.96	Peak=0.9 Slope=-1.3
2021/07	CMS=0.91	Peak=0.9 Slope=-1.1
2021/08	CMS=0.83	Peak=0.9 Slope=-0.9
2021/09	CMS=0.85	Peak=0.9 Slope=-1.1
2021/10	CMS=0.81	Peak=0.9 Slope=-0.9
2021/11	CMS=0.89	Peak=0.9 Slope=-1.1
2021/12	CMS=0.92	Peak=0.9 Slope=-1.1
2022/01	CMS=0.93	Peak=0.9 Slope=-1.2
2022/02	CMS=0.86	Peak=0.9 Slope=-1.1
2022/03	CMS=0.88	Peak=0.9 Slope=-1.1
2022/04	CMS=0.63	Peak=0.9 Slope=-1.1
2022/05	CMS=0.95	Peak=2.1 Slope=-1.5
2022/06	CMS=0.90	Peak=2.1 Slope=-1.2
2022/07	CMS=0.93	Peak=2.1 Slope=-1.2
2022/08	CMS=0.83	Peak=2.1 Slope=-1.5
2022/09	CMS=0.86	Peak=2.1 Slope=-1.5
2022/10	CMS=0.92	Peak=2.1 Slope=-1.5
2022/11	CMS=0.71	Peak=2.1 Slope=-1.5

Table 17: Monthly phenotype characteristics

7.5 Discussion

The sampling for this study was set to capture 5 sec samples (window=5) every 15 seconds (interval=15), which is equivalent to 20 sec per minute. Only one patient (33%) appeared able to monitor uniform LRE samples across the recruitment period. Still, no issues were raised installing RApp™ and interaction with the app was minimal, as intended by design. The other two patients (66%) recorded only a partial amount of LRE samples with significant gaps across the recruitment period. Nevertheless, the samples were recorded with the expected sample rate and high CMS. This indicates that Android power save might have interrupted long-term recording but didn't affect the LRE sampling that occurred. Unlike study 3, the sampling and CMS observed by the researcher showed consistent sampling with high CMS on multiple Android smartphones swapped over the years. These results indicate that power saving alters the recording of samples but are lessened when smartphones are used as primary device.

Fitbit promotes performing 10k steps per day which offer a simple way to compare the trends over time. Unlike daily step counts, comparing physical activity using LRE is more complex and looking only at daily LRE aggregates would be misleading. The relation between LRE

characteristics should not be used as isolated metrics and instead be looked at in context. For instance, the peak shows the LRE most recorded, which could identify p74 as being sedentary compared to p73 and p75. However, it might also be due to smartphone recording more samples on charge and placed on a table. The CMS of p73 is lower than p74 and appears close to p75´s CMS, but p73 recorded a broader timeframe than the others. So, the CMS accurately evaluates that p74 was the closest to capturing the expected amount of samples but only on specific time slots rather than the entire period. Therefore, CMS comparison between patients should be considered alongside the same timeframe. Similarly, using the slopes to compare the trend of physical activity should be done on a common timeframe to be usable.

Moreover, patient differences might also occur because of habits and personal ways of handling their phones. This concern is alleviated when comparing LRE within the same patient. The LRE phenotype captured by the researcher allows us to identify patterns and changes over the years. For instance, the peak gradually decreases from 1.1 in 2018 to 0.9 in 2022. Although the CMS fluctuate and amount of months logged varies over the years, this decrease can be considered an accurate representation. This indicates that the slope can be interpreted without the need to impute data and shows that 2018 and 2022 were more active years than 2019 and 2020. In this case, the year-on-year variations can easily be associated with pandemic restrictions and lockdowns caused by Covid in the UK. Therefore, it is possible to identify drastic changes in physical activity trends caused by changes in disease activity.

7.6 Conclusion

This study confirms that smartphones used as primary devices alleviate the impact of power saving in obtaining LRE. In free-living conditions, there is a strong dependency on the way that patients handle and interact with their phones. LRE obtained do not reach the performance in ideal conditions seen in study 2 but are more reliable and accurate than in study 3.

The phenotype and its characteristics provide a methodology to compare LRE patterns and physical activity among patients, but it is far more effective when comparing for an individual. A trend evaluated over three months does not show significant physical activity changes and cannot be correlated with disease activity. However, an impactful event, such as the lockdown restrictions caused by Covid, is noticeable over an LRE monitoring timeframe of 5 years.

Chapter 8 Conclusions and future research

8.1 General Discussion

This thesis developed, tested and verified a methodology to record load rate estimates using smartphone sensors continuously and remotely. The app's ability, RApp™, was tested and validated at the project's end to record load rate estimates on cloud-based services. The methodology developed was able to interpret these estimates to draw individual characteristics of physical ability, which can be used in further research.

In Chapter 2, the value for monitoring load rate estimate was explained in the context of healthcare and musculoskeletal disease. The benefits of being active and exercising are already known, and metrics such as step count and MET-min help quantify physical activity as aggregated. Evaluating the impact of LRE is more complex by nature, as too much, too less, or sudden bursts could lead to detrimental effects on the joints and bones. This project's novelty is using smartphones, which provide cheap and broadly available technology, to monitor LREs outside of specialised rehabilitation and clinical facilities continuously.

In Chapter 4 (study 1), rheumatoid arthritis patients were engaged to gather feedback and validate the approach and challenges associated to using technology in healthcare. The insight obtained through the PPI session validated and expanded the design of RApp™ to being a smartphone app that anyone could use rather than a specialised clinical tool. The challenges in recording and validating LRE were raised and used to design the subsequent studies. It was also possible to conclude early on that drawing correlation with disease activity is associated to the patient´s symptoms (pain and swelling). These cannot be enforced or controlled as this project is observational rather than a medical intervention.

In Chapter 5 (study 2), the smartphone´s ability and performance to record LRE were assessed in ideal usage conditions and compared to smartwatches. Smartphones were set up with power save disabled to estimate the impact on the sampling. As seen in study 1, Fitbit wristbands focus on estimating daily step counts and do not permit access to raw accelerometer data. So, smartwatches were used as the baseline of LRE comparison at the wrist to validate LRE obtained from smartphones. The concept of phenotype was introduced and tested to evaluate LRE´s characteristics: the trend, CMS, slope and peak.

In Chapter 6 (study 3), smartphones were used in free-living conditions as part of a clinical trial with osteoarthritis patients. Considering the lack of a benchmark to measure LRE, Fitbit were once more used to validate smartphone monitoring but considering daily step count as a more broadly

known and used unit of measure. The trial consisted in giving a knee injection to OA patients who acted as a trigger to physical activity changes expected to be seen in disease progression. The results showed a correlation with Fitbit in estimating daily step count. However, a smartphone´s monitoring performance is substantially affected when used as a secondary device. The interval between samples recorded and the sample rate (i.e., sample size) was significantly lower than expected.

In Chapter 7 (study 4), smartphone LRE monitoring was evaluated on the primary device of rheumatoid arthritis patients and the researcher for longer timeframes than in previous studies. The methodology confirmed that the phenotype interpretation of LRE is valid, but comparing patients' trends is not straightforward. LRE interpretation needs to consider the characteristics in the overall context rather than aggregated and isolated values. In that sense, comparing LRE variations of a patient is possible and allows us to see changes that occur over time. The timeline coincided with the Covid pandemic and lockdown restrictions in the UK, indicating that significant physical activity changes caused by disease or other factors could be perceivable.

8.2 Research contribution and future work

The work presented in this thesis focused on evaluating smartphones´ ability to monitor and remotely record LRE continuously. Smartphones are not designed as wearable devices, and the user handling biases the sampling. OApp™ has been used in another PhD thesis to develop an activity classifier (Nazirizadeh, 2018). More studies are needed to quantify the amount of samples needed to assess LRE, and statistical models should be further evaluated accurately. For example, a large amount of data might be used in machine learning models, e.g., to impute missing LRE based on historical values recorded. This project was explicitly focused on smartphone´s performance when used as an individual device. Other smartphone sensors and activity data might provide further insights into the smartphone´s handling (e.g., on call or playing games), which could be used to improve LRE accuracy further. The CMS provides confidence in the sample reliability, but data captured from other wearable devices could be used to cross-check and improve LRE accuracy further. Besides, smartphones were used as a surrogate for the lower limbs but are not always worn at the hips. Since the first introduction of smartwatches and wristbands, a considerable amount of smart devices has been made commercially available, in data collection and connected care use cases, including smart in-sole and textiles (Vijayan et al., 2021). Smartwatches, wristbands and other IoT devices might allow subtracting the noise generated by arm swings and upper limbs activity.

Furthermore, energy expenditure, step count and MET-min provide a known unit of measure for a range of activities that have been well documented (Mancuso et al., 2007) (White et al., 2016). The measure of load rate is not standardised, and further studies might be able to map LRE magnitude to these metrics. Maintaining a diary is recommended to identify patterns of physical activity associated with seasonality (e.g., Christmas, summer holidays, travel, etc..). The questions presented in study 1 were reviewed with TRIFoRM (Hooper et al., 2015), which evaluates the psychological impact of using technology (Hooper et al., 2015). The dashboard of platforms like Strava and Fitbit has shown that the gamification (i.e., receiving trophies) and overall layout can motivate in performing more physical activity. This project was observational, so it was decided not to make the LRE analytics available to patients. More studies should be made on the impact of physio-promoting exercise and making recommendations based on remote monitoring and integrated dashboard as used in the Remora project (Dixon and Michaud, 2018).

British Society of Rheumatology accepted the project's abstract (Study 1) for a poster presentation in Birmingham in 2017 (see Appendix G). The overall principle of the project received very positive feedback, precisely as no other known projects are executing similar scope, with a focus on patients with arthritis. Following the event, the project has been referenced in the context of smartphone apps that may aid home rheumatoid arthritis monitoring (Freeman, 2017). The project has also been referenced, amongst the CRF studies, on Southampton General Hospital website and attracted interest for recruitment (Southampton, 2016) with patients reaching out to take part as well as for general interest, notably by a consultant for Arthritis research Australia that would like to introduce the project in Australia. Study 3 is a joint effort with the University of Sydney, with results submitted to OARSI 2020 world congress and published (Yu et al., 2022). This project aimed to correlate physical activity with disease activity in the context of musculoskeletal disease. Specifically, passive monitoring has shown that bursts of LRE occur at a magnitude up to 800 times higher than the mean LRE. Further studies should be done to evaluate the significance of musculoskeletal conditions and whether these represent micro fractures in the bone structure leading towards osteoarthritis.

Finally, the methodology of this project was applied to a study to monitor LRE in the context of occupational activity, and a paper has been published using OApp™ (Wang et al., 2021). Studies have been made tracking weather and disease severity with smartphones (Reade et al., 2017). Therefore, studies beyond healthcare use cases could use passive smartphone monitoring e.g., to evaluate patterns in sleep, and digital dependencies.

8.3 Summary

In summary, this project produced novel contributions in providing a methodology for monitoring load rate estimates using smartphones. The results of this thesis have shown a means to interpret load rate estimates that might support patients with musculoskeletal disease and other use cases. The app developed is now available on Google Play and is considered for commercialisation through a spinout with the University of Southampton. Further studies are needed, but the protocol has already been applied and published in the context of osteoarthritis and occupational paper with the key benefit of providing remote monitoring.

Appendix A PPI Questionnaire

1.1 Q: When you see your rheumatologist, can you recall how active your arthritis has been during the period since your previous visit? How would your rate this from 0 to 5?

A: 0 – cannot remember at all

1 – can recall the past week in detail

2 – the past month in detail

3 – the past two months in detail

4 – the past three months in detail

5 – can recall the whole period in detail

1.2 Q: When you see your rheumatologist, can you easily list areas of pain and inflammation during the period since your previous visit? How would you rate this from 0 to 5?

A: 0 – cannot remember at all

1 – can recall the past week in detail

2 – the past month in detail

3 – the past two months in detail

4 – the past three months in detail

5 – can recall the whole period in detail

1.3 Q: Would you like your rheumatologist/GP to monitor your overall status (including pain and physical activity) between visits?

A: Yes Additional comments:

No

Not sure

1.4 Q: Are you interested in how your physical activity (e.g., amount of walking) may influence your arthritis?

A: Yes Additional comments:

No

Not sure

2.1 Q: Do you currently have a smartphone (the type of phone with a touch-sensitive screen)?

A: Yes

No

Not sure

2.1.1 Q: If not, would you like to have one?

A: Yes

No

Not sure

2.2 Q: Which type?

A: iPhone

Android

Other

Not sure

2.3 Q: Do you usually have your phone/smartphone with you?

A: 0 – never

1 – only pick it up to answer calls

2 – sometimes/not sure

3 – most of the time, e.g. when at work

4 – whenever you are out of the house

5 – all the time

2.4 Q: How big is your phone/smartphone?

A: Please see annexes

2.5 Q: How often do you use smartphone apps to manage your everyday life (e.g. calendar, email, exercise, calorie counter)?

A: 0 – never

1 – very unlikely

2 – unlikely

3 – neutral/not sure

4 – likely

5 – very likely

2.6 Q: If you do use your smartphone in that way, which apps do you use? Please list them here.

A:

2.7 Q: Do you think a smartphone app could support the treatment of your arthritis? How would you rate this from 0 to 5?

A: 0 – never Additional comments:

1 – very unlikely

2 – unlikely

3 – neutral/not sure

4 – likely

5 – very likely

2.8 Q: Would you be willing to use a smartphone app to support the treatment of your arthritis? How would you rate this from 0 to 5?

A: 0 – never Additional comments:

1 – very unlikely

2 – unlikely

3 – neutral/not sure

4 – likely

5 – very likely

2.9 Q: Would you be willing to use a smartphone app, which continuously monitors and records your physical activity in relation to your arthritis (e.g. whether you are sitting or walking)? How would you rate this from 0 to 5?

A: 0 – never Additional comments:

1 – very unlikely

2 – unlikely

3 – neutral/not sure

4 – likely

5 – very likely

2.10 Q: To record your physical activity, you would need to be carrying your smartphone. Would you be prepared to do this?

A: 0 – never Additional comments:

1 – very unlikely

2 – unlikely

3 – neutral/not sure

4 – likely

5 – very likely

2.11 Q: To support the management of your condition, you would also need to input information about levels of pain and activity. How likely would you be prepared to do this?

A: 0 – never Additional comments:

1 – very unlikely

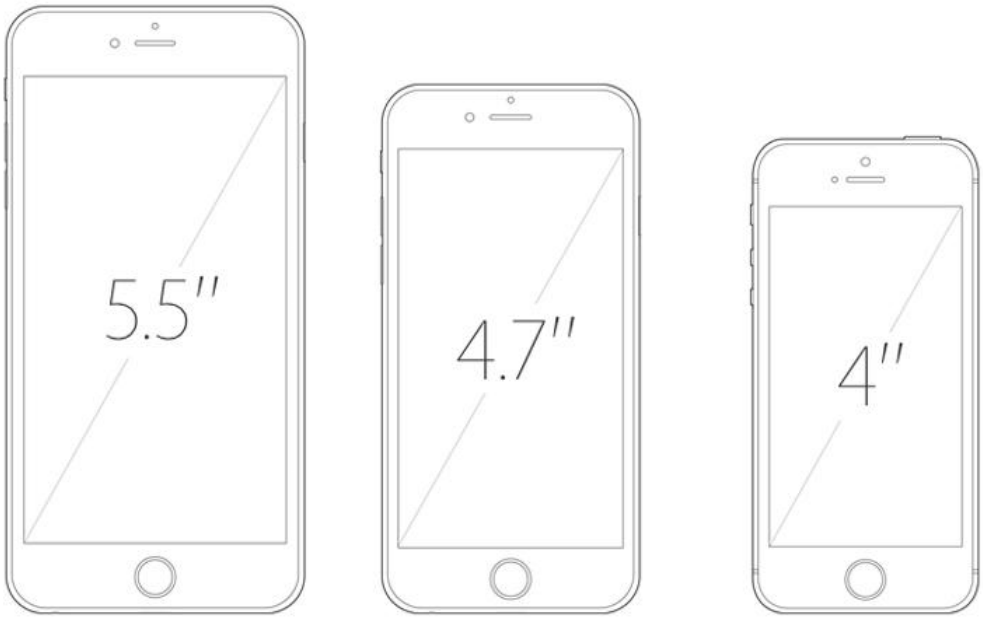
2 – unlikely

3 – neutral/not sure

4 – likely

5 – very likely

Annexe: Screen size comparison (Q 2.1.3):


5.5''
4.7''
4''

Appendix B PPI survey summary

Summary		
Project	Continous Monitoring Patient Assessment Rhematoid Arthritis	
Scope	Survey provided to fill prior to PPI session	
Date	11th Feb 2015	
Participants	Jo, Chris, Susan and Jimmy	
PPI	9 participants	
Survey results	Survey Results'	
Questions	Answer	Summary Answer
CURRENT TREATMENT QUESTIONS		
1.1 When you see your rheumatologist, can you recall how active your arthritis has been during the period since your previous visit? How would your rate this from 0 to 5?	0 - cannot remember at all	
	1 - can recall the past week in detail	1
	2 - the past month in detail	3
	3 - the past two months in detail	
	4 - the past three months in detail	1
	5 - can recall the whole period in detail	4
1.2 When you see your rheumatologist, can you easily list areas of pain and	0 - cannot remember at all	
	1 - can recall the past week in detail	
	2 - the past month in detail	4
	3 - the past two months in detail	

inflammation during the period since your previous visit? How would you rate this from 0 to 5?	4 - the past three months in detail	
	5 - can recall the whole period in detail	5
1.3 Would you like your rheumatologist/GP to monitorÂ your overall status (including pain and physical activity) between visits?	1 - Yes	6
	2- No	2
	3 - Not sure	1
1.4 Are you interested in how your physical activity (e.g., amount of walking) may influence your arthritis?	1 - Yes	8
	2- No	
	3 - Not sure	1
APP QUESTIONS		
2.1 Do you currently have a smartphone (the type of phone with a touch-sensitive screen)?	1 - Yes	8
	2- No	1
	3 - Not sure	
2.1.1 Would you like to have one?	1 - Yes	
	2- No	
	3 - Not sure	1
2.2 Which type?	1 - iPhone	3

	2- Android	5
	3 - Other	1
	4 - Not sure	
2.3 Do you usually have your phone/smartphone with you?	0 - Never	
	1 - only pick it up to answer calls	
	2 - sommetimes/not sure	
	3 - most of the time	
	4 - whenever you are out of the house	4
	5 - all the time	5
2.4 How big is your phone/smartphone ?	5.5 inches	4
	4.7 inches	2
	4 inches	1
	N/A	1
2.5 How often do you use smartphone apps to manage your everyday life (e.g. calendar, email, exercise, calorie counter)?	0 - never	1
	1 - very unlikely	
	2 - unlikely	1
	3 - neutral/not sure	
	4 - likely	2
	5 - very likely	5
2.6 If you do use your smartphone in		

that way, which apps do you use? Please list them here.		
2.7 Do you think a smartphone app could support the treatment of your arthritis? How would you rate this from 0 to 5?	0 - never	
	1 - very unlikely	
	2 - unlikely	
	3 - neutral/not sure	5
	4 - likely	3
	5 - very likely	1
2.8 Would you be willing to use a smartphone app to support the treatment of your arthritis? How would you rate this from 0 to 5?	0 - never	
	1 - very unlikely	
	2 - unlikely	
	3 - neutral/not sure	4
	4 - likely	1
	5 - very likely	4
2.9 Would you be willing to use a smartphone app, which continuously monitors and records your physical activity in relation to your arthritis (e.g. whether you are	0 - never	
	1 - very unlikely	
	2 - unlikely	1
	3 - neutral/not sure	2
	4 - likely	2
	5 - very likely	4

sitting or walking)? How would you rate this from 0 to 5?		
2.10 To record your physical activity, you would need to be carrying your smartphone. Would you be prepared to do this?	0 - never	
	1 - very unlikely	
	2 - unlikely	1
	3 - neutral/not sure	1
	4 - likely	2
	5 - very likely	5
2.11 To support the management of your condition, you would also need to input information about levels of pain and activity. How likely would you be prepared to do this?	0 - never	
	1 - very unlikely	
	2 - unlikely	
	3 - neutral/not sure	1
	4 - likely	3
	5 - very likely	5

Appendix C Clinic shadowing notes

Notes from shadowing Southampton General Hospital's Monday clinics.

