The Use of Gaming Technology in the Rehabilitation of Patients Following Total Knee Replacement Surgery

Jonathan James Negus

SID : 309303990
MA(Cantab.) MSc MBBS MRCS

A dissertation submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Faculty of Medicine
Northern Clinical School
Institute of Bone and Joint Research
University of Sydney, Australia

and

Sydney Orthopaedic Research Institute

October 2015

ORIGINALITY STATEMENT

I hereby declare that this submission is my own work and to the best of my knowledge it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at the University of Sydney or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at the University of Sydney or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of the thesis is the product of my own work, except to the extent that the assistance from others in the project's design and conception or in style, presentation, and linguistic expression is acknowledged.

Jonathan James Negus

November 2014

ACKNOWLEDGEMENTS

PUBLICATIONS RESULTING FROM THIS THESIS

Published

Negus, J.J., Cawthorne, D., Chen, J.S., Scholes, C.J., Parker, D.A., March, L.M. (2014) Patient Outcomes using Wii Enhanced Rehabilitation after Total Knee Replacement – The TKR-POWER study. (