> Reassurance and Communication is key with patients,

- They feel most vulnerable when at doctor appointment
  > Women that spoke about her divorce, husband having an affair when not asked.
- Patient come to hospital are usually worried and want to leave with a feeling that they've been taken care of and reassured – not with doubt
  > Young girl had her MRI ok, not much pain but complaining of random right knee pain (chronic pain)
- Another one following knee pain surgery got reassured by being explained and walked through. Happy to avoid MRI once joint examination completed
- Once
- Reassure can be done by testing even when there is nothing

> Many patients routed to RA visit even when not needed

- Need to check if biomechanical rather than inflammatory
  - Patient that had surgeries when shouldn't have. When still in pain, ended up in RA department
  - Women in her 50s suffering of back pain, shoulder and neck pain. Should go to physio as more exercises and a pillow might help
- Need to check if other disease related (diabetes, overweight)
  > patient blood sugar checked – no inflammatory but blood sugar fluctuates

> Routine visits

- when no specific disease activity – however this reflects current state rather than full period
- next visit in 6 months
- routine visit in 1 year

> Medicine and disease awareness

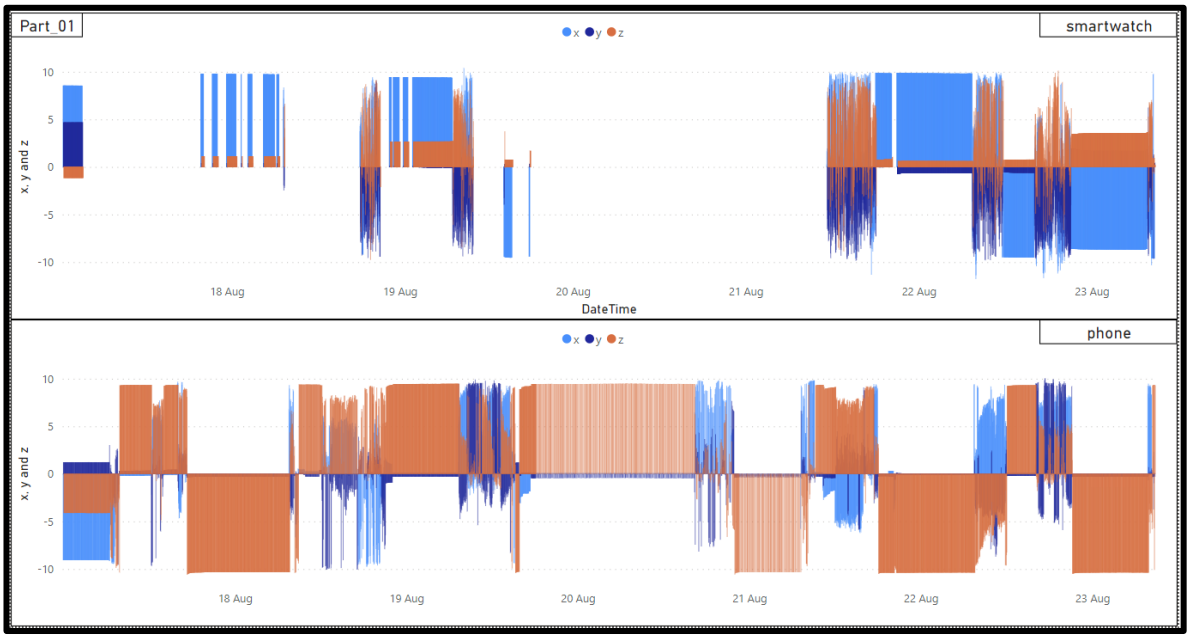
- patient there for her daughter (uncapable to stand for herslf – incapacitated by JIA arthritis) knew exactly what her daughter tried and
- husband supporting his wife was answering the questions
- medication is always discussed to reach an agreement
  - should we experiment a bit more or less?
  - Should we try a different injection?
  - Patient that wanted to stay on tablet rather than injections due to potential side effect (hair loss)

Appendix D Study 1: Patient feedback

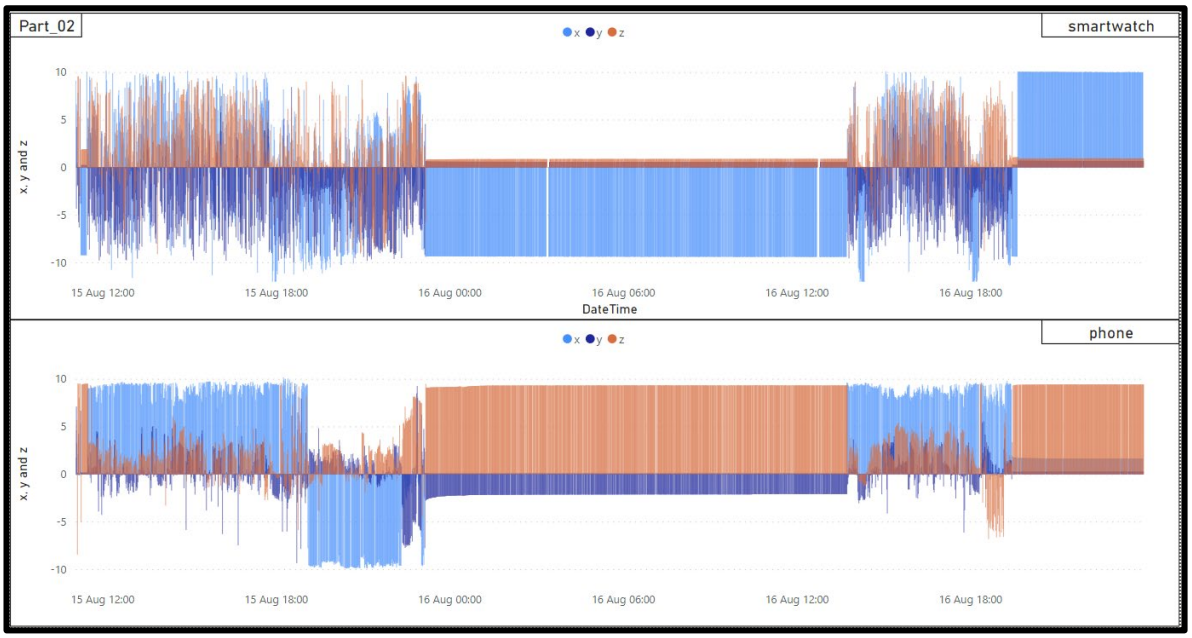
Patients id	001		002		003		004		005	
Visit 1 (Patient's introduction/recruitment)										
1) Introduction, instructions and tutorial on using the smartphone and RApp										
2) How easy does RApp appear to use?	1/10	it seems ok	3/10	seems ok	0/10	Very easy.			7/10	Very easy. Questions are not clear.
3) Do you think using RApp could affect your Rheumatoid Arthritis and its treatment in anyway?	0/10	No	0/10			No				In the future, possibly.
4) How do you feel about using RApp?	2/10	Apprehensive of using it	10/10	Good as could help other people (being part of a research)		No worries, even if not for me it's still beneficial.				Up for it, good, would be better if could use on any phone.
5) Generally how confident are you that Internet based technologies actually do what they claim to do?	7/10	Love using internet and can find anything with Google.	8/10	Confident	5/10	that they do what they claim.			10/10	Very confident. The way it will be in the future.
6) Generally how helpful do you find Internet technologies in supporting you to achieve your goals?	8/10	Very good		Don't use muh tech, know it's there if need it, amount used would be 5/10. Don't rely much on tech but likes it when need it.	7/10	Not too used, Satnav and FB.			10/10	Extremely helpful.
7) Generally how predictable do you think Internet technologies in carrying out what you expect them to do?	8/10	Very good	9/10	High expectations of technology	7/10	Very reliable			10/10	Very predicatble.
Other comments:	Uses Samsung		Iphone 5 for text and emails		Used to have Xperia Z5 compact (ie same) then bought X5. Uses the Xperia as primary (put personal Sim and provided with a cover)		- The app is missing hips and neck joints.		- Questions: Maybe add a bubble with clarification of what they mean (similar to hospital forms). - Titel of P&I could change color of P&I.	
Visit 2 (Patient follow up)										
Other comments:	- Fitibit is encouraging to do 10k steps - Does more in the morning, tired in the afternoon - Would like to see the reporting even when feel good. For eg, it is rewarding to see when doing 5k steps		- Iphone has already an health app formonitoring steps. Fitbit seems to track more steps (806 through iphone vs 1251 for Fibit) but wear Fitbit longer. - Could fill the form every day. - No logs for period of 3rd to 21st of Sept 2016 but happy to extend participation in the study.		- Would use as ok for now. - Not too intensive. - Just bought a new phone but will stick to Xperia - When switch between P&I, it should be 0 (not the last value entered). - The dates should show the reports.		- The scale goes back to previous screen. - Questionnaire summary, swipe to go back takes to the beginning. - P&I, enter pain then tick does not go to Inflamation. - Missing joints: Ankle, hips, neck, back. - Questions (RAPID3): "physical well-being". Label "without any", difficulty… so scale is not clear. Should be on the scale. - Skeleton: which is left or right (is it mirrored image of you?). Maybe highlight if left or right. Click on the joint is nice but should be clearer. - P&I: 0 to 5 is nice to record.		- Not complicated. - The weekly questionnaire doesn't seem related to RA. - Phone keeps running out of battery (1.5 day), sometime dead battery. - Plays drum 3 times per week (for 2-3 hours). Could affect thte fitbit reading. Don't keep the phone.	
Visit 3 (Patient's closing review)										
1) How easy was it to use RApp?	1/10	Easy once explained.	1/10	Very straigth forward.	0/10	Easy	0/10	Easy	8/10	Easily used. P&I: doesn't always work when press the buttons. Q: Easy to do. Never drag the VAS, just tap. Difficult and hard to know how to address, likely to always be the same. Knee was really bad in that period. If was run by specialist, would listen the recomendation in form filling. - Liked RAPID3 scoring better. Easier to do than 0 to 10 as 0 to 10 is very difficult to know how to gage. 10 implies disability, 7 hurts but can live with it, it depends on people. Scale from "without any difficulty" to "unable to do" is more useful for patient perspective. Explanation is very important to know how to answer.
2) How easy was it to understand RApp?	2/10	Questions every day is too much. Long questionnaire 3 times per week would be better. Same for P&I but to keep open for when flare happen. Using became part of routine and was not hard to do, takes 2 min.	1/10	Self explanatory. Flows on its own. Very easy. Not complicated in any ways.	0/10	Easy but reports not always working. Good to have: to see the difference after stopped the medication (2 weeks before) as had MTX stopped for 2 weeks then back on lower doses. Could see the trend. Not used when was in Venice.	8-9/10		8/10	Questions were not that easy.
3) Did using RApp make you feel positive about your Rheumatoid Arthritis?	8/10	Encourage to move more, did lots of steps. Positive as need to move with RA.		Didn't make feel any different.		As positive as can be, it's got its purpose. Would like to see the trend. Ability to look back would be usefull for GP&nurses.		- Not in pain so don't think about pain. So more of a negative because had to think about it while feel well. - Report everydat is too much when no pain. - But if was in pain, it would be a comfort and positive.		Positive. Scale helps put things in contrast.
4) Did using RApp make you feel negative about your Rheumatoid Arthritis?		No, you got to live with it.		Didn't make feel any different.		Not at all.		See above.		Didn't feel negative at all.
5) Could you highlight 5 positives comments?		Tech is the way forward.		Info callected can be good. Not difficult to use.		Good app and good for RA. Ability to monitor. Easy to use and to understand. People would use it. Layout is good and self explanatory for users.		- Easy to use. - Notification would be good for questionnaire, medication. - Reporting (Fitbit style). - Simplified UI would be good.		- Easy to use. - Made feel good, felt as was being recorded, good psychologically. - Valid doctor Q, helps understand RA as a disease. - Light weight app, doesn't take long to load.
6) Could you highlight 5 negative comments?		Can't think of anything negative. No hardship in using RApp.		Nothing negative. Would be good on iPhone. Not all joints are included (for eg, elbow)		Not fully working.		- If not in pain, make think about it. - Enter infos but don't get anything from it.		Xperia is bad, would use it on his phone. Couldn't see results/reporting.
7) Any additional comments?		Battery was draining even without using the phone. Last for about 2 days.		Nothing.		When is it available, please keep up to date when can be used.		Son has diabeteses and uses medical dev ice to record and upload the results. Then goes every 3 months at the clinic. Also self manage a bit by checking once per week.		
8) Would you use RApp again?		Yes		Yes, no personal use but for other people. Could be used to show history to clinicians. Would use it if on App store (even for nominal payment).		Yes		It depends on disease activity, so not really. But would use if newly diagnosed.	10/10	Yes, not without stats to use for doctor recommendations.
9) How well do you think this app achieves your goals?		App side is more directed to clinicians. Need to incorporate more to be patient focus (for eg, include steps).		Don't know how to answer.		Good.		See above.		If there were a doctor involved, it would achieve taking the results.
10) How much do you think you can rely on it?	10/10	Perfect, only problem was upload to server.	10/10	Reliable except for logging. Battery charge lasts only for a day (vs iphone that last 3 days)		No problems, no issues, no crashes.	9/10	Completely	10/10	Never crashed. Logging to re-enable when phone dies.
11) How much do you trust the app's performance?	9/10	Trust that it's doing its job with taking the information. Same as when submiting something online.	7/10	Good. Logging not working and if battery is flat, needs to re-enable.		Haven't seen the report but fine.	9/10	No reasons to questions it. Trustworthy.		See above.
12) How secure do you think the processes behind the app are?	10/10	Very secure I would think. Security is very important. Anything that I don't understand, i don't touch. For eg, don't accept the friends requests from people I don't know.		Fine		Fine, no personal data. If someone wants to see what I've done for a day, that's ok.	9/10	Hadnt thought about it so must mean that trust it. Don't put anything confidential, it doesn't matter who gets hold of it (no credit card details).	10/10	Trust it
13) Do you think anyone could pretend to be you or access your data?		No		People get your data. If nothing to hide, it doesn't matter.		No.		Yes. Anyone in the family could but not sure anyone would want to. There is no PIN or Password.		Nothing personal and believe it's secured. If someone does, doesn't really care.
14) How predictable do you think the app when handling your input or recording your activity?		Very predictable, the questions are always the same.		Fine for recording. Can't see any problems.		Can't put so many different questions.		Don't understand the questions.		Predicatable, questions are the same every day. Mixing might be good. No reward.
15) What are your thoughts about how much you can trust the processes behind the app?		Can't trust completely.		Nothing specifial.		No thoughts.		Hadnt thought about it. Would trust it.		See above.
16) Do you think that the people who might provide the app have your best interests at heart?		Yes, of course. Otherwise would not be able to sell it.		- Yes, if not charged for a lot of money. - If funded by NHS, for well being. - If Apple, for the money. Even if 1gbp, for the money.		Yes. If payable, would use it if tie with a doctor/GP.		Yes, cant think of any other reasions to do it.		Guess so. They want to keep making money. Yes and No.
17) How confident are you that the app will continue to provide benefits in the future?		Yes, only will get better as this is for research.		Very confident.		Confident.		Yes, I think it will after a few teaks: usefull for patients and clinicians. Makes you feel more in control and involved.	10/10	More info is always better.
18) Tell us whether or not you trust the purposes to which the app is being used to help your care.		Yes, trust completely also because done through the hospital. Credibility of Medical body is important.		Yes, definitely to help.		Trust it.		I do trust the purpose.		Knew what was going into and that wouldn't actually being used for the diseases. Don't trust it yet and "benefits you, not me". The bigger picture is that when it goes live, it will help.
Other comments:	- Made walk more - Would definitely use the app if monitoring was included (climb staris..), include heart monitor. - Questions don't mean much, just things that happen. Can't see the point of it. Infos beneficiat to professionals, not patient. - Feels like need to record P&I. For patient, it's just pain. No point in knowing it and would just take a painkiller. - Some people like things monitored but are not really monitoring often.		- Questions not relevant for daily use. Once a week would be more generalised. - Questions might be more relevant for people with higher disease activity. - P&I once per day is ok. Intensity is usefull (1 to 10 would be too much). One of the wirts hurts more than the other and so can record. - On Questionnaire, the scale is the wrong way around.		- Can't remember daily, great for nurses to look at. - VAS scale a bit difficult to use. Want to set it to 0 but when N/A or would log it. N/A means 0. - If app is available in 6 months (ie study 2), would use it if can be used as part of treatment progress or results visible to nurses. - In GP clinics, the form gives an input just for the day, not the last month. When stop the medication, pain can come back > more input. - Happy to use even if not part of the study. - I can see a future in it.				Hates the phone as battery doesn't last.

Appendix E Study 2: SP & SW (x, y and z)

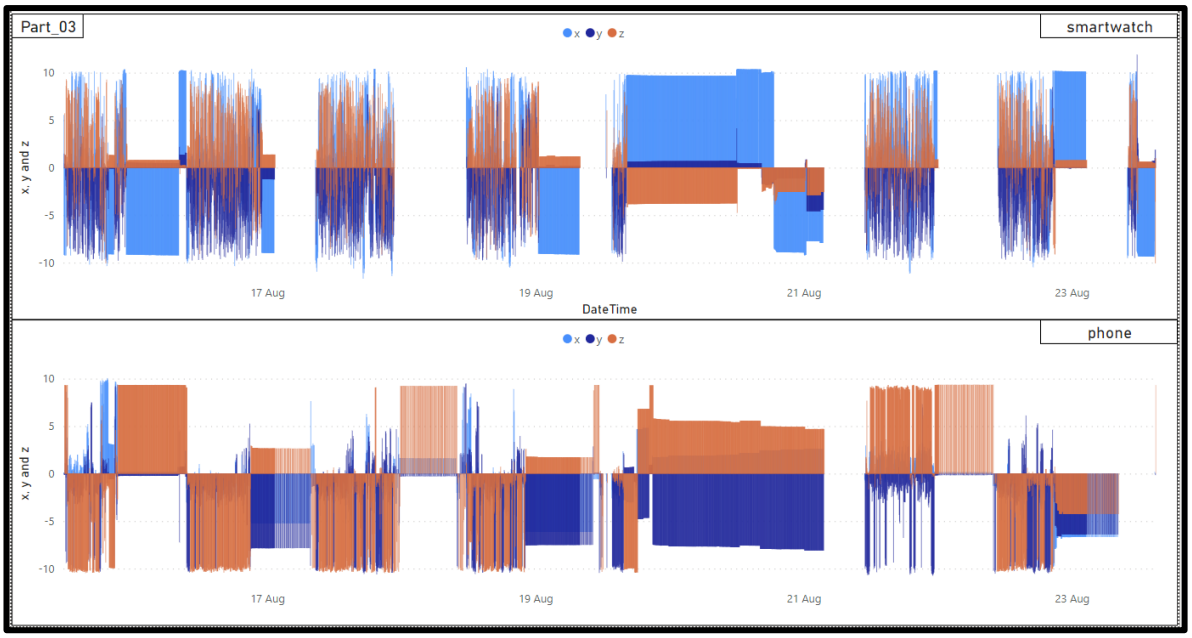
The figures below show the accelerometer's variations of the smartphone and smartwatches worn over the recruitment period on all axis (x, y and z) for each participant.



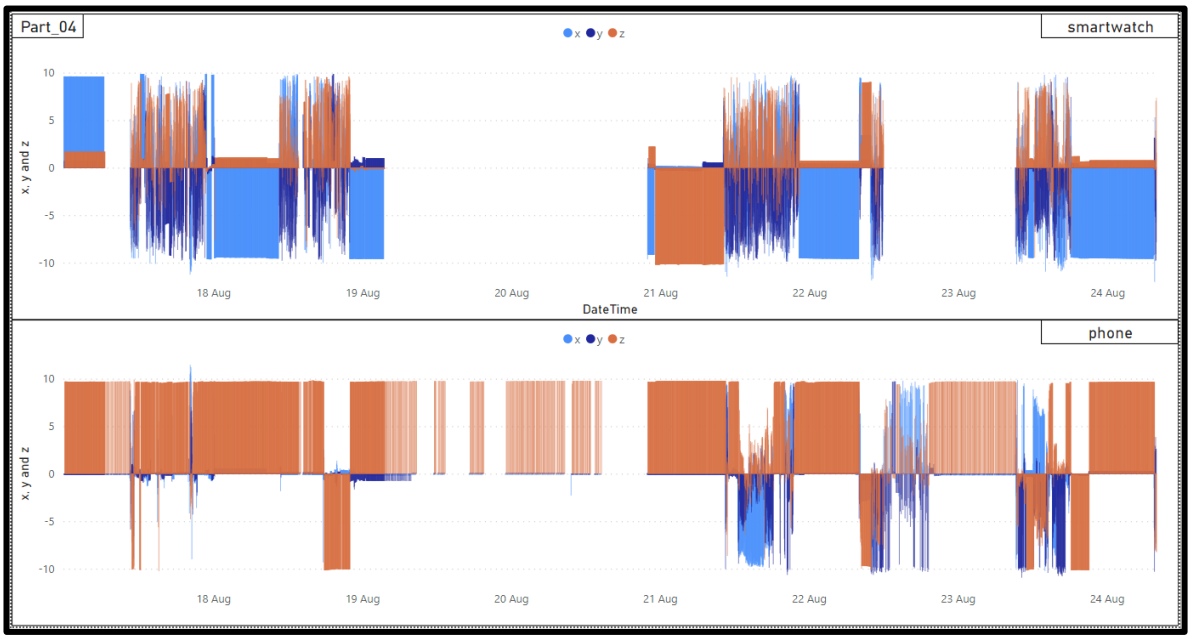
SP & SW raw acceleromter (x, y and z) over time (Participant 1)



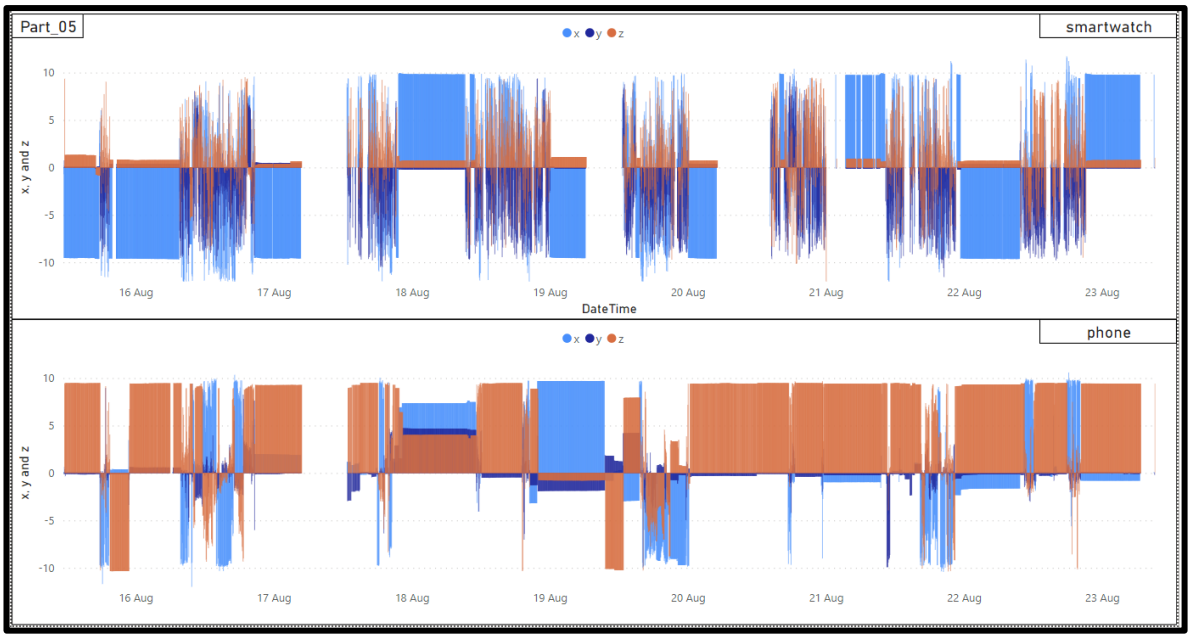
SP & SW raw acceleromter (x, y and z) over time (Participant 2)



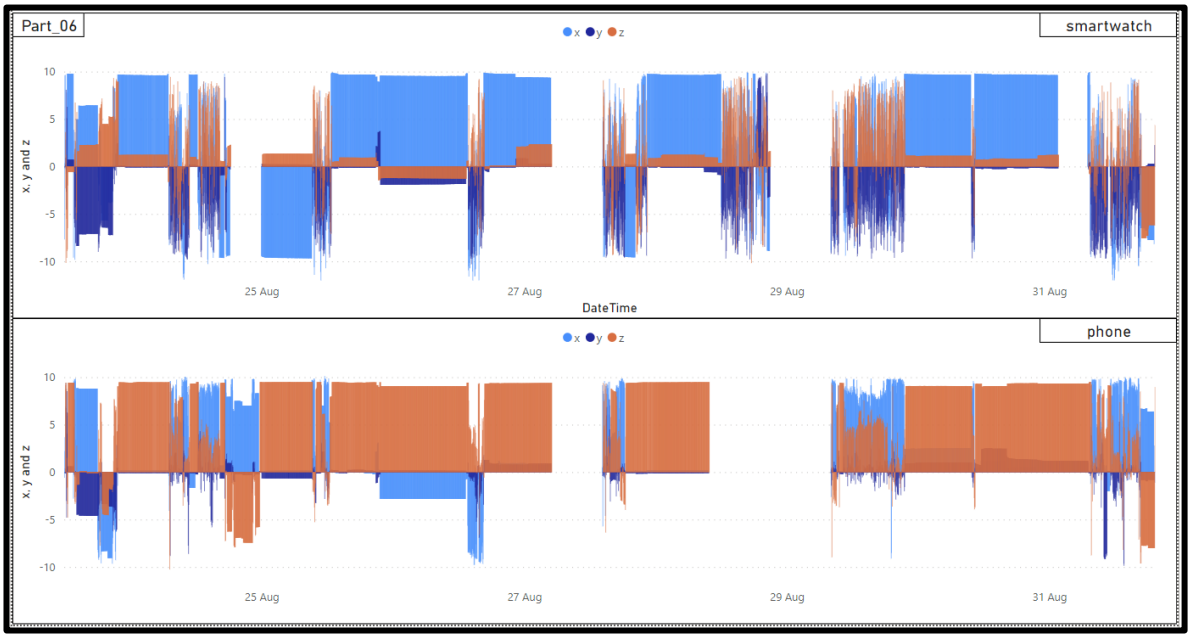
SP & SW raw acceleromter (x, y and z) over time (Participant 3)



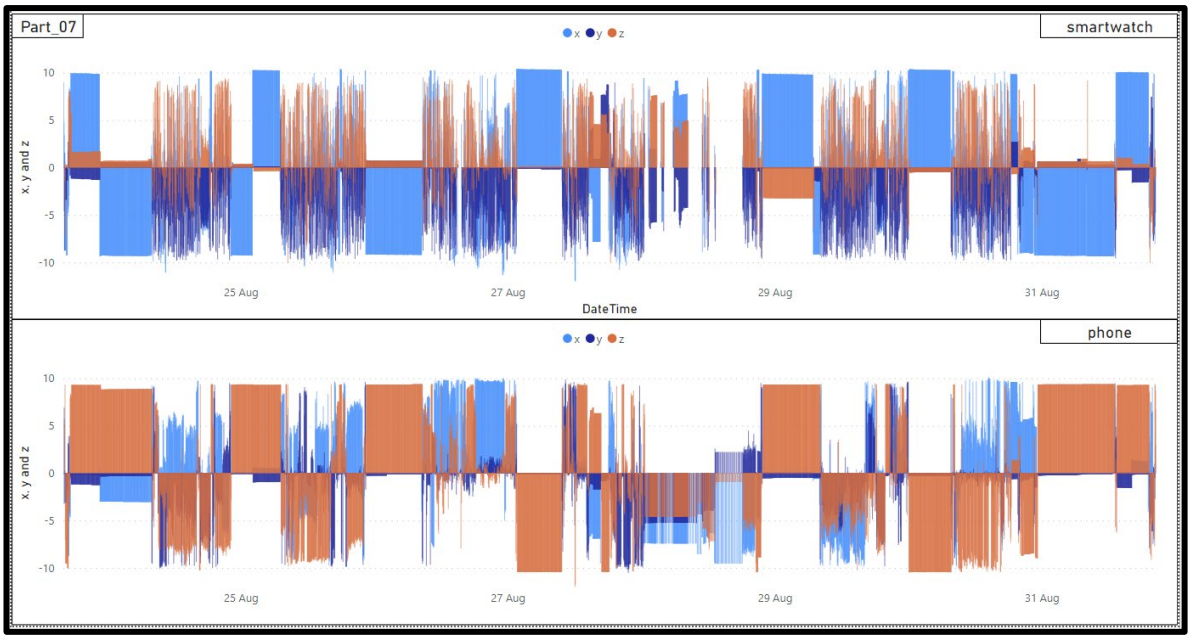
SP & SW raw acceleromter (x, y and z) over time (Participant 4)



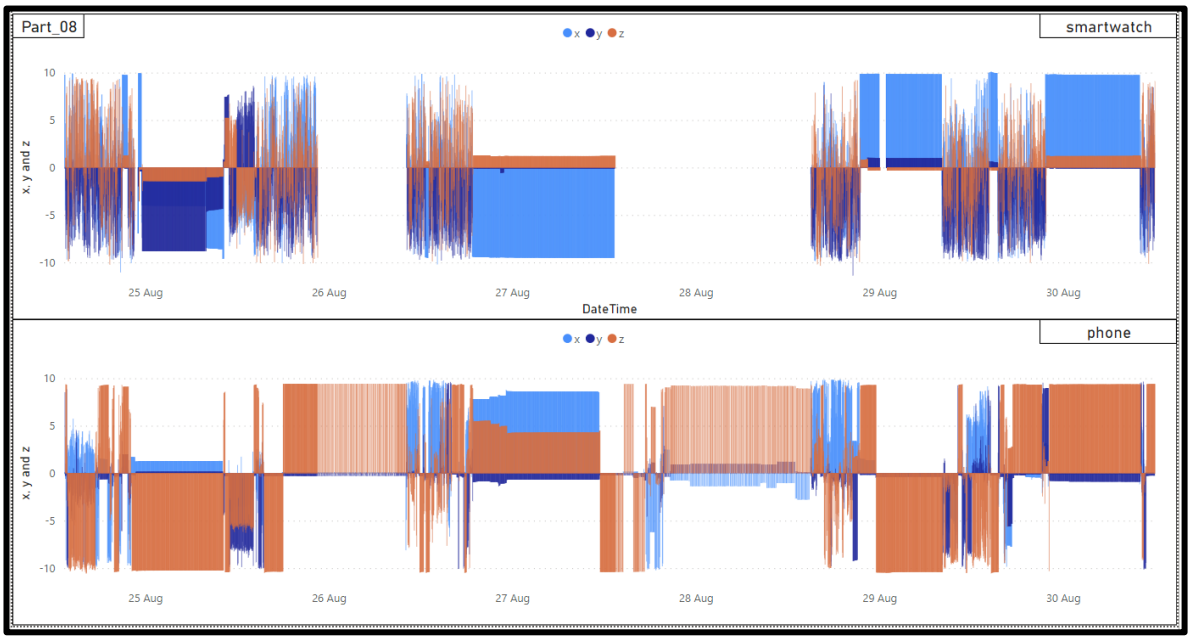
SP & SW raw acceleromter (x, y and z) over time (Participant 5)



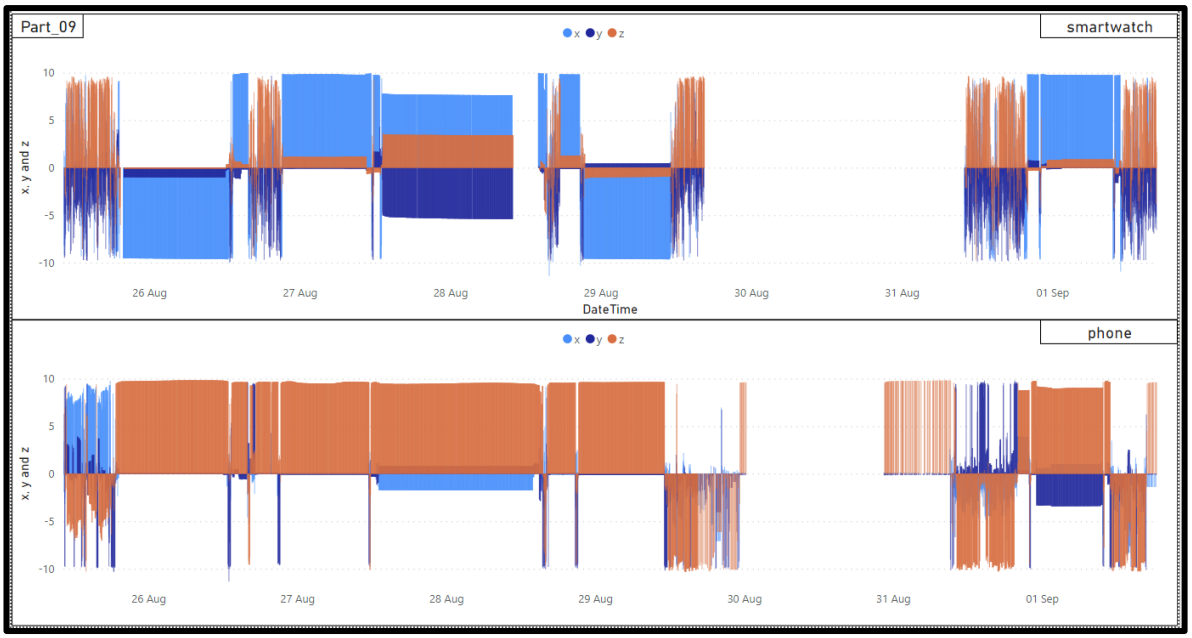
SP & SW raw acceleromter (x, y and z) over time (Participant 6)



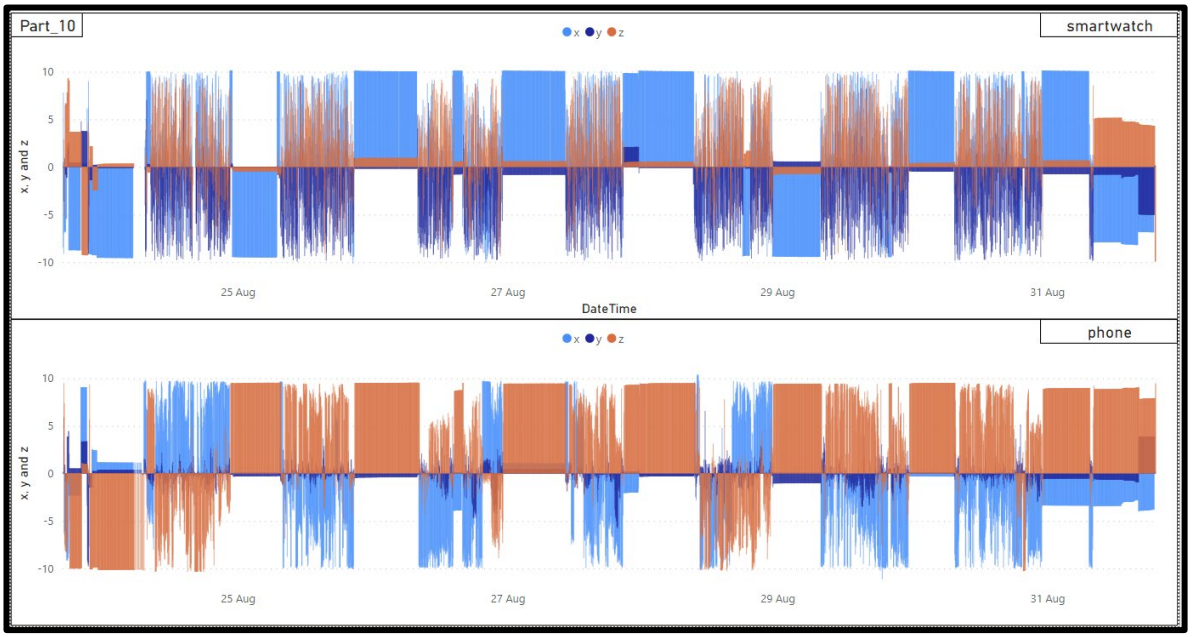
SP & SW raw acceleromter (x, y and z) over time (Participant 7)



SP & SW raw acceleromter (x, y and z) over time (Participant 8)



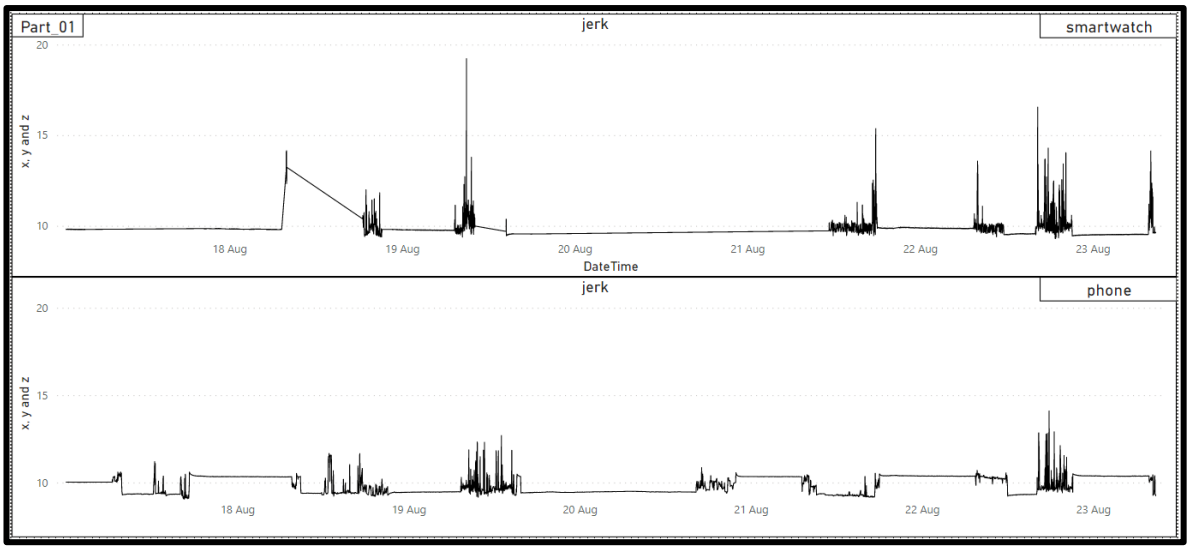
SP & SW raw acceleromter (x, y and z) over time (Participant 9)



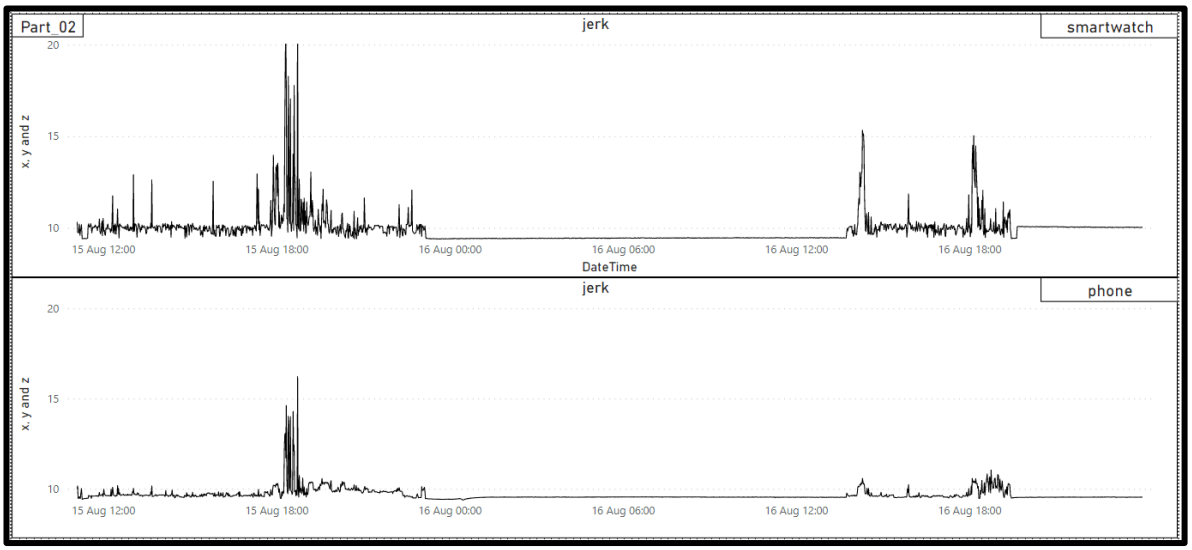
SP & SW raw acceleromter (x, y and z) over time (Participant 10)

Appendix F Study 2: SP & SW jerk (mean)

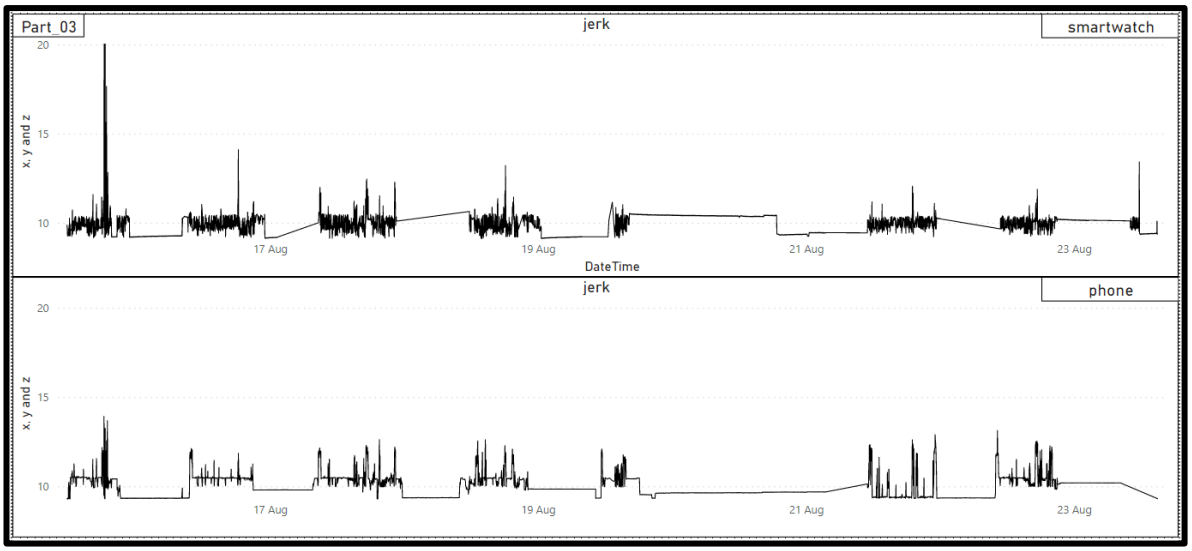
The plots below show the mean of jerks calculated from each axis over time for both SW and SP.



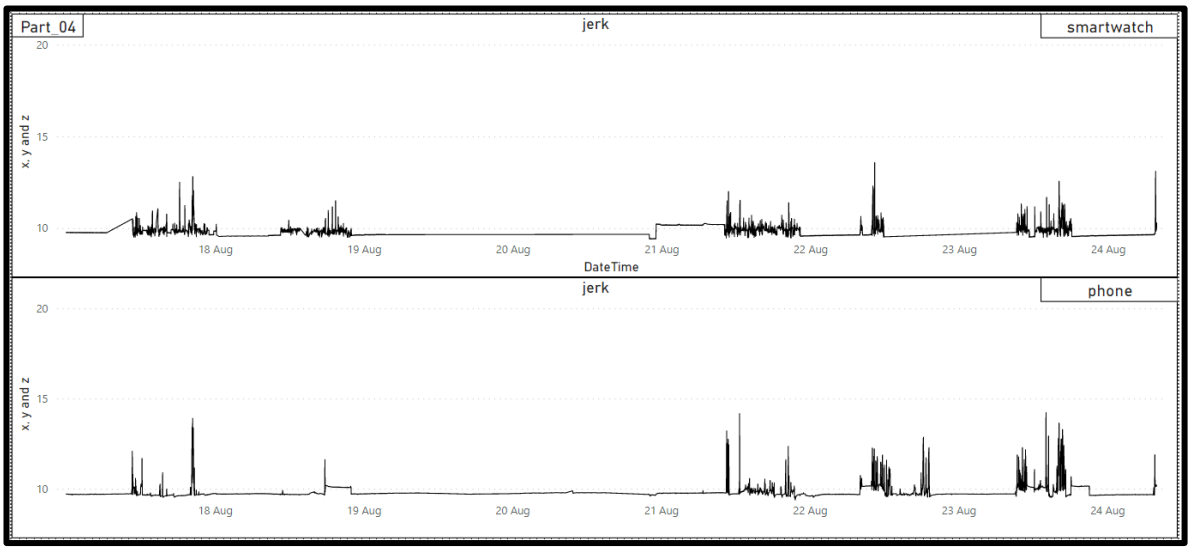
SP & SW jerk (mean) over time (Participant 1)



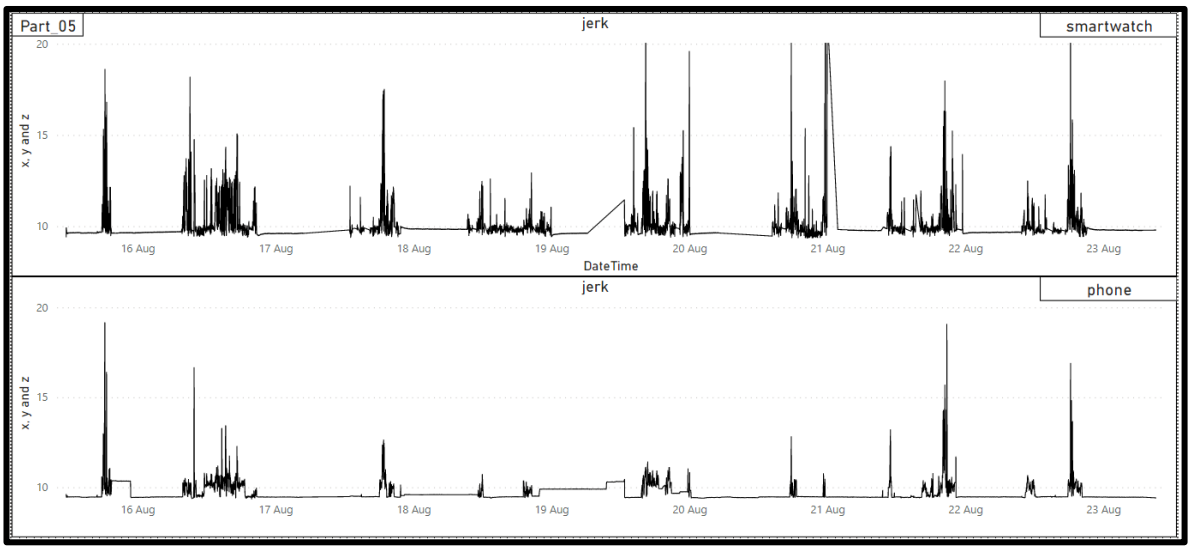
SP & SW jerk (mean) over time (Participant 2)



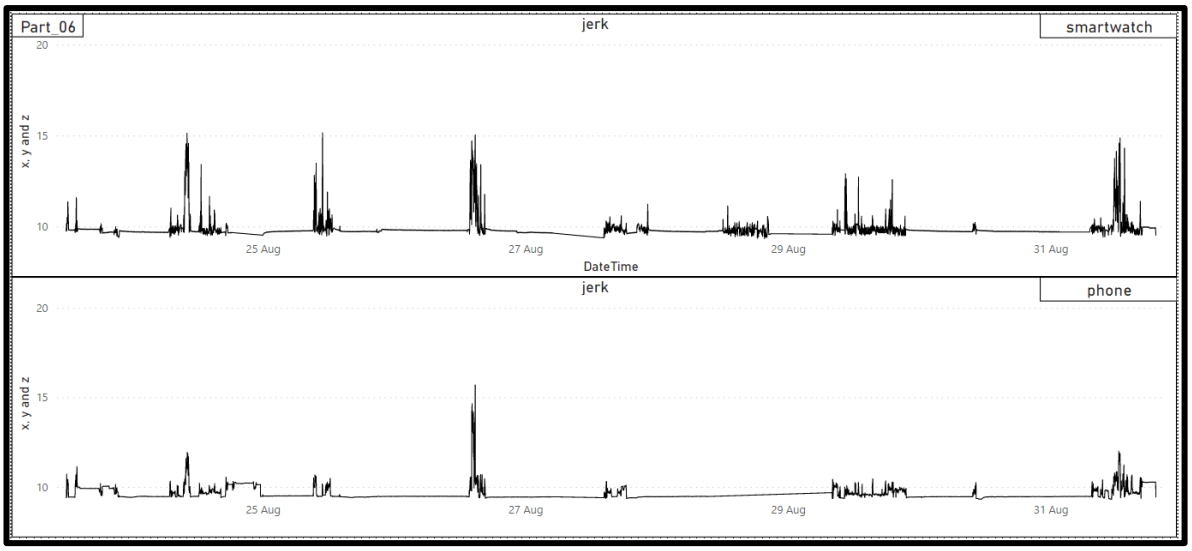
SP & SW jerk (mean) over time (Participant 3)



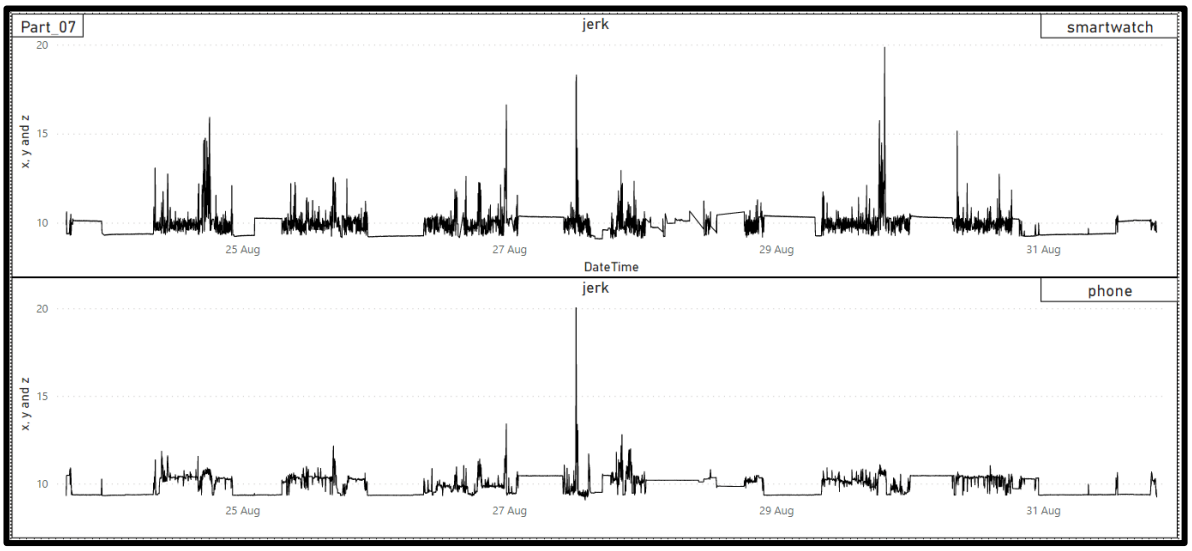
SP & SW jerk (mean) over time (Participant 4)



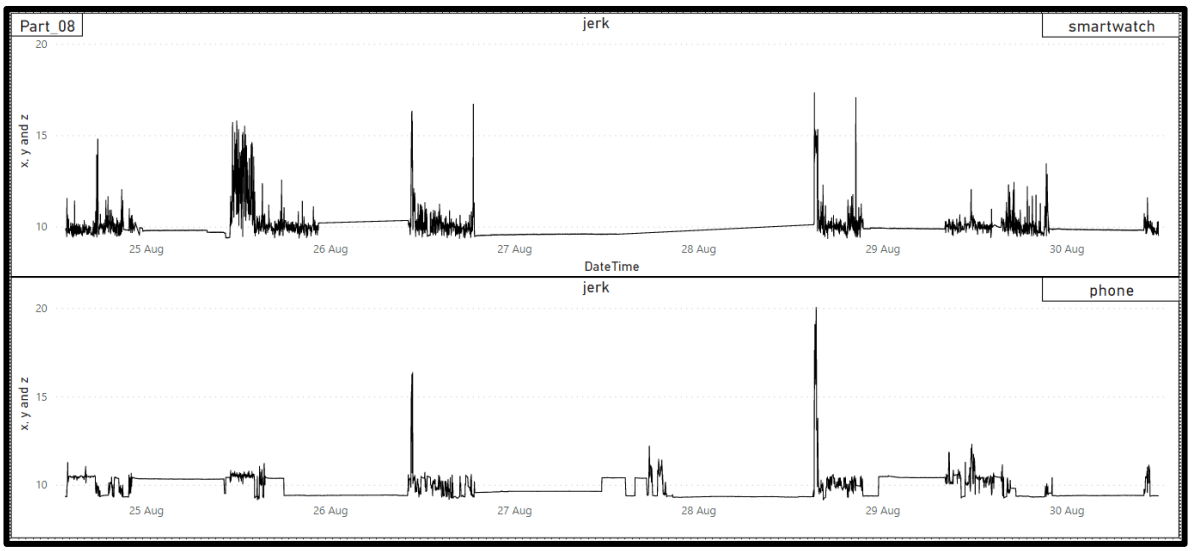
SP & SW jerk (mean) over time (Participant 5)



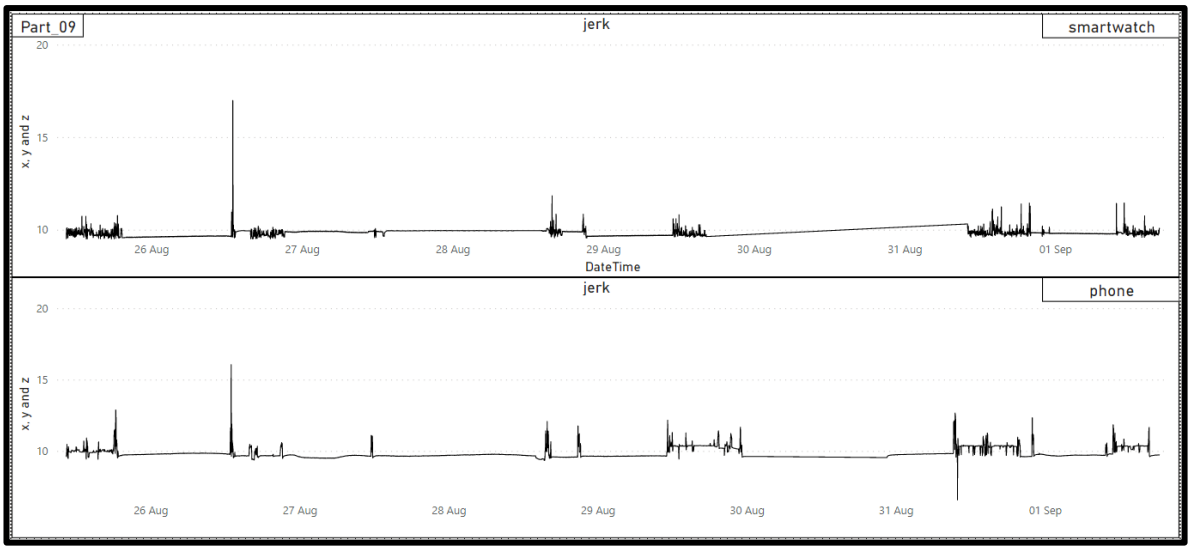
SP & SW jerk (mean) over time (Participant 6)



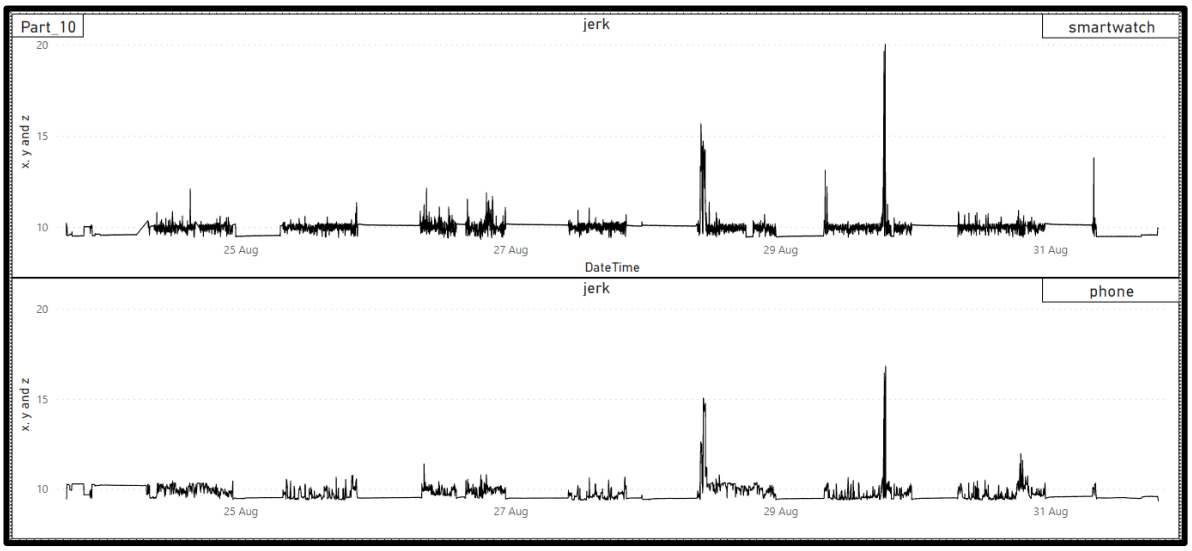
SP & SW jerk (mean) over time (Participant 7)



SP & SW jerk (mean) over time (Participant 8)



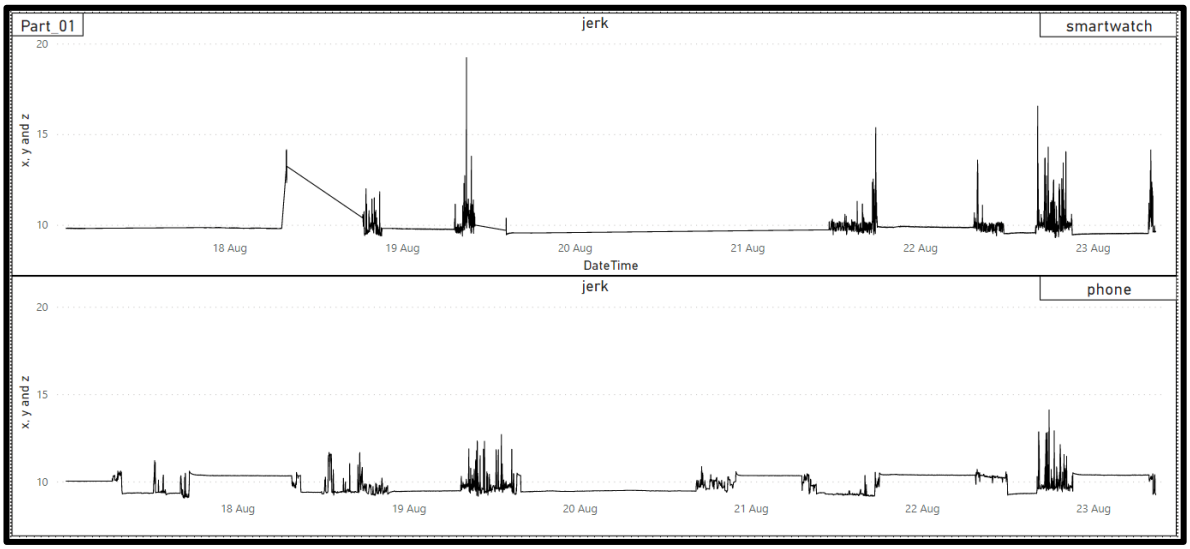
SP & SW jerk (mean) over time (Participant 9)



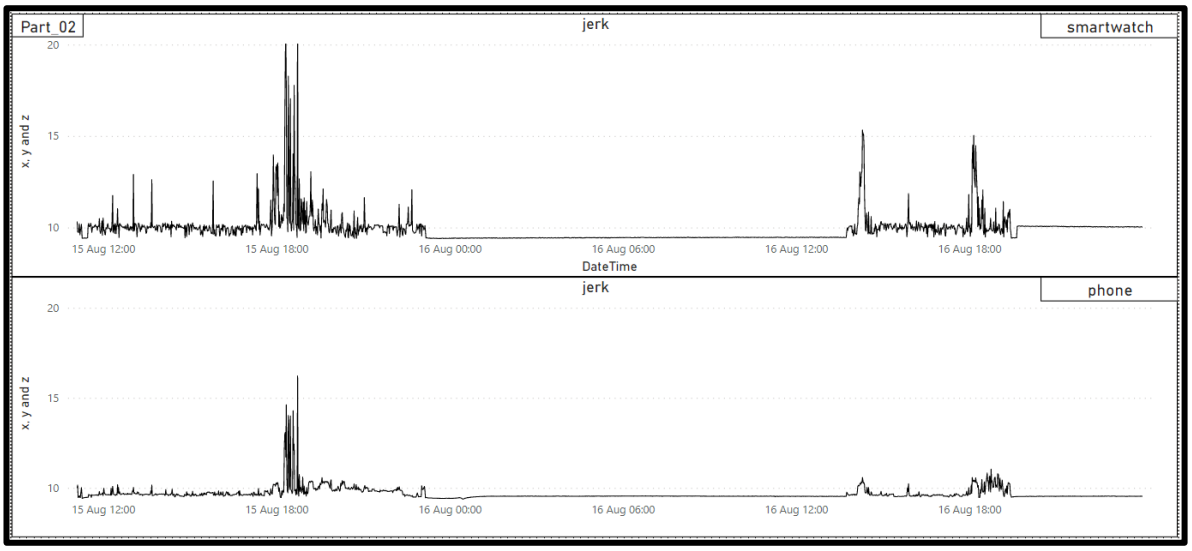
SP & SW jerk (mean) over time (Participant 10)

Appendix G Study 2: SP & SW Phenotypes

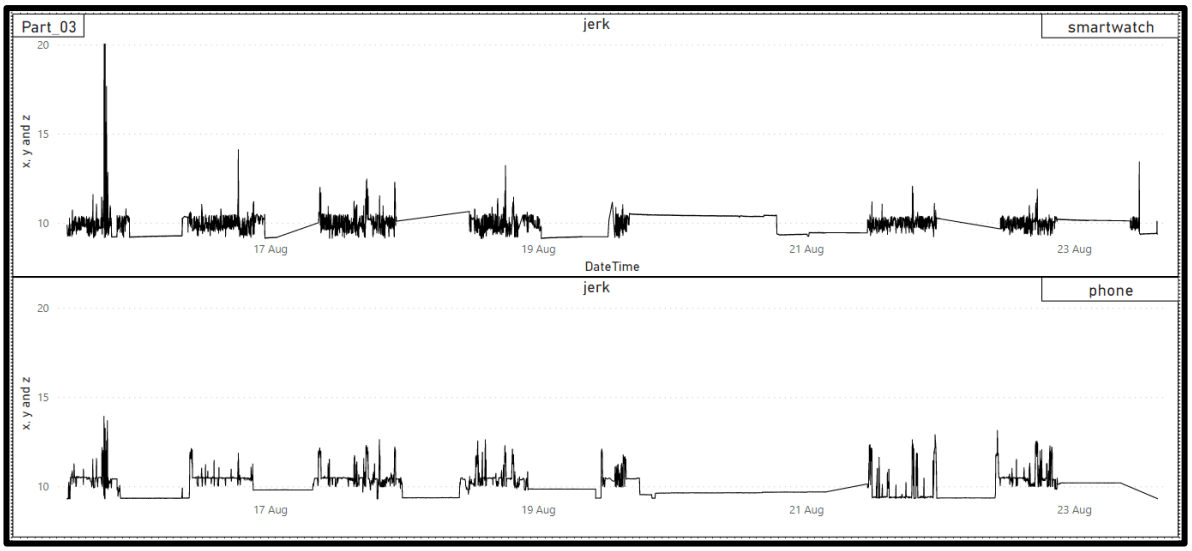
The figures below show the phenotypes and characteristics generated for all participants on both SP & SW.



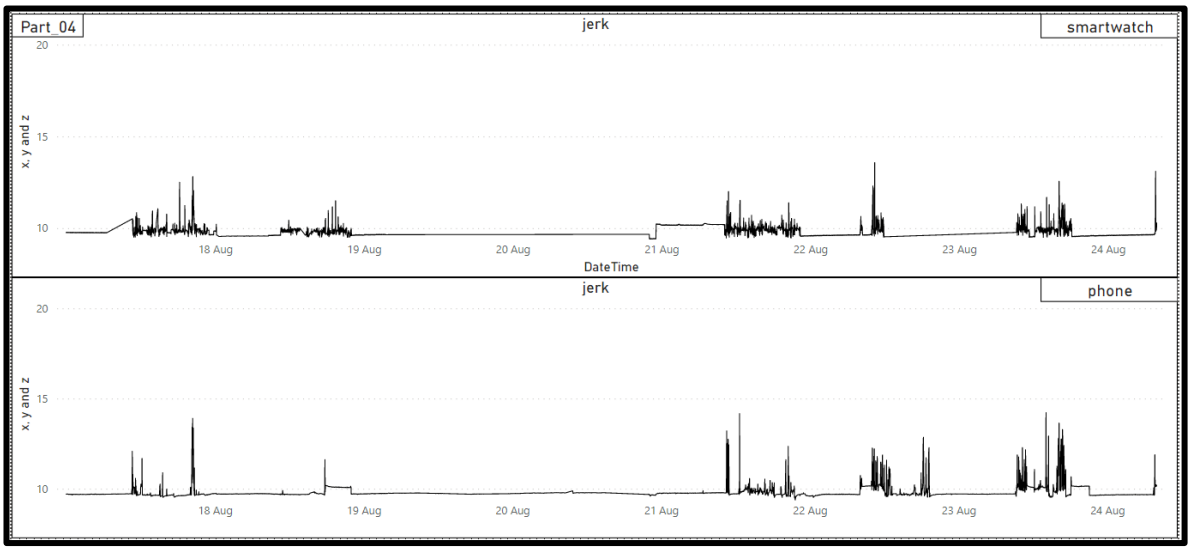
SP & SW jerk (mean) over time (Participant 1)



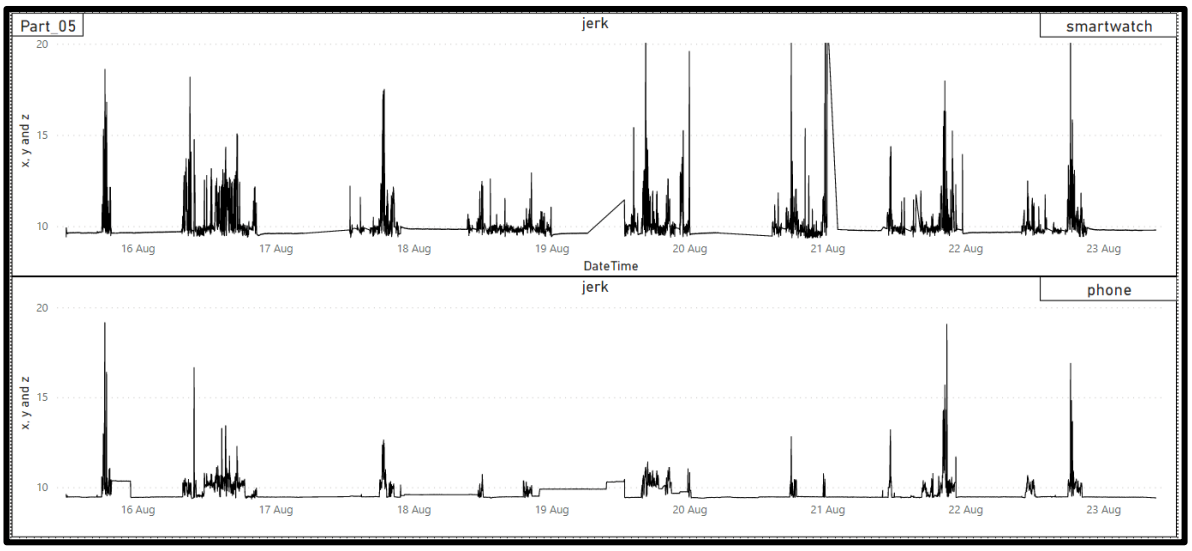
SP & SW jerk (mean) over time (Participant 2)



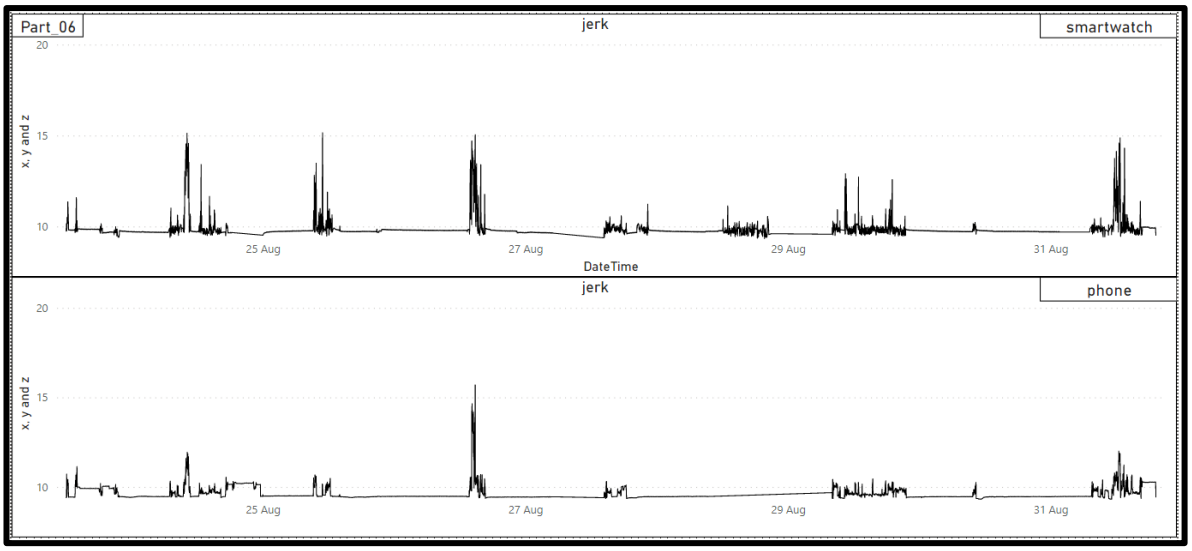
SP & SW jerk (mean) over time (Participant 3)



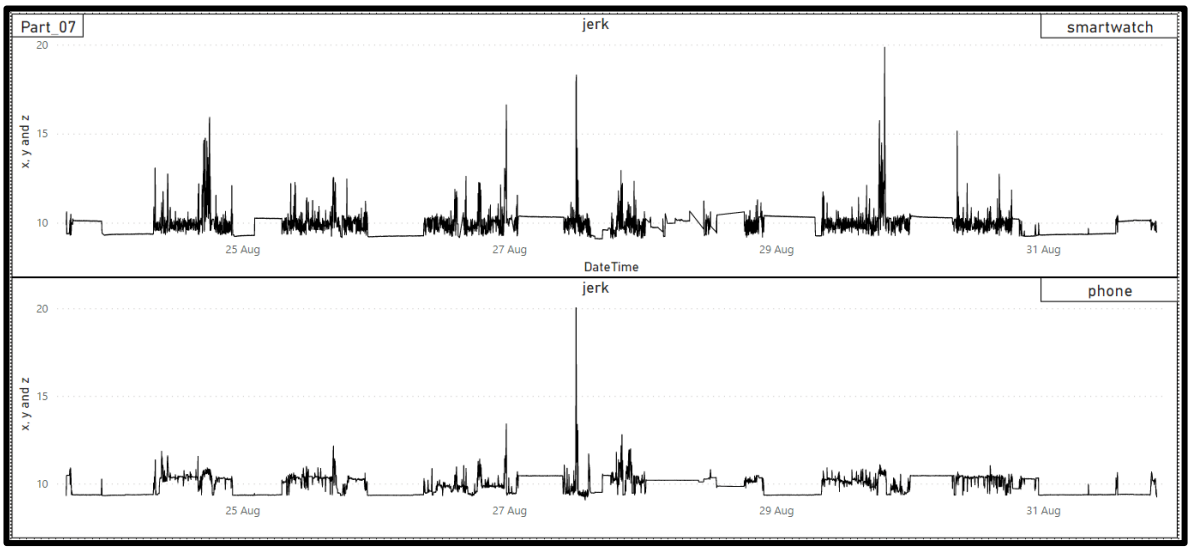
SP & SW jerk (mean) over time (Participant 4)



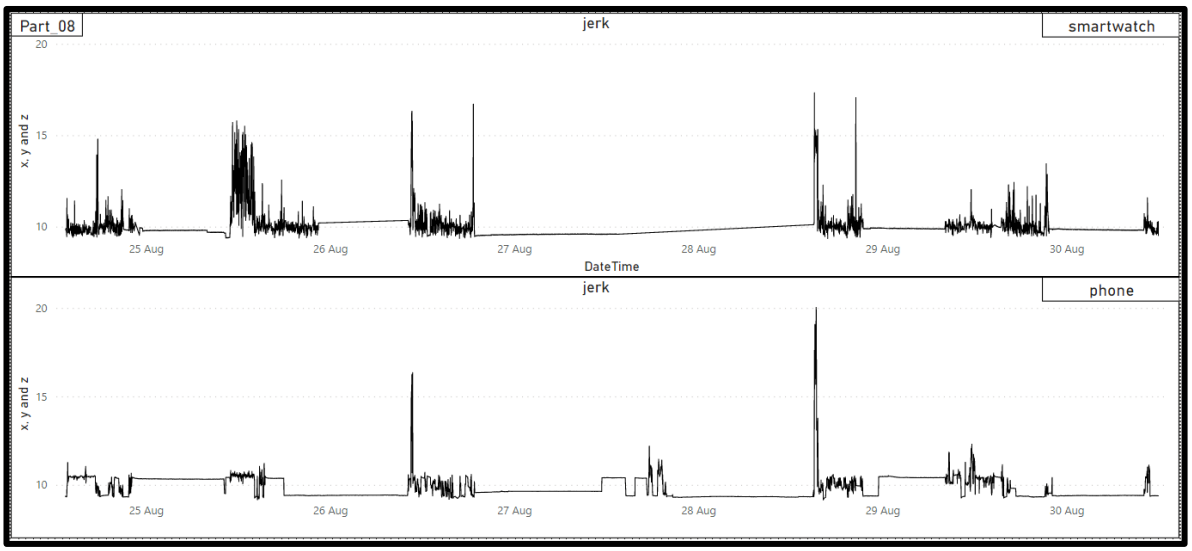
SP & SW jerk (mean) over time (Participant 5)



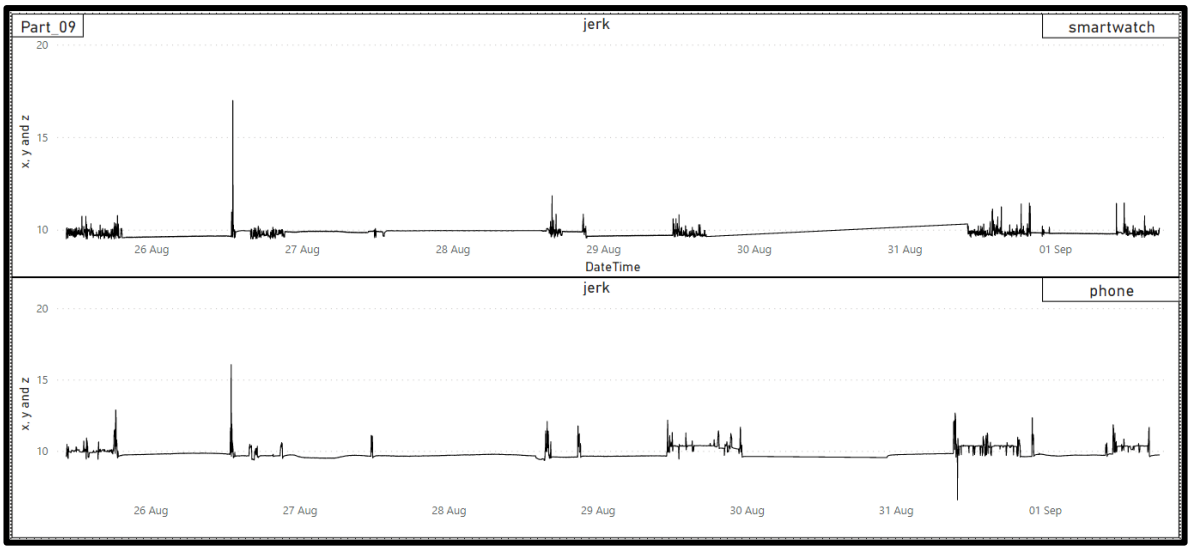
SP & SW jerk (mean) over time (Participant 6)



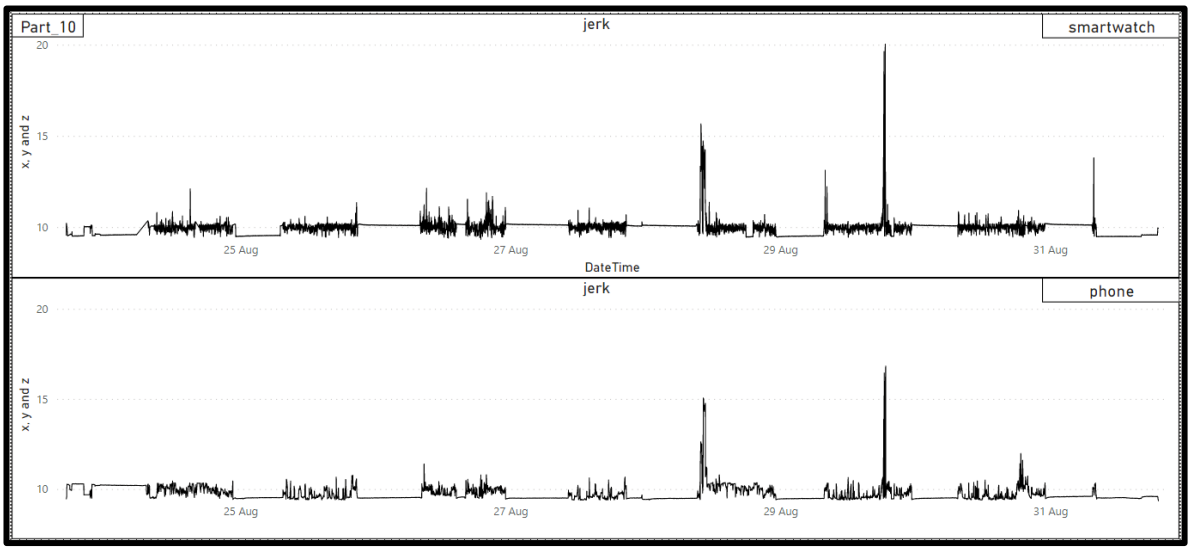
SP & SW jerk (mean) over time (Participant 7)



SP & SW jerk (mean) over time (Participant 8)



SP & SW jerk (mean) over time (Participant 9)



SP & SW jerk (mean) over time (Participant 10)

Appendix H Study 3: Patient list

	Baseline			2 Months		
Patient ID	Phone ID	Start	End	Phone ID	Start	End
1	Samsung 1	04/06/2018				
2	Samsung 2	26/06/2018	03/07/2018	Samsung 11	05/10/2018	11/10/2018
3	Samsung 3	19/06/2018	26/06/2018	Samsung 9	13/10/2018	19/10/2018
4	Samsung 4	03/07/2018	09/07/2018	Samsung 4	25/09/2018	01/10/2018
5	Samsung 5	26/06/2018	02/07/2018	Samsung 2	26/09/2018	02/10/2018
6	Samsung 6	02/07/2018	08/07/2018	Samsung 3	04/10/2018	11/10/2018
7	Samsung 7	03/07/2018	10/07/2018	Samsung 4	09/10/2018	15/10/2018
8	Samsung 8	09/07/2018	15/07/2018			
9	Samsung 9	09/07/2018	15/07/2018	Samsung 11	11/10/2018	17/10/2018
10	Samsung 10	10/07/2018	16/07/2018	Samsung 5	27/09/2018	03/10/2018
11	Samsung 2	10/07/2018	16/07/2018	Samsung 2	04/10/2018	11/10/2018
12	Samsung 3	17/07/2018	23/07/2018			
13	Samsung 4	16/07/2018	22/07/2018	Samsung 7	11/10/2018	17/10/2018
14	Samsung 5	16/07/2018	23/07/2018	Samsung 2	17/10/2018	23/10/2018
15	Samsung 6	16/07/2018	23/07/2018	Samsung 8	11/10/2018	17/10/2018
16	Samsung 7	23/07/2018	29/07/2018	Samsung 10	30/10/2018	05/11/2018
17	Samsung 8	23/07/2018	30/07/2018	Samsung 4	29/10/2018	04/11/2018
18	Samsung 9	30/07/2018	05/08/2018	Samsung 6	28/10/2018	03/11/0218
19	Samsung 10	30/07/2018	05/08/2018	Samsung 8	28/10/2018	04/11/2018
20	Samsung 2	31/07/2018	06/08/2018			
21	Samsung 3	31/07/2018	06/08/2018	Samsung 3	19/10/2018	25/10/2018
22	Samsung 4	06/08/2018	12/08/2018	Samsung 4	08/11/2018	14/11/2018
23	Samsung 5	06/08/2018	12/08/2018	Samsung 11	10/11/2018	16/11/2018
24	Samsung 6	07/08/2018	14/08/2018	Samsung 6	09/11/2018	15/11/2018
25	Samsung 7	08/08/2018	14/08/2018	Samsung 2	06/12/2018	12/12/2018
26	Samsung 9	13/08/2018	19/08/2018	Samsung 10	09/11/2018	15/11/2018
27	Samsung 10	13/08/2018	19/08/2018		16/11/2018	22/11/2018
28	Samsung 2	14/08/2018	20/08/2018	Samsung 5	08/11/2018	14/11/2018
29	Samsung 4	20/08/2018	26/08/2018	Samsung 10	30/11/2018	06/12/2018
30	Samsung 5	21/08/2018	27/08/2018	Samsung 4	29/11/2018	06/12/2018
31	Samsung 6	21/08/2018	27/08/2018			
32	Samsung 7	21/08/2018	27/08/2018	Samsung 6	27/11/2018	03/12/2018
33	Samsung 8	28/08/2018	03/09/2018	Samsung 13	31/12/2018	06/01/2019

34	Samsung 10	28/08/2018	03/09/2018	Samsung 8	05/12/2018	11/12/2018
35	Samsung 2	03/09/2018	09/09/2018	Samsung 3	13/12/2018	19/12/2018
36	Samsung 3	04/09/2018	10/09/2018	Samsung 11	04/12/2018	10/12/2018
37	Samsung 5	04/09/2018	10/09/2019	Samsung 12	01/12/2018	07/12/2018
38	Samsung 6	10/09/2018	16/09/2018	Samsung 8	21/12/2018	27/12/2018
39	Samsung 4	11/09/2018	17/09/2018	Samsung 9	03/12/2018	10/12/2018
40	Samsung 9	01/10/2018	07/10/2018	Samsung 4	15/12/2018	21/12/2018
41	Samsung 7	17/09/2018	23/09/2018	Samsung 6	02/01/2019	08/01/2019
42	Samsung 8	18/09/2018	24/09/2018	Samsung 2	21/12/2018	27/12/2018
43	Samsung 10	25/09/2018	01/10/2018	Samsung 11	14/12/2018	20/12/2018
44	Samsung 6	01/10/2018	07/10/2018	Samsung 10	14/12/2018	20/12/2018
45	Samsung 7	01/10/2018	07/10/2018	Samsung 5	14/12/2018	20/12/2018
46	Samsung 8	02/10/2018	08/10/2018	Samsung 7	14/12/2018	20/12/2018
47	Samsung 5	08/10/2018	14/10/2018	Samsung 12	26/12/2018	01/01/2019
48	Samsung 10	08/10/2018	14/10/2018	Samsung 10	27/12/2018	02/01/2019
49	Samsung 6	09/10/2018	15/10/2018	Samsung 10	22/01/2019	28/01/2019
50	Samsung 3	05/11/2018	11/11/2018	Samsung 2	02/02/2019	08/02/2019
51	Samsung 2	05/11/2018	11/11/2018	Samsung 11	21/02/2019	27/02/2019
52	Samsung 7	05/11/2018	11/11/2018	Samsung 11	31/01/2019	06/02/2019
53	Samsung 9	06/11/2018	12/11/2018			
54	Samsung 12	12/11/2018	19/11/2018	Samsung 8	04/02/2019	10/02/2019
55	Samsung 8	13/11/2018	19/11/2018	Samsung 13	31/01/2019	06/02/2019
56	Samsung 3	19/11/2018	25/11/2018	Samsung 4	19/02/2019	25/02/2019
57	Samsung 5	26/11/2018	02/12/2018	Samsung 12	22/02/2019	28/02/2019
58	Samsung 2	08/01/2019	14/12/2018			
59	Samsung 3	28/01/2019	03/02/2019	Samsung 8	10/05/2019	16/05/2019
60	Samsung 9	29/01/2019	04/02/2019	Samsung 7	24/04/2019	30/04/2019
61	Samsung 12	29/01/2019	04/02/2019	Samsung 3	24/04/2019	30/04/2019
62	Samsung 10	04/02/2019	10/02/2019	Samsung 12	29/05/2019	04/06/2019
63	Samsung 6	11/02/2019	17/02/2019	Samsung 6	11/06/2019	17/06/2019
64	Samsung 7	18/02/2019	24/02/2019	Samsung 10	29/05/2019	04/06/2019
65	Samsung 13	26/02/2019	04/03/2019	Samsung 5	28/05/2019	03/06/2019
66	Samsung 2	05/03/2019	11/03/2019	Samsung 3	06/06/2019	13/06/2019

Appendix I Study 3: Phenotype characteristics

The table below describes the phenotype characteristics for all patients.

		Raw			Interpolated		
ID	Session	Peak	Slope	CMS	Peak	Slope	CMS
2	1	1.2	-1.5	0.45	0.2	-1.5	0.98
2	2	1	-1.5	0.84	0.2	-1.5	0.99
3	1	1	-1.7	0.92	0.2	-1.5	0.98
4	1	1	-1.6	0.66	0.2	-1.5	0.98
5	1	1.1	-1.4	0.72	0.2	-1.4	0.98
5	2	1	-1.3	0.71	0.2	-1.4	0.99
7	2	1	-1.3	0.76	0.2	-1.4	0.98
8	2	1.3	-1.1	0.37	0.1	-1.1	0.98
9	2	1	-1.6	0.63	0.1	-1.4	1
11	2	1	-1.2	0.56	0.2	-1.2	0.98
12	2	1.3	-1.1	0.35	0.1	-1.1	0.98
14	1	1	-1.5	0.83	0.1	-1.6	1
15	1	1	-1.6	0.78	0.2	-1.6	0.98
16	1	0.9	-1.3	0.69	0.2	-1.5	0.98
16	2	1.1	-1.2	0.89	0.1	-1.4	1
18	1	0.9	-1.3	0.58	0.2	-1.2	0.94
18	2	0.9	-1.2	0.54	0.1	-1.1	0.99
19	1	0.9	-1.1	0.69	0.1	-1.1	1
20	2	1.3	-1.1	0.36	0.1	-1.1	0.98
21	2	1	-1.1	0.66	0.2	-1.3	1
22	1	1	-1.7	0.78	0.2	-1.4	0.97
23	2	0.9	-1.2	0.48	0.2	-1.1	0.95
25	2	1	-1.3	0.59	0.2	-1.3	1
27	1	1	-1.6	0.56	0.1	-1.5	0.98
27	2	0.9	-1.6	0.65	0.1	-1.5	0.98
28	1	1	-1.8	0.82	0.2	-1.6	0.98
28	2	1.1	-1.2	0.72	0.2	-1.3	0.99
29	1	1	-1.1	0.84	0.1	-1.6	0.99
30	2	1.1	-0.5	0.58	0.1	-1.9	1
31	2	1.3	-1.1	0.36	0.1	-1.1	0.98
32	2	1	-1.3	0.5	0.1	-1.3	1
33	1	0.8	-1.4	0.89	0.2	-1.6	0.96
33	2	1.1	-1.5	0.91	0.2	-1.5	0.98
34	1	1.1	-1.3	0.58	0.1	-1.4	1
34	2	1.2	-1.7	0.78	0.1	-1.5	1
35	2	1.1	-0.5	0.71	0.1	-1.3	1
36	1	0.6	-1.5	0.66	0.2	-1.4	0.98
37	1	1	-1.6	0.77	0.2	-1.7	0.98
39	1	1	-1.3	0.7	0.2	-1.4	0.99
40	1	0.8	-1.3	0.61	0.1	-1.3	0.99
43	2	1	-1.6	0.81	0.2	-1.6	0.98
44	1	0.7	-1.2	0.81	0.2	-1.2	0.99
44	2	1	-1.2	0.58	0.2	-1.2	0.99
47	1	1	-1.1	0.6	0.1	-0.9	0.99
47	2	0.7	-1.1	0.8	0.2	-1.1	0.99
49	1	1	-1.1	0.8	0.2	-1.5	1
50	1	0.7	-1.2	0.86	0.1	-1.2	0.99
51	1	0.8	-1.3	0.78	0.1	-1.4	1
51	2	0.8	-1.6	0.53	0.1	-1.5	0.97
52	2	1	-1.4	0.8	0.2	-1.4	0.98
53	1	1	-1.3	0.54	0.2	-1.4	0.98
53	2	1.3	-1.1	0.37	0.1	-1.1	0.98
54	2	1.1	-1.3	0.87	0.2	-1.4	0.99
55	2	1.1	-2	0.77	0.1	-2.4	0.99
58	2	1.3	-1.1	0.36	0.1	-1.1	0.98
60	1	1	-1.4	0.43	0.1	-1.5	1
61	1	1	-1.2	0.9	0.1	-1.5	1
61	2	1.1	-1.2	0.77	0.1	-1.4	1
63	2	0.9	-1.4	0.75	0.1	-1.4	0.97
64	1	1	-1.5	0.58	0.2	-1.4	0.98
65	1	1	-1.4	0.61	0.2	-1.4	0.98
65	2	1	-1.4	0.57	0.2	-1.3	0.97
66	1	1.1	-1.4	0.75	0.1	-1.5	1

Appendix J NHRA Docs

NHS
National Institute for Health Research

CERTIFICATE of ACHIEVEMENT

This is to certify that

Jimmy Caroupapoulle

has completed the course

Good Clinical Practice (GCP) Refresher: eLearning

14 March 2019

A practical guide to ethical and scientific quality standards in clinical research

Including EU Directives, Medicines for Human Use (Clinical Trials) Regulations & the Department of Health UK Policy Framework for Health & Social Care Research, as applied to the conduct of Clinical Trials & other studies conducted in the NHS

Modules completed:

Core
Team Roles
Eligibility
Safety Reporting
Electronic Studies and Source Text
Summary

This course is worth 3 CPD credits

Delivering research to make patients, and the NHS, better

Appendix K BSR (April 2017) Poster & Abstract

Rheumatoid Arthritis, a study of continuous monitoring of physical activity using smartphones (RApp)

University of Southampton

Jimmy Caroupapoullé[1], Dr Alex Forrester[2], Prof. Chris Edwards[3], Prof. Cyrus Cooper[4]

Background

- Have you ever wondered how your smartphone knows which way up you are holding your phone, and twists the screen around to fit? Most now have built-in devices that can sense movement. It's these sensors that allow RApp to track physical activity.
- Measuring disease activity to assess the effects of medical treatment for patients with rheumatoid arthritis (RA) is well reported but far less information on quantity and quality of day-to-day physical activity is available. RA causes pain and stiffness that discourages physical movement and reduces mobility while exercising helps to relieve symptoms.
- The aim of our study is to demonstrate that continuous measurement of RA patient activity, recorded using a non-intrusive smartphones-based approach can be used as an objective and reliable input to current disease activity tools and help promote physical activity.

Methodology

Designed with patients

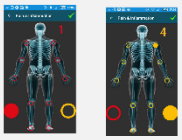
Analysis
- Commercial activity monitors focus on step counts and calories burned.
- Our approach is to analyse rates of acceleration (jerk) and calculate metrics related to impact and joint loading (load rate).

Design
- Exploratory study with a group of men and women aged 36-65 diagnosed with RA for 4-23 years.
- PPI focus group and questionnaire to assess the design of RApp and scope the requirements.

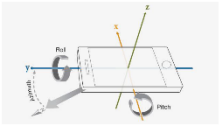
Evaluation
- **Study 1**: Assess the stability, usability and overall data gathering approach.
- **Study 2&3**: Expand to more RA patients and release RApp on Google Play.

Patient self assessment

Using RApp, patients can record:
- Self-assessed pain, inflammation and perceived physical activity via daily questionnaire (MDHAQ/RAPID3).
- Self-reported DAS28 count, with an intensity factor introduced to highlight how painful and/or swollen a joint is (0-4).



Statistical sampling



- On-board accelerometer sampling to evaluate the activity of a smartphone user.
- On-board computing power to perform data analysis which solves 'big data' of 24/7 monitoring.
- Sampling and analysis run in the background, consuming negligible power and without affecting other processes on the phone.

- We developed an intermittent random 'Monte-Carlo' strategy that breaks the sampling into small chunks performed at random times.
- From this we can calculate the jerk *(j)* which shows the change of acceleration.
- And the load rate *(L)* which indicates the force applied to the joints.

$$\dot{f}_L = \frac{df_L}{dt} = m\,\frac{da}{dt} = m\,j$$

Results

- *Probability density function*: likelihood of the mean for the specified day
- *Mean jerk*: value that the sensors (ie joints) withstand in N/kg/s

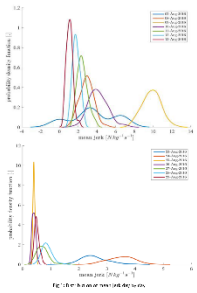
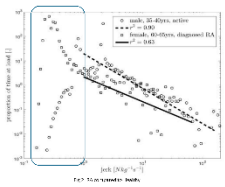


Figure 1 shows the activity as a daily load distribution:
- The activity is much higher on the 09th of August and consistently lower for the rest of the week. Similar trend can be seen on the week of the 24th of August.
- Understanding the data in this way allows us to see the consistency in physical activity.
- Here, the patient has consistently a peak of physical activity on the Tuesdays compared to the rest of the week.

Figure 2 shows the trend of an RA patient compared to a healthy volunteer:
- The most frequent activity level for RA patient is ~0.25 N/kg/s
- 10x higher for healthy participant.
- Beyond the most frequent activity, both distributions become linear but with noticeably different trends



Online viewing/Portal

- Activity and self-assessment are regularly uploaded to a cloud based Azure server
- Upload to Azure mitigates the smartphone's storage limitation
- Online access allows analysis and visualization of results through web-based interface.
- This could enable RA patients to take a role in monitoring their own health.

Rapp admin	GP/Clinician	RA patient
Admin overview	Portal to input medical data: Blood samples, ESR ...	Portal with aggregated data: Medical data, physical activity, self assessment...

Conclusion

- It is possible to design an app looking at the quality of quantitative movement which RA patients find both usable and acceptable.
- Preliminary data suggests the promise of using activity probability distributions to assess a patient's physical activity.
- Online access provides scope for potential design of an integrated self management platform.

1 PhD Research Student, Computational Engineering and Design Research Group, Faculty of Engineering and the Environment, University of Southampton
2 Associate Professor in Computational Engineering and Design Research Group, Faculty of Engineering and the Environment, University of Southampton
3 Consultant Rheumatologist, University Hospital Southampton NHS Foundation Trust. Honorary Chair of Clinical Rheumatology, Professor in Rheumatology, Faculty of Medicine, University of Southampton, Associate Director, Southampton NIHR Welcome Trust Clinical Research Facility
4 Director & Professor of Rheumatology, MRC Lifecourse Epidemiology Unit; Vice-Dean, Faculty of Medicine, University of Southampton; and Professor of Epidemiology, University of Oxford.

Title:

Rheumatoid Arthritis, a study of continuous monitoring of physical activity using smartphones

Background:

Measuring disease activity to assess the effects of medical treatment for patients with rheumatoid arthritis (RA) is well reported. However, far less information on quantity and quality of day-to-day physical activity is available. RA causes pain and stiffness that discourages physical movement and reduces mobility. The aim of our study is to demonstrate that continuous measurement of RA patient activity, recorded using a non-intrusive smartphones-based approach, can be used as an objective and reliable input to current disease activity tools used.

Method:

Commercial activity monitors focus on quantitative data including steps count and calories burned. The novelty of our approach is to gather smartphone sensor data to analyse rates of acceleration (jerk) and calculate metrics related to impact and joint loading. Men and women aged 36-65 diagnosed with rheumatoid arthritis for 4-23 years completed a questionnaire and gave feedback during a focus group session to assess the design of a smartphone app and scope the requirements according to real RA patients.

Truly continuous logging of acceleration is too resource intensive and affecting everyday use of the smartphone. We therefore developed an intermittent random 'Monte-Carlo' strategy of sampling activity that runs in the background on the phone, unnoticed by the user. From this we can obtain statistically rigorous parameters relating to physical activity.

Using the app, patients can also record self-assessed pain and inflammation as well as perceived physical activity via daily questionnaire (BSR and RAPID3). The user interface also includes self-reporting of the DAS28 count, with addition of an intensity factor introduced to highlight how painful and/or swollen a joint is (0-4). Activity and self-assessment are regularly uploaded to a cloud based Azure server to mitigate the smartphone's storage limitation and to allow analysis/visualization of results through a web-based interface.

Results:

The app is currently being trialled by patients (six weeks) and we have preliminary data for diagnosed RA and non-RA active participants. Commercial pedometer worn at the wrist shows a daily average of step count, sedentary and active time. The activity recorded through smartphone sensors shows different information and that the most frequent low activity level follows a linear trend ~0.25 N/kg/s for the RA patient and roughly x10 higher for the healthy participant. Beyond the region of peak activity, both distributions become linear (on the log/log scale) but with

noticeably different trends. Furthermore, the spikes of mean jerk recorded highlight the intensity of load generated that the joints withstand.

Conclusion:

It is possible to design an app looking at the quality of quantitative movement that is both usable and acceptable by RA patient. Preliminary data suggests the promise of using activity probability distribution to assess patient' physical activity.

Appendix L Git repository and best practices

Branch strategy:

1) Master:

To simplify and limit the setup changes, we keep the "Master" branch as the testing branch, i.e. this the branch of code to test with.

2) Release:

It's important to ensure that we have a working version at any point of time. The "Release" branch is to be kept as the branch of working version of the app. Once we all contributors agree with a feature (that means that is has been tested and confirmed working), then the feature branch can be merged to this Release branch.

3) Feature specific:

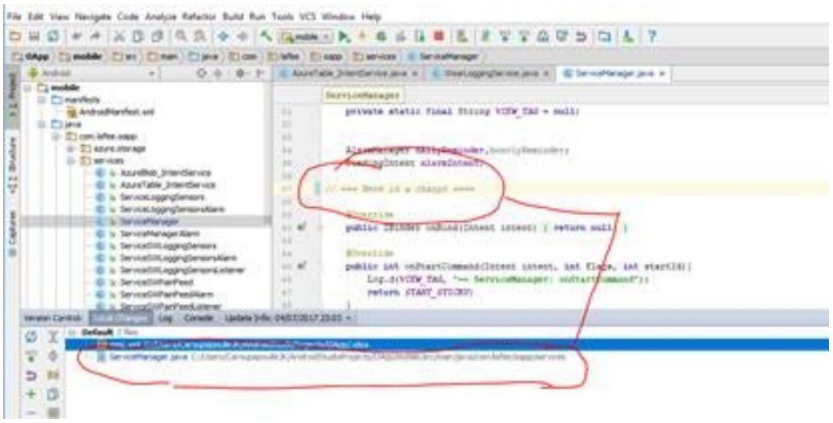
The changes are very difficult to track and in general, it is best practice to work on a specific features/work/fixes at a time and so for each new major work, a new branch should be added to be specific to that specific feature, e.g. new branches added specific to new db changes, service logic, SW notifications…

ADT/Git:

There is no need to reinstall or create a new project each time, tracking can be done on the local version.

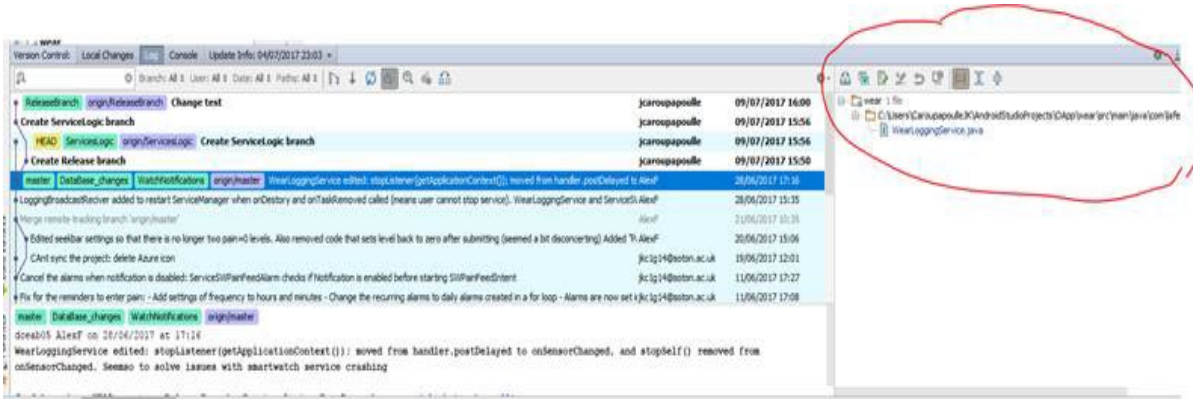
1) ***Version control:***
   Very useful to check that the code on the local environment is the code you meant to be.

   - **Local changes tab:**
     Open the "version control" view at the bottom of ADT and if you make a change to the code in your local environment, it will straight away.

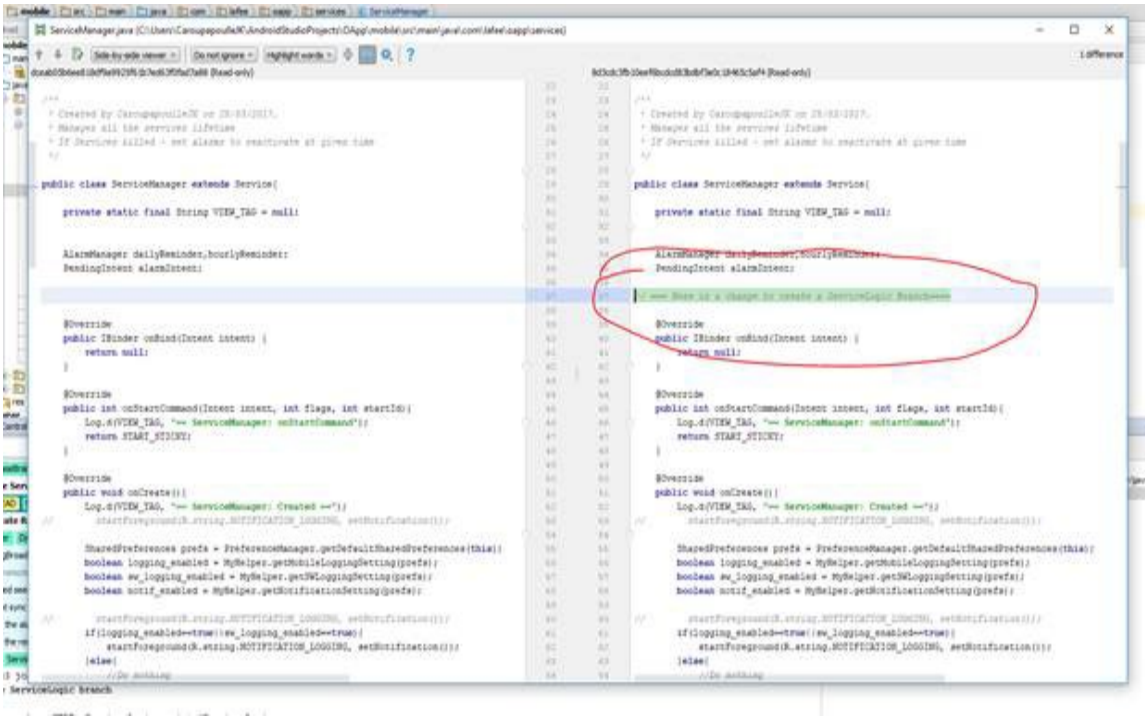


   - **Log tab:**
     This shows you all the branch versioning.

If you click on any of the commits, you can see the details on the right panel.



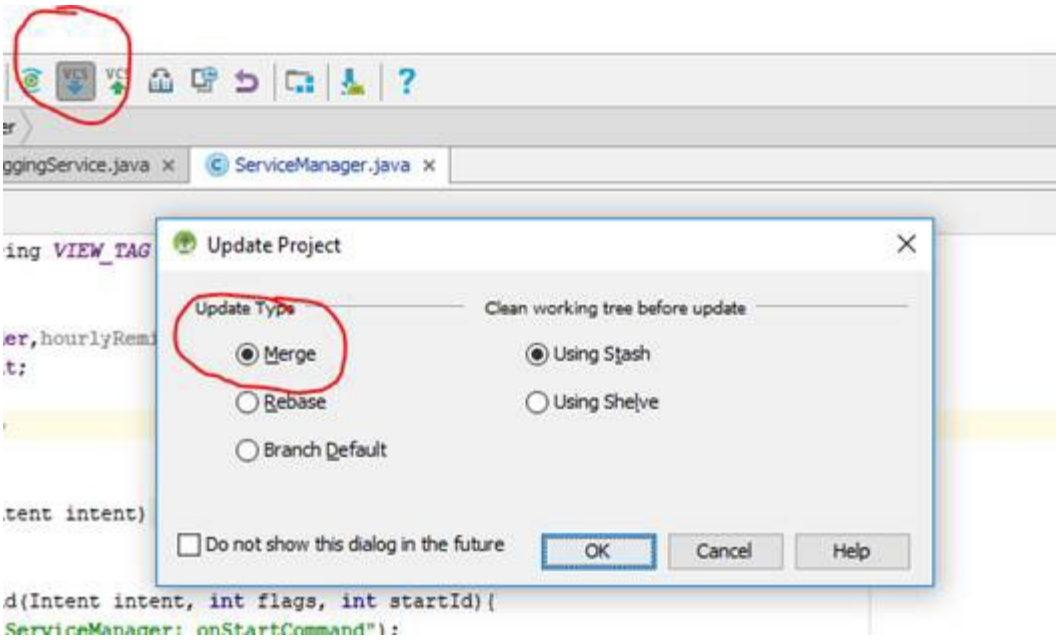
- **Files diff:**
  Double click on the file that you want to further look at from the right panel will open a diff panel so you can browse the highlighted changes.



2) ***Update the project:***

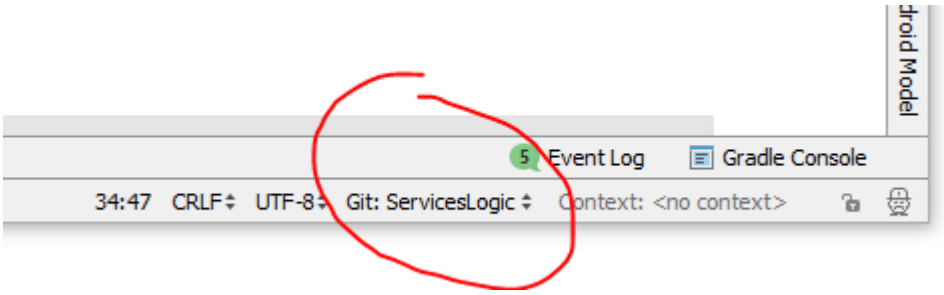
To pull the code, the blue ADT shortcut can be used. Assuming being in the correct local folder, merge will make sure to have the latest code from the said branch.

- Merge:
  Eq to git fetch and git merge
- Rebase:
  Eq to git fetch and git rebase



3) ***Branch selection:***

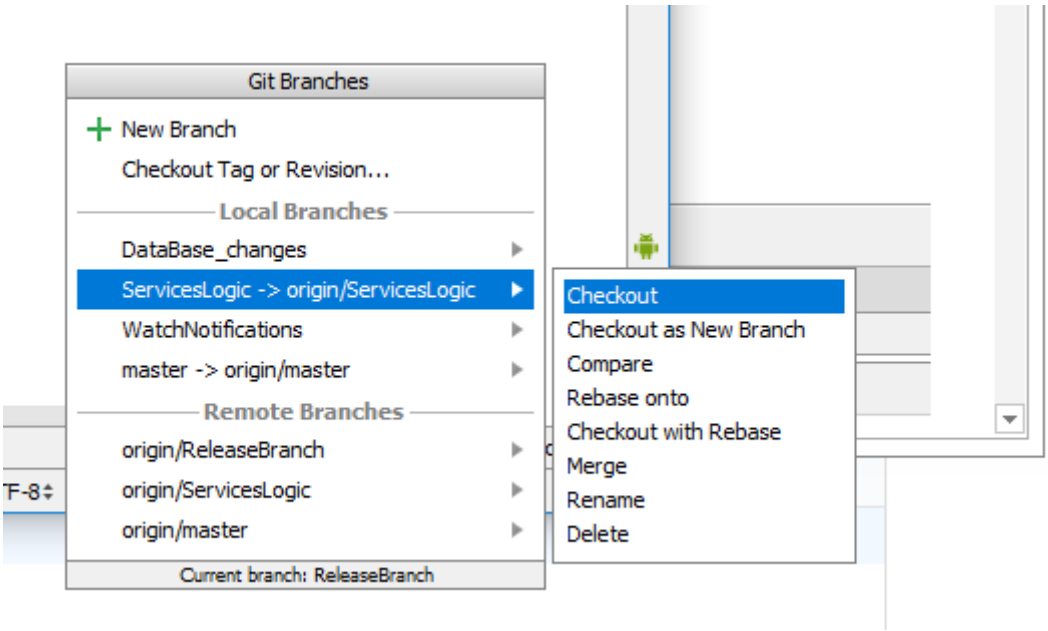
The Git branch that worked on can be seen at the bottom right of ADT.



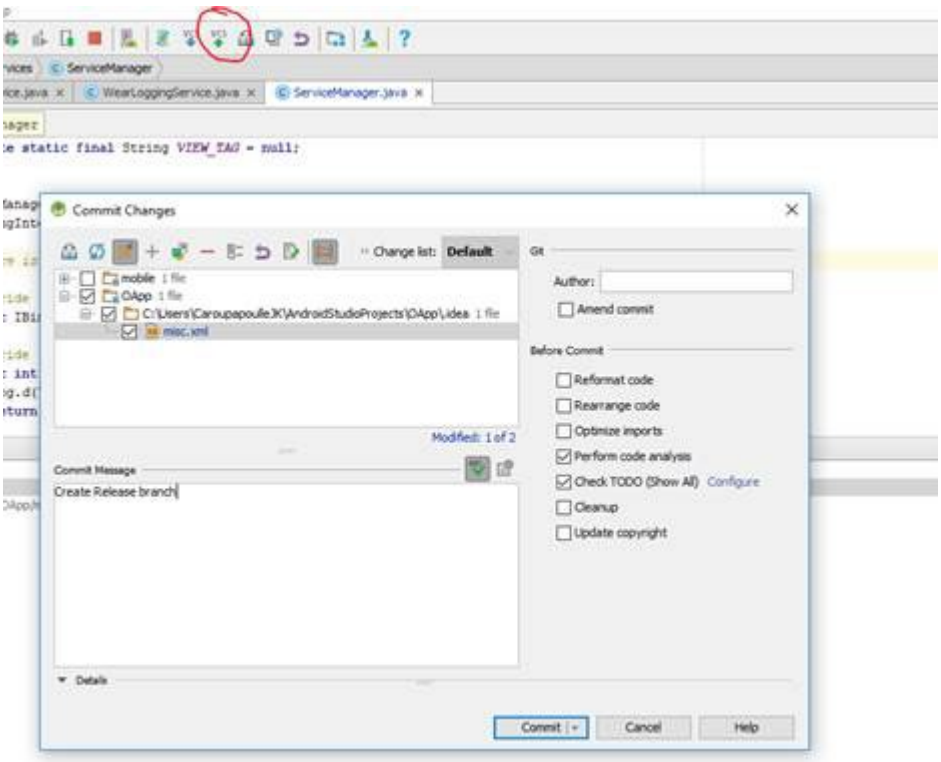
From there, you can switch between branches:

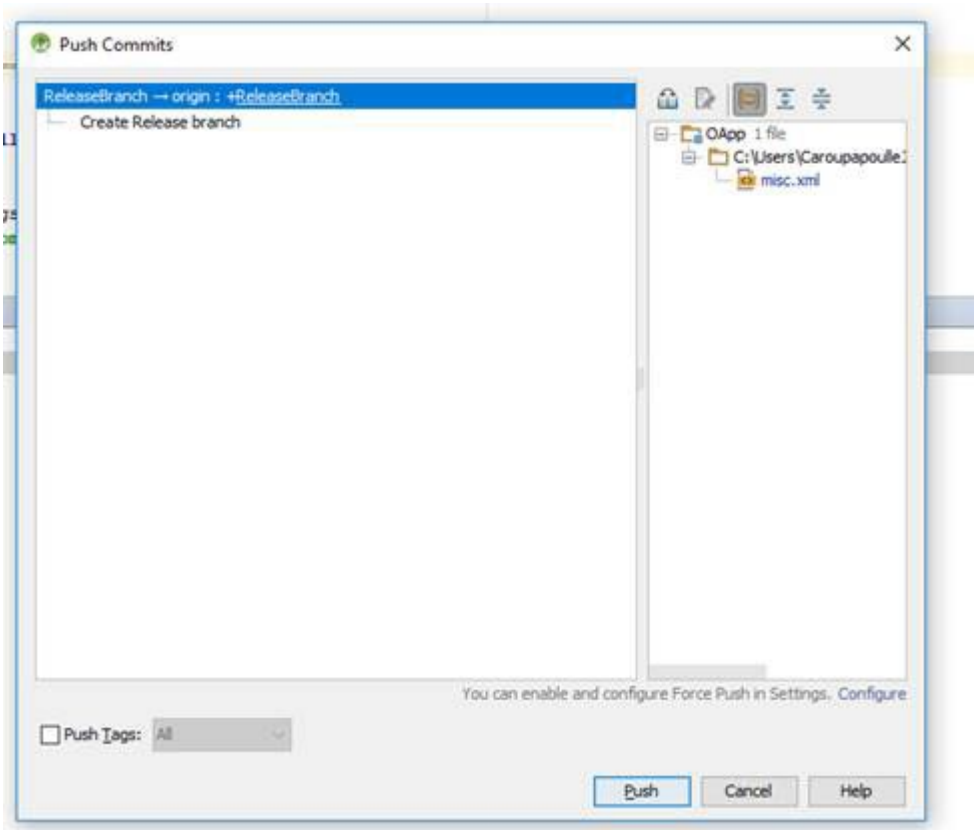
- "Local branches":
  If working on multiple features at a time, allows to switch the local version.
- "Remote branches":
  This is what actually does get uploaded to Git.
- "Checkout":

  Loads the version of the files associated to the branch selected. It will prompt to save/commit your changes if you try to checkout from git while having worked on something.



ADT shortcut to commit the changes to git.





ADT/Git setup:

1) Sign up and create a GitHub account in www.github.com.
2) Download git from https://git-scm.com/downloads and install it in your system.
3) Open the project in android studio and go to File -> Settings -> Version Control -> Git.
4) Click on test button to test "path to Git executables". If successful message is shown everything is ok, else navigate to git.exe from where you installed git and test again.
5) Go to File -> Settings -> Version Control -> GitHub. Enter your email and password used to create GitHub account and click on OK button.
6) Then go to VCS -> Import into Version Control -> Share Project on GitHub. Enter Repository name, Description and click Share button.
7) In the next window check all files inorder to add files for initial commit and click OK.
8) Now the project will be uploaded to the GitHub repository and when uploading is finished we will get a message in android studio showing "Successfully shared project on GitHub". Click on the link provided in that message to go to GitHub repository.

Bibliography

ABIRESEARCH. 2019. *Smartphones, Mobile Broadband & Wearables | ABI Research* [Online]. @abiresearch. Available: https://www.abiresearch.com/market-research/service/smartphones-mobile-broadband-wearables/ [Accessed].

ADAM NOAH, J., SPIERER, D. K., GU, J. & BRONNER, S. 2013. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *J Med Eng Technol,* 37**,** 456-62.

AGRAWAL, A. 2019. Musculoskeletal etymology: What's in a name? *Journal of Clinical Orthopaedics and Trauma,* 10**,** 387-394.

AINSWORTH, B. E., HASKELL, W. L., LEON, A. S., JACOBS, D. R., MONTOYE, H. J., SALLIS, J. F. & PAFFENBARGER, R. S. 1993. Compendium of Physical Activities: classification of energy costs of human physical activities. *Medicine & Science in Sports & Exercise,* 25**,** 71-80.

AINSWORTH, B. E., HASKELL, W. L., WHITT, M. C., IRWIN, M. L., SWARTZ, A. M., STRATH, S. J., O'BRIEN, W. L., BASSETT, D. R., SCHMITZ, K. H., EMPLAINCOURT, P. O., JACOBS, D. R. & LEON, A. S. 2000. Compendium of Physical Activities: an update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise,* 32**,** S498-S516.

ÅKERBERG, A., SÖDERLUND, A. & LINDÉN, M. 2016. Investigation of the validity and reliability of a smartphone pedometer application. *European Journal of Physiotherapy,* 18**,** 185-193.

ALAIN SARAUX, J. M. B., GÉRARD CHALÈS, CATHERINE LE HENAFF, JEAN B. THOREL, SYLVIE HOANG, ISABELLE VALLS, VALÉRIE DEVAUCHELLE, ANTOINE MARTIN, DOMINIQUE BARON, YVON PENNEC 2001. Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later - Saraux - 2001 - Arthritis & Rheumatism - Wiley Online Library.

APPLE. 2018. *Apple - Support - Technical Specifications* [Online]. Available: https://support.apple.com/specs/iphone [Accessed].

APPLE. 2019. *Business Models and Monetization - App Store - Apple Developer* [Online]. Available: https://developer.apple.com/app-store/business-models/ [Accessed].

ARTHRITISRESEARCHUK 2019. Introduction to musculoskeletal assessment.

ARTHRITISRESEARCHUK 2021. State of Musculoskeletal Health 2021.

ASSOCIATION, A. R. 1988. ACR 1987.

ATALLAH, L., LO, B., KING, R. & YANG, G.-Z. 2011. Sensor Positioning for Activity Recognition Using Wearable Accelerometers. *IEEE Transactions on Biomedical Circuits and Systems,* 5**,** 320-329.

ATTILA REISS, D. S. 2013. Aerobic activity monitoring: towards a longterm approach. *UNIVERSAL ACCESS IN THE INFORMATION SOCIETY*.

AUSTIN, AFTON, L. H., GUY, S., MICHAELA, S.-M., JOHAN, W. S. V., CÉSAR, F.-D.-L.-P., SARAH, J. R., STEFAN BERGMAN, R. G., CÉCILE, L. O., ROBIN, C., PERNILLA, Å., SUSANA, C., KAREN, L. H., MAI ELIN, P. H., ALBÈRE, J. A. K., ZOE, P., IRENE, A. P., CARINE, S. & JUDITH 2018. EULAR recommendations for the health professional's approach to pain management in inflammatory arthritis and osteoarthritis.

AUSTRALIA, M. 2021. *Musculoskeletal Australia* [Online]. Available: https://msk.org.au/ [Accessed].

AYHAN, E. 2014. Intraarticular injections (corticosteroid, hyaluronic acid, platelet rich plasma) for the knee osteoarthritis. *World Journal of Orthopedics,* 5**,** 351.

AZEVEDO, R., BERNARDES, M., FONSECA, J. & LIMA, A. 2015. Smartphone application for rheumatoid arthritis self-management: cross-sectional study revealed the usefulness, willingness to use and patients' needs. *Rheumatol Int*.

BALOGUN, J. A., AMUSA, L. O. & ONYEWADUME, I. U. 1988. Factors affecting Caltrac and Calcount accelerometer output. *Phys Ther,* 68**,** 1500-4.

BANNURU, R. R., OSANI, M. C., VAYSBROT, E. E., ARDEN, N. K., BENNELL, K., BIERMA-ZEINSTRA, S. M. A., KRAUS, V. B., LOHMANDER, L. S., ABBOTT, J. H., BHANDARI, M., BLANCO, F. J., ESPINOSA, R., HAUGEN, I. K., LIN, J., MANDL, L. A., MOILANEN, E., NAKAMURA, N., SNYDER-MACKLER, L., TROJIAN, T., UNDERWOOD, M. & MCALINDON, T. E. 2019. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthritis and Cartilage,* 27**,** 1578-1589.

BASSETT, D. R., TOTH, L. P., LAMUNION, S. R. & CROUTER, S. E. 2017. Step Counting: A Review of Measurement Considerations and Health-Related Applications. *Sports Medicine,* 47**,** 1303-1315.

BEKHET, A. K. & ZAUSZNIEWSKI, J. A. 2013. Methodological triangulation: an approach to understanding data. *Nurse Res,* 20**,** 40-3.

BENEDETTI, M. G., FURLINI, G., ZATI, A. & LETIZIA MAURO, G. 2018. The Effectiveness of Physical Exercise on Bone Density in Osteoporotic Patients. *BioMed Research International,* 2018**,** 4840531.

BERNHARDT, K. A., BECK, L. A., LAMB, J. L., KAUFMAN, K. R., AMIN, S. & WUERMSER, L.-A. 2012. Weight Bearing Through Lower Limbs in a Standing Frame with and Without Arm Support and Low-Magnitude Whole-Body Vibration in Men and Women with Complete Motor Paraplegia. *American Journal of Physical Medicine & Rehabilitation,* 91**,** 300-308.

BOBBERT, M. F. 2014. Effect of Unloading and Loading on Power in Simulated Countermovement and Squat Jumps. *Medicine & Science in Sports & Exercise,* 46.

BOUTEN, C. V. C., SAUREN, A. A. H. J., VERDUIN, M. & JANSSEN, J. D. 1997. Effects of placement and orientation of body-fixed accelerometers on the assessment of energy expenditure during walking. *Medical & Biological Engineering & Computing,* 35**,** 50-56.

BOYD, L. J., BALL, K. & AUGHEY, R. J. 2011. The Reliability of MinimaxX Accelerometers for Measuring Physical Activity in Australian Football. *International Journal of Sports Physiology and Performance,* 6**,** 311-321.

BRAJDIC, A. & HARLE, R. Walk detection and step counting on unconstrained smartphones. 2013 2013. ACM.

BRAMBILLA, C., MALOSIO, M., RENI, G. & SCANO, A. 2022. Optimal Biomechanical Performance in Upper-Limb Gestures Depends on Velocity and Carried Load. *Biology (Basel),* 11.

BREWER, W., SWANSON, B. T. & ORTIZ, A. 2017. Validity of Fitbit's active minutes as compared with a research-grade accelerometer and self-reported measures. *BMJ Open Sport & Exercise Medicine,* 3**,** e000254.

BROOKE-WAVELL, K., SKELTON, D. A., BARKER, K. L., CLARK, E. M., DE BIASE, S., ARNOLD, S., PASKINS, Z., ROBINSON, K. R., LEWIS, R. M., TOBIAS, J. H., WARD, K. A., WHITNEY, J. & LEYLAND, S. 2022. Strong, steady and straight: UK consensus statement on physical activity and exercise for osteoporosis. *British Journal of Sports Medicine,* 56**,** 837-846.

CAMERON, K. L., DRIBAN, J. B. & SVOBODA, S. J. 2016. Osteoarthritis and the Tactical Athlete: A Systematic Review. *Journal of Athletic Training,* 51**,** 952-961.

CASPERSEN, C. J., POWELL, K. E. & CHRISTENSON, G. M. 1985. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep,* 100**,** 126-31.

CHANDRASEKARAN, R., KATTHULA, V. & MOUSTAKAS, E. 2020. Patterns of Use and Key Predictors for the Use of Wearable Health Care Devices by US Adults: Insights from a National Survey. *J Med Internet Res,* 22**,** e22443.

CHARAN, J. & BISWAS, T. 2013. How to calculate sample size for different study designs in medical research? *Indian journal of psychological medicine,* 35**,** 121-126.

CHOI, B. C. K., PAK, A. W. P. & CHOI, J. C. L. 2007. Daily step goal of 10,000 steps: A literature review. *Clinical & Investigative Medicine,* 30**,** 146.

DALSKY, G. P. 1987. Exercise: its effect on bone mineral content. *Clin Obstet Gynecol,* 30**,** 820-32.

DAOUD, A. I., GEISSLER, G. J., WANG, F., SARETSKY, J., DAOUD, Y. A. & LIEBERMAN, D. E. 2012. Foot Strike and Injury Rates in Endurance Runners: A Retrospective Study. *Medicine & Science in Sports & Exercise,* 44.

DAVIS, B. L. & CAVANAGH, P. R. 1993. Simulating reduced gravity: a review of biomechanical issues pertaining to human locomotion. *Aviat Space Environ Med,* 64**,** 557-66.

DELFIEN VAN DYCK, K. D. G., BENEDICTE DEFORCHE, JOHANNES RUIGE, JACQUES BOUCKAERT, CATRINE E. TUDOR-LOCKE, JEAN-MARC KAUFMAN, ILSE DE BOURDEAUDHUIJ 2013. The relationship between changes in steps/day and health outcomes after a pedometer-based physical activity intervention with telephone support in type 2 diabetes patients. *Health Education Research,* 28**,** 539-545.

DIAZ, K. M., KRUPKA, D. J., CHANG, M. J., PEACOCK, J., MA, Y., GOLDSMITH, J., SCHWARTZ, J. E. & DAVIDSON, K. W. 2015. Fitbit(R): An accurate and reliable device for wireless physical activity tracking. *Int J Cardiol,* 185**,** 138-40.

DIERICK, F., PENTA, M., RENAUT, D. & DETREMBLEUR, C. 2004. A force measuring treadmill in clinical gait analysis. *Gait & Posture,* 20**,** 299-303.

DIRICAN, A. & AKSOY, S. 2017. Step Counting Using Smartphone Accelerometer and Fast Fourier Trransform. *Sigma Journal of Engineering and Natural Sciences,* 8**,** 175-182.

DIXON, W. G. & MICHAUD, K. 2018. Using technology to support clinical care and research in rheumatoid arthritis. *Curr Opin Rheumatol,* 30**,** 276-281.

DONTJE, M. L., DE GROOT, M., LENGTON, R. R., VAN DER SCHANS, C. P. & KRIJNEN, W. P. 2015. Measuring steps with the Fitbit activity tracker: an inter-device reliability study. *J Med Eng Technol***,** 1-5.

DYRSTAD, S. M., HANSEN, B. H., HOLME, I. M. & ANDERSSEN, S. A. 2013. Comparison of self-reported versus accelerometer-measured physical activity. *Med Sci Sports Exerc,* 46**,** 99-106.

EDGE. 2018. *Intelligent Research Management* [Online]. Available: https://www.edge.nhs.uk [Accessed].

ELVAN, A. & OZYUREK, S. 2020. Principles of kinesiology. Elsevier.

ERIC B HEKLER, M. P. B., LAUREN GRIECO, MARY ROSENBERGER, SANDRA J WINTER, WILLIAM HASKELL, ABBY C KING 2015. Validation of Physical Activity Tracking via Android Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living Validation Studies. *JMIR Mhealth Uhealth*.

ETSI 2016a. 3GPP 24.008 - Mobile radio interface Layer 3 specification; Core network protocols; Stage 3.

ETSI 2016b. 3GPP 25.331 - Universal Mobile Telecommunications System (UMTS); Radio Resource Control (RRC); Protocol specification

EVENSON, K. R., GOTO, M. M. & FURBERG, R. D. 2015. Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity,* 12**,** 159.

FAUTREL, S. H., BENJAMIN, H., STÉPHANE, M., LAURE, G. & BRUNO 2017. Ability of disease-modifying antirheumatic drugs to prevent or delay rheumatoid arthritis onset: a systematic literature review and meta-analysis.

FEEHAN, L. M., GELDMAN, J., SAYRE, E. C., PARK, C., EZZAT, A. M., YOO, J. Y., HAMILTON, C. B. & LI, L. C. 2018. Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR mHealth and uHealth,* 6**,** e10527.

FELSON, D. T., GOGGINS, J., NIU, J., ZHANG, Y. & HUNTER, D. J. 2004. The effect of body weight on progression of knee osteoarthritis is dependent on alignment. *Arthritis and rheumatism,* 50 12**,** 3904-9.

FENG, J. M. D., CHEN, Q. P., YU, F. P., WANG, Z. P., CHEN, S. M. D., JIN, Z. P., CAI, Q. M. D. P., LIU, Y. E. B. & HE, J. M. D. P. 2016. Body Mass Index and Risk of Rheumatoid Arthritis: A Meta-Ana... : Medicine.

FERGUSON, T., ROWLANDS, A. V., OLDS, T. & MAHER, C. 2015. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act,* 12**,** 42.

FITABASE 2022. Fitabase.

FITBIT. 2016. *Charge HR* [Online]. Available: https://www.fitbit.com/be/chargehr [Accessed].

FREEMAN, S. 2017. *Smartphone apps may aid home rheumatoid arthritis monitoring* [Online]. MDedge. Available: https://www.mdedge.com/rheumatology/article/138166/rheumatoid-arthritis/smartphone-apps-may-aid-home-rheumatoid-arthritis [Accessed].

FÜZÉKI, E., ENGEROFF, T. & BANZER, W. 2017. Health Benefits of Light-Intensity Physical Activity: A Systematic Review of Accelerometer Data of the National Health and Nutrition Examination Survey (NHANES). *Sports Medicine,* 47**,** 1769-1793.

GABEL, L., LIPHARDT, A.-M., HULME, P. A., HEER, M., ZWART, S. R., SIBONGA, J. D., SMITH, S. M. & BOYD, S. K. 2022. Incomplete recovery of bone strength and trabecular microarchitecture at the distal tibia 1 year after return from long duration spaceflight. *Scientific Reports,* 12.

GAL, R., MAY, A. M., VAN OVERMEEREN, E. J., SIMONS, M. & MONNINKHOF, E. M. 2018. The Effect of Physical Activity Interventions Comprising Wearables and Smartphone Applications on Physical Activity: a Systematic Review and Meta-analysis. *Sports Medicine - Open,* 4.

GATES, L. S., LEYLAND, K. M., SHEARD, S., JACKSON, K., KELLY, P., CALLAHAN, L. F., PATE, R., ROOS, E. M., AINSWORTH, B., COOPER, C., FOSTER, C., NEWTON, J. L., BATT, M. E. & ARDEN, N. K. 2017. Physical activity and osteoarthritis: a consensus study to harmonise self-reporting methods of physical activity across international cohorts. *Rheumatology International,* 37**,** 469-478.

GATO-CALVO, L., MAGALHAES, J., RUIZ-ROMERO, C., BLANCO, F. J. & BURGUERA, E. F. 2019. Platelet-rich plasma in osteoarthritis treatment: review of current evidence. *Therapeutic Advances in Chronic Disease,* 10**,** 204062231982556.

GECHT, M. R., CONNELL, K. J., SINACORE, J. M. & PROHASKA, T. R. 1996. A survey of exercise beliefs and exercise habits among people with arthritis. *Arthritis Care Res,* 9**,** 82-8.

GIANGREGORIO, L. M., PAPAIOANNOU, A., MACINTYRE, N. J., ASHE, M. C., HEINONEN, A., SHIPP, K., WARK, J., MCGILL, S., KELLER, H., JAIN, R., LAPRADE, J. & CHEUNG, A. M. 2014. Too Fit To Fracture: exercise recommendations for individuals with osteoporosis or osteoporotic vertebral fracture. *Osteoporosis International,* 25**,** 821-835.

GOMEZ, R., BOULOS, H., JOACHIM, K., EDWARD, C. K., TORE, K. K., IAIN, M., EMILIO, M.-M., CARLOMAURIZIO, M., MONIKA, S., DESIRÉE VAN DER HEIJDE, J. S. S., DANIEL, A., JOHANNES, W. J. B., FERDINAND, C. B., DIMITRIOS, B., GERD, B., BERNARD, C., MAURIZIO, C., MAARTEN DE, W., MAXIME, D., PAUL, E., ALAN, G. & JUAN, J. 2010. Treating rheumatoid arthritis to target: recommendations of an international task force.

GOOGLE. 2016. *Sensors interaction* [Online]. Available: https://source.android.com/devices/sensors/interaction [Accessed].

GOOGLE 2018a. Compatibility Definition.

GOOGLE. 2018b. *Sensor types* [Online]. Available: https://source.android.com/devices/sensors/sensor-types [Accessed].

GOOGLE. 2022a. *Android Doc - Use the step counter sensor* [Online]. Available: https://developer.android.com/reference/android/hardware/Sensor#TYPE_STEP_COUNTER [Accessed].

GOOGLE. 2022b. *Android Power management restrictions* [Online]. Available: https://developer.android.com/topic/performance/power/power-details [Accessed].

GOV.UK 2017. Improving Lives. The Future of Work, Health and Disability. *In:* PENSIONS, D. F. W. A. (ed.).

GOV.UK 2018a. The Government's revised mandate to NHS England for 2017-18. *In:* CARE, D. O. H. A. S. (ed.).

GOV.UK 2018b. Medical devices: software applications (apps). *In:* CARE, H. A. S. (ed.).

GOV.UK. 2021. *Musculoskeletal health: applying All Our Health* [Online]. Available: https://www.gov.uk/government/publications/musculoskeletal-health-applying-all-our-health/musculoskeletal-health-applying-all-our-health [Accessed].

GRIFFIN, T. M. & GUILAK, F. 2005. The Role of Mechanical Loading in the Onset and Progression of Osteoarthritis. *Exercise and Sport Sciences Reviews,* 33.

GRIFFITHS, A. J., WHITE, C. M., THAIN, P. K. & BEARNE, L. M. 2018. The effect of interactive digital interventions on physical activity in people with inflammatory arthritis: a systematic review. *Rheumatology International,* 38**,** 1623-1634.

HARE, S. E. S., MICHAEL, F. C., DAVID, G. M., JEREMY, P., RALPH, H. G., ANDREW, D. W., VORAMONT, B., DEIDRE, T., HENRY, K. & DAVID, L. 2004. Moderate-intensity resistance exercise training in patients with chronic heart failure improves strength, endurance, heart rate variability, and forearm blood flow. *Journal of Cardiac Failure,* 10**,** 21-30.

HARRISON, M. J., BOONEN, A., TUGWELL, P. & SYMMONS, D. P. 2009. Same question, different answers: a comparison of global health assessments using visual analogue scales. *Qual Life Res,* 18**,** 1285-92.

HARTMAN, S. J., NELSON, S. H. & WEINER, L. S. 2018. Patterns of Fitbit Use and Activity Levels Throughout a Physical Activity Intervention: Exploratory Analysis from a Randomized Controlled Trial. *JMIR mHealth and uHealth,* 6**,** e29.

HASLAM, C. 2018. *Best workout apps we've used: improve your fitness in just 20 minutes a day* [Online]. @techradar. Available: https://www.techradar.com/news/best-fitness-apps-best-hiit-apps-best-strength-apps [Accessed].

HEIDARI, B. 2011. Rheumatoid Arthritis: Early diagnosis and treatment outcomes. *Caspian J Intern Med.*

HENRIKSEN, M., AABOE, J. & BLIDDAL, H. 2012. The relationship between pain and dynamic knee joint loading in knee osteoarthritis varies with radiographic disease severity. A cross sectional study. *The Knee,* 19**,** 392-398.

HERNANDEZ-HERNANDEZ, V., FERRAZ-AMARO, I. & DIAZ-GONZALEZ, F. 2014. Influence of disease activity on the physical activity of rheumatoid arthritis patients. *Rheumatology,* 53**,** 722-731.

HOOPER, C., PICKERING, B., PRICHARD, J. & ASHLEIGH, M. 2015. TRIFoRM Final Report: TRust in IT: Factors, metRics, Models.

HTC. *Using power saver mode* [Online]. Available: https://www.htc.com/mea-en/support/htc-10/howto/power-saver-mode.html [Accessed].

HURKMANS, H. L. P., BUSSMANN, J. B. J., BENDA, E., VERHAAR, J. A. N. & STAM, H. J. 2003. Techniques for measuring weight bearing during standing and walking. *Clinical Biomechanics,* 18**,** 576-589.

JENNIFER K. COONEY, R.-J. L., VERENA MATSCHKE, ANDREW B. LEMMEY, JONATHAN P. MOORE, YASMEEN AHMAD, JEREMY G. JONES, PETER MADDISON, JEANETTE M. THOM 2010. Benefits of Exercise in Rheumatoid Arthritis. *Journal of Aging Research*.

JETTÉ, M., SIDNEY, K. & BLÜMCHEN, G. 1990. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clinical Cardiology,* 13**,** 555-565.

KALLBERG, H., DING, B., PADYUKOV, L., BENGTSSON, C., RONNELID, J., KLARESKOG, L. & ALFREDSSON, L. 2010. Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Ann Rheum Dis,* 70**,** 508-11.

KANG, X., HUANG, B. & QI, G. 2018. A Novel Walking Detection and Step Counting Algorithm Using Unconstrained Smartphones. *Sensors,* 18**,** 297.

KANIS, J. A., DELMAS, P., BURCKHARDT, P., COOPER, C. & TORGERSON, D. 1997. Guidelines for diagnosis and management of osteoporosis. *Osteoporosis International,* 7**,** 390-406.

KAPOOR, M. C. 2016. Types of studies and research design. *Indian journal of anaesthesia,* 60**,** 626-630.

KARAS, M., BAI, J., STRĄCZKIEWICZ, M., HAREZLAK, J., GLYNN, N. W., HARRIS, T., ZIPUNNIKOV, V., CRAINICEANU, C. & URBANEK, J. K. 2019. Accelerometry Data in Health Research: Challenges and Opportunities. *Statistics in Biosciences,* 11**,** 210-237.

KATZ PP, M. M., GREGORICH S, KAPLAN S, RUSH S, TRUPIN L. 2015. A Randomized Controlled Trial for a Physical Activity Intervention for RA Fatigue [abstract]. *Arthritis Rheumatol.,* 67 (suppl 10).

KAY, J. & UPCHURCH, K. S. 2012. ACR/EULAR 2010 rheumatoid arthritis classification criteria. *Rheumatology,* 51.

KELL, R. T., BELL, G. & QUINNEY, A. 2001. Musculoskeletal Fitness, Health Outcomes and Quality of Life. *Sports Medicine,* 31**,** 863-873.

KERN-STEINER, R., WASHECHECK, H. S. & KELSEY, D. D. 1999. Strategy of Exercise Prescription Using an Unloading Technique for Functional Rehabilitation of an Athlete with an Inversion Ankle Sprain. *Journal of Orthopaedic & Sports Physical Therapy,* 29**,** 282-287.

KHAN, A. M., TUFAIL, A., KHATTAK, A. M. & LAINE, T. H. 2014. Activity Recognition on Smartphones via Sensor-Fusion and KDA-Based SVMs. *International Journal of Distributed Sensor Networks,* 10**,** 503291.

KHEDR, M. & EL-SHEIMY, N. 2017. A Smartphone Step Counter Using IMU and Magnetometer for Navigation and Health Monitoring Applications. *Sensors,* 17**,** 2573.

KILTZ, KEEGAN, K., MICHAEL, N., SANDRA, P., GUY, S., THIJS WILLEM, S., IRENE, A. P., LOUISE, W., ZHIVKO, Y., THEODORA P M VLIET VLIELAND, A.-K. R. O., KARIN, N., JÜRGEN, B., JO, A., NINA, B., HANNE, D., TUNCAY, D., BENTE APPEL, E., KLAUS-PETER, G., EMAILIE, H., CARSTEN BOGH, J., NORELEE, K. & UTA 2018. 2018 EULAR recommendations for physical activity in people with inflammatory arthritis and osteoarthritis.

KINH, T., TU, L. & TIEN, D. A high-accuracy step counting algorithm for iPhones using accelerometer. 2012 2012. IEEE.

KIRSTEN JACK, S. M. M., JENNIFER KLABER MOFFETT, ERIC GARDINER 2010. Barriers to treatment adherence in physiotherapy outpatient clinics: A systematic review. *Manual therapy*.

KUJALA, U. M., KAPRIO, J. & SARNO, S. 1994. Osteoarthritis of weight bearing joints of lower limbs in former elite male athletes. *BMJ,* 308**,** 231-234.

LEACH, M. Z. 2018. *5 Things Patients With RA Want Their Doctors to Know* [Online]. rheumatologynetwork.com. Available: [http://www.rheumatologynetwork.com/rheumatoid-arthritis/5-things-patients-ra-want-their-doctors-know [Accessed].

LEE, H.-H., CHOI, S. & LEE, M.-J. 2015. Step Detection Robust against the Dynamics of Smartphones. *Sensors,* 15**,** 27230-27250.

LEON, A. C., DAVIS, L. L. & KRAEMER, H. C. 2011. The role and interpretation of pilot studies in clinical research. *Journal of psychiatric research,* 45**,** 626-629.

MANCUSO, C. A., RINCON, M., SAYLES, W. & PAGET, S. A. 2007. Comparison of energy expenditure from lifestyle physical activities between patients with rheumatoid arthritis and healthy controls. *Arthritis & Rheumatism,* 57**,** 672-678.

MARCZYK, G., DEMATTEO, D. & FESTINGER, D. 2005. *Essentials of research design and methodology,* Hoboken, NJ, US, John Wiley & Sons Inc.

MATHIE, M. J., COSTER, A. C. F., LOVELL, N. H. & CELLER, B. G. 2004. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement,* 25**,** R1-R20.

MAYAGOITIA, R. E., LÖTTERS, J. C., VELTINK, P. H. & HERMENS, H. 2002. Standing balance evaluation using a triaxial accelerometer. *Gait & Posture,* 16**,** 55-59.

MCCALLUM, C., ROOKSBY, J. & GRAY, C. M. 2018. Evaluating the Impact of Physical Activity Apps and Wearables: Interdisciplinary Review. *JMIR mHealth and uHealth,* 6**,** e58.

MCLEAN, B. D., CUMMINS, C., CONLAN, G., DUTHIE, G. & COUTTS, A. J. 2018. The Fit Matters: Influence of Accelerometer Fitting and Training Drill Demands on Load Measures in Rugby League Players. *Int J Sports Physiol Perform,* 13**,** 1083-1089.

METSIOS, G. S., STAVROPOULOS-KALINOGLOU, A., VELDHUIJZEN VAN ZANTEN, J. J. C. S., NIGHTINGALE, P., SANDOO, A., DIMITROULAS, T., KITAS, G. D. & KOUTEDAKIS, Y. 2014. Individualised exercise improves endothelial function in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases,* 73**,** 748.

MEY, N. D. 2017. *How to Make Money With An App (and other business model ideas to explore) - Board of Innovation* [Online]. @boardofinno. Available: https://www.boardofinnovation.com/blog/how-to-make-money-app-business-model-ideas/ [Accessed].

MICROSOFT. 2017a. *Azure certifications* [Online]. Available: https://azure.microsoft.com/en-gb/overview/trusted-cloud [Accessed].

MICROSOFT. 2017b. *Azure for health* [Online]. Available: https://azure.microsoft.com/en-gb/industries/healthcare/ [Accessed].

MICROSOFT. 2017c. *Azure storage* [Online]. Available: https://docs.microsoft.com/en-us/azure/storage/ [Accessed].

MICROSOFT. 2017d. *MyDriving* [Online]. Available: https://azure.microsoft.com/en-gb/campaigns/mydriving [Accessed].

MILNER, C. E., FERBER, R., POLLARD, C. D., HAMILL, J. & DAVIS, I. S. 2006. Biomechanical factors associated with tibial stress fracture in female runners. *Med Sci Sports Exerc,* 38**,** 323-8.

MORCOS, M. W., TEETER, M. G., SOMERVILLE, L. E. & LANTING, B. 2020. Correlation between hip osteoarthritis and the level of physical activity as measured by wearable technology and patient-reported questionnaires. *Journal of Orthopaedics,* 20**,** 236-239.

MYO, W. W., WETTAYAPRASIT, W. & AIYARAK, P. 2018. A More Reliable Step Counter using Built-in Accelerometer in Smartphone. *Indonesian Journal of Electrical Engineering and Computer Science,* 12**,** 775-782.

NAO 2009. Services for people with rheumatoid arthritis. HC 823 Session 2008-2009 ed. National Audit Office.

NATIONAL ACADEMIES OF SCIENCES, E., AND MEDICINE 2020. *Selected Health Conditions and Likelihood of Improvement with Treatment*.

NAZIRIZADEH, S. 2018. *Load Rate Estimates on the Lower Limb Joints Using the Accelerometer of Wearable During Physical Activity.* Doctor of Philosophy, UNIVERSITY OF SOUTHAMPTON.

NEWMAN, D. J., ALEXANDER, H. L. & WEBBON, B. W. 1994. Energetics and mechanics for partial gravity locomotion. *Aviat Space Environ Med,* 65**,** 815-23.

NHS. 2018. *Benefits of exercise* [Online]. Available: https://www.nhs.uk/live-well/exercise/exercise-health-benefits/#health-benefits [Accessed].

NRAS. 2014. *The Importance of Exercise* [Online]. Available: https://www.nras.org.uk/the-importance-of-exercise [Accessed].

O'FALLON, S. E. G. C. S. C. M. 2001. The epidemiology of rheumatoid arthritis in Rochester, Minnesota, 1955–1985 - Gabriel - 1999 - Arthritis & Rheumatism - Wiley Online Library.

O'BRIEN, M. 2001. Exercise and osteoporosis. *Irish Journal of Medical Science,* 170**,** 58-62.

OZDALGA, E., OZDALGA, A. & AHUJA, N. 2012. The smartphone in medicine: a review of current and potential use among physicians and students. *J Med Internet Res,* 14**,** e128.

OZLEM PALA, F. C. 2006. Joint Counts.

PANUSH, R. S. & LANE, N. E. 1994. Exercise and the musculoskeletal system. *Baillière's Clinical Rheumatology,* 8**,** 79-102.

PATERSON, K. L., HUNTER, D. J., METCALF, B. R., EYLES, J., DUONG, V., KAZSA, J., WANG, Y., BUCHBINDER, R., CICUTTINI, F., FORBES, A., HARRIS, A., YU, S. P., WANG, B. H., CONNELL, D., LINKLATER, J. & BENNELL, K. L. 2018. Efficacy of intra-articular injections of platelet-rich plasma as a symptom- and disease-modifying treatment for knee osteoarthritis - the RESTORE trial protocol. *BMC Musculoskeletal Disorders,* 19.

PAUL, S. S., TIEDEMANN, A., HASSETT, L. M., RAMSAY, E., KIRKHAM, C., CHAGPAR, S. & SHERRINGTON, C. 2015. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. *BMJ Open Sport & Exercise Medicine,* 1**,** e000013.

PRENDERGAST, P. J. & HUISKES, R. 1995. The Biomechanics of Wolff's law: Recent advances. *Irish Journal of Medical Science,* 164**,** 152-154.

PRINCE, S. A., ADAMO, K. B., HAMEL, M., HARDT, J., CONNOR GORBER, S. & TREMBLAY, M. 2008. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity,* 5**,** 56.

PRIORESCHI, A., HODKINSON, B., AVIDON, I., TIKLY, M. & MCVEIGH, J. A. 2013. The clinical utility of accelerometry in patients with rheumatoid arthritis. *Rheumatology,* 52**,** 1721-1727.

QUALCOMM. 2013. *designing-mobile-devices-for-low-power-and-thermal-efficiency* [Online]. Available: https://www.qualcomm.com/products/snapdragon/battery-life [Accessed].

QUALCOMM. 2014. *Behind the sixth sense of smartphones: the Snapdragon processor sensor engine | Qualcomm* [Online]. Available:

https://www.qualcomm.com/news/onq/2014/04/24/behind-sixth-sense-smartphones-snapdragon-processor-sensor-engine [Accessed].

QUALCOMM. 2022. *Quick Charge* [Online]. Available: https://www.qualcomm.com/products/features/quick-charge [Accessed].

READE, S., SPENCER, K., SERGEANT, J. C., SPERRIN, M., SCHULTZ, D. M., AINSWORTH, J., LAKSHMINARAYANA, R., HELLMAN, B., JAMES, B., MCBETH, J., SANDERS, C. & DIXON, W. G. 2017. Cloudy with a Chance of Pain: Engagement and Subsequent Attrition of Daily Data Entry in a Smartphone Pilot Study Tracking Weather, Disease Severity, and Physical Activity in Patients With Rheumatoid Arthritis. *JMIR Mhealth Uhealth,* 5**,** e37.

REDMAYNE, M. 2017. Where's Your Phone? A Survey of Where Women Aged 15-40 Carry Their Smartphone and Related Risk Perception: A Survey and Pilot Study. *PLOS ONE,* 12**,** e0167996.

RIVERA, J. C., WENKE, J. C., BUCKWALTER, J. A., FICKE, J. R. & JOHNSON, A. E. 2012. Posttraumatic Osteoarthritis Caused by Battlefield Injuries. *Journal of the American Academy of Orthopaedic Surgeons,* 20**,** S64-S69.

ROSENBAUM, L. 2019. Should You Really Take 10,000 Steps A Day? Available from: https://blog.fitbit.com/should-you-really-take-10000-steps-a-day/.

ROWLANDS, A. V. & STILES, V. H. 2012. Accelerometer counts and raw acceleration output in relation to mechanical loading. *Journal of Biomechanics,* 45**,** 448-454.

RUFF, C., HOLT, B. & TRINKAUS, E. 2006. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *American Journal of Physical Anthropology,* 129**,** 484-498.

RUNNER'SWORLD. 2018. *The Best Running Apps for Tracking, Motivation, Music, and More* [Online]. @runnersworld. Available: https://www.runnersworld.com/gear/a20865699/best-running-apps/ [Accessed].

SALLIS, J. F. & SAELENS, B. E. 2000. Assessment of Physical Activity by Self-Report: Status, Limitations, and Future Directions. *Research Quarterly for Exercise and Sport,* 71**,** 1-14.

SCHRAM, B., ORR, R., POPE, R., CANETTI, E. & KNAPIK, J. 2020. Risk factors for development of lower limb osteoarthritis in physically demanding occupations: A narrative umbrella review. *Journal of Occupational Health,* 62.

SCOFIELD, D. E. & KARDOUNI, J. R. 2015. The Tactical Athlete: A Product of 21st Century Strength and Conditioning. *Strength & Conditioning Journal,* 37**,** 2-7.

SEEGER, C., VAN LAERHOVEN, K. & BUCHMANN, A. 2014. MyHealthAssistant: an event-driven middleware for multiple medical applications on a smartphone-mediated body sensor network. *IEEE J Biomed Health Inform,* 19**,** 752-60.

SEEGER, C., VAN LAERHOVEN, K. & BUCHMANN, A. 2015. MyHealthAssistant: An Event-driven Middleware for Multiple Medical Applications on a Smartphone-Mediated Body Sensor Network. *IEEE Journal of Biomedical and Health Informatics,* 19**,** 752-760.

SHAKOOR, N. & MOISIO, K. 2004. A biomechanical approach to musculoskeletal disease. *Best Pract Res Clin Rheumatol,* 18**,** 173-86.

SHANES, J. 2017. The evolution of Fitness trackers. Polar.

SHIN, G., JARRAHI, M. H., FEI, Y., KARAMI, A., GAFINOWITZ, N., BYUN, A. & LU, X. 2019. Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review. *Journal of Biomedical Informatics,* 93**,** 103153.

SIMILARWEB. 2019. *Follow the leaders: highest ranking Apps in Google Play Store, United Kingdom* [Online]. @similarweb. Available: https://www.similarweb.com [Accessed].

SOKKA, T., HAKKINEN, A., KAUTIAINEN, H., MAILLEFERT, J. F., TOLOZA, S., MORK HANSEN, T., CALVO-ALEN, J., ODING, R., LIVEBORN, M., HUISMAN, M., ALTEN, R., POHL, C., CUTOLO, M., IMMONEN, K., WOOLF, A., MURPHY, E., SHEEHY, C., QUIRKE, E., CELIK, S., YAZICI, Y., TLUSTOCHOWICZ, W., KAPOLKA, D., SKAKIC, V., ROJKOVICH, B., MULLER, R., STROPUVIENE, S., ANDERSONE, D., DROSOS, A. A., LAZOVSKIS, J., PINCUS, T. & GROUP, Q.-R. 2008. Physical inactivity in patients with rheumatoid arthritis: data from twenty-one countries in a cross-sectional, international study. *Arthritis Rheum,* 59**,** 42-50.

SOLIGARD, T., SCHWELLNUS, M., ALONSO, J.-M., BAHR, R., CLARSEN, B., DIJKSTRA, H. P., GABBETT, T., GLEESON, M., HÄGGLUND, M., HUTCHINSON, M. R., RENSBURG, C. J. V., KHAN, K. M., MEEUSEN, R., ORCHARD, J. W., PLUIM, B. M., RAFTERY, M., BUDGETT, R. & ENGEBRETSEN, L. 2016. How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury.

SONY. 2019. *How can STAMINA mode and Extended Standby Mode reduce battery power consumption?* [Online]. Available: https://www.sony.com/electronics/support/articles/SX038901 [Accessed].

SOUTHAMPTON, C. R. I. 2016. *New smartphone app for arthritis patients* [Online]. UHS NHS. Available: http://www.uhs.nhs.uk/ClinicalResearchinSouthampton/Research/Facilities/NIHR-Southampton-Clinical-Research-Facility/News-and-updates/Articles/New-smartphone-app-for-arthritis-patients.aspx [Accessed].

SPECTOR, T. D., HARRIS, P. A., HART, D. J., CICUTTINI, F. M., NANDRA, D., ETHERINGTON, J., WOLMAN, R. L. & DOYLE, D. V. 1996. Risk of osteoarthritis associated with long-term weight-bearing sports: A radiologic survey of the hips and knees in female ex-athletes and population controls. *Arthritis & Rheumatism,* 39**,** 988-995.

STAVNICHUK, M., MIKOLAJEWICZ, N., CORLETT, T., MORRIS, M. & KOMAROVA, S. V. 2020. A systematic review and meta-analysis of bone loss in space travelers. *npj Microgravity,* 6.

STONE, A. A., SHIFFMAN, S., SCHWARTZ, J. E., BRODERICK, J. E. & HUFFORD, M. R. 2003. Patient compliance with paper and electronic diaries. *Control Clin Trials,* 24**,** 182-99.

SUGIYAMA, D., NISHIMURA, K., TAMAKI, K., TSUJI, G., NAKAZAWA, T., MORINOBU, A. & KUMAGAI, S. 2009. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Ann Rheum Dis,* 69**,** 70-81.

SYLVIA, L. G., BERNSTEIN, E. E., HUBBARD, J. L., KEATING, L. & ANDERSON, E. J. 2014. Practical Guide to Measuring Physical Activity. *Journal of the Academy of Nutrition and Dietetics,* 114**,** 199-208.

TAKACS, J., POLLOCK, C. L., GUENTHER, J. R., BAHAR, M., NAPIER, C. & HUNT, M. A. 2014. Validation of the Fitbit One activity monitor device during treadmill walking. *J Sci Med Sport,* 17**,** 496-500.

TUDOR-LOCKE, C. & BASSETT, D. R. 2004. How Many Steps/Day Are Enough? *Sports Medicine,* 34**,** 1-8.

TUDOR-LOCKE, C., CRAIG, C. L., BROWN, W. J., CLEMES, S. A., DE COCKER, K., GILES-CORTI, B., HATANO, Y., INOUE, S., MATSUDO, S. M., MUTRIE, N., OPPERT, J.-M., ROWE, D. A., SCHMIDT, M. D., SCHOFIELD, G. M., SPENCE, J. C., TEIXEIRA, P. J., TULLY, M. A. & BLAIR, S. N. 2011. How many steps/day are enough? for adults. *International Journal of Behavioral Nutrition and Physical Activity,* 8**,** 79.

TUDOR-LOCKE, C., JONES, G. R., MYERS, A. M., PATERSON, D. H. & ECCLESTONE, N. A. 2002. Contribution of Structured Exercise Class Participation and Informal Walking for Exercise to Daily Physical Activity in Community-Dwelling Older Adults. *Research Quarterly for Exercise and Sport,* 73**,** 350-356.

TURNER, C. H. 1998. Three rules for bone adaptation to mechanical stimuli. *Bone,* 23**,** 399-407.

TURNER, C. H. & ROBLING, A. G. 2003. Designing Exercise Regimens to Increase Bone Strength. *Exercise and Sport Sciences Reviews,* 31.

VAN DEN BERG, M. H., DE BOER, I. G., LE CESSIE, S., BREEDVELD, F. C. & VLIET VLIELAND, T. P. 2007. Are patients with rheumatoid arthritis less physically active than the general population? *J Clin Rheumatol,* 13**,** 181-6.

VERHOEVEN, F., TORDI, N., PRATI, C., DEMOUGEOT, C., MOUGIN, F. & WENDLING, D. 2016. Physical activity in patients with rheumatoid arthritis. *Joint Bone Spine,* 83**,** 265-270.

VERONICA HACKETHAL, M. 2018. *6 Top Rheumatology Apps for Patients* [Online]. Available: http://www.rheumatologynetwork.com/psoriatic-arthritis/6-top-rheumatology-apps-patients [Accessed].

VERVLOESEM, N., VAN GILS, N., OVAERE, L., WESTHOVENS, R. & VAN ASSCHE, D. 2012. Are personal characteristics associated with exercise participation in patients with rheumatoid arthritis? A cross-sectional explorative survey. *Musculoskeletal Care,* 10**,** 90-100.

VIJAYAN, V., CONNOLLY, J. P., CONDELL, J., MCKELVEY, N. & GARDINER, P. 2021. Review of Wearable Devices and Data Collection Considerations for Connected Health. *Sensors,* 21**,** 5589.

WAGNER, E. B., GRANZELLA, N. P., SAITO, H., NEWMAN, D. J., YOUNG, L. R. & BOUXSEIN, M. L. 2010. Partial weight suspension: a novel murine model for investigating adaptation to reduced musculoskeletal loading. *J Appl Physiol (1985),* 109**,** 350-7.

WAHL, Y., DÜKING, P., DROSZEZ, A., WAHL, P. & MESTER, J. 2017. Criterion-Validity of Commercially Available Physical Activity Tracker to Estimate Step Count, Covered Distance and Energy Expenditure during Sports Conditions. *Frontiers in Physiology,* 8.

WALLIS, J. A., WEBSTER, K. E., LEVINGER, P. & TAYLOR, N. F. 2013. What proportion of people with hip and knee osteoarthritis meet physical activity guidelines? A systematic review and meta-analysis. *Osteoarthritis and Cartilage,* 21**,** 1648-1659.

WANG, M., DONOVAN-HALL, M., HAYWARD, H. & ADAMS, J. 2014. People's Perceptions and Beliefs about their Ability to Exercise with Rheumatoid Arthritis: A Qualitative Study. *Musculoskeletal Care*.

WANG, X., PERRY, T. A., CAROUPAPOULLÉ, J., FORRESTER, A., ARDEN, N. K. & HUNTER, D. J. 2021. Monitoring work-related physical activity and estimating lower-limb loading: a proof-of-concept study. *BMC Musculoskeletal Disorders,* 22.

WANMIN WU, S. D., ERNESTO E RAMIREZ, CARLYN PETERSON, GREGORY J NORMAN 2012. Classification Accuracies of Physical Activities Using Smartphone Motion Sensors. *Journal of Medical Internet Research*.

WHITE, T., WESTGATE, K., WAREHAM, N. & BRAGE, S. 2016. Estimation of Physical Activity Energy Expenditure during Free-Living from Wrist Accelerometry in UK Adults. *PLOS ONE,* 11**,** e0167472.

WOOLLASTON, V. 2014. How often do YOU look at your phone? *Dailymail.co.uk* [Online]. Available from: http://www.dailymail.co.uk/sciencetech/article-2783677/How-YOU-look-phone-The-average-user-picks-device-1-500-times-day.html.

YU, S. P., FERREIRA, M. L., DUONG, V., CAROUPAPOULLÉ, J., ARDEN, N. K., BENNELL, K. L. & HUNTER, D. J. 2022. Responsiveness of an activity tracker as a measurement tool in a knee osteoarthritis clinical trial (ACTIVe-OA study). *Ann Phys Rehabil Med,* 65**,** 101619.

ZHANG, S., ROWLANDS, A. V., MURRAY, P. & HURST, T. L. 2012. Physical activity classification using the GENEA wrist-worn accelerometer. *Med Sci Sports Exerc,* 44**,** 742-8.

ZHONG, B. Q., MIN, Y., HAITAO, F., NING, M., TINGTING, W., QINGQIN, T., ZHIDE, H., YAN, L., ZAIXING, Y. & RENQIAN 2015. Body mass index and the risk of rheumatoid arthritis: a systematic review and dose-response meta-analysis. *Arthritis Research & Therapy,* 17**,** 86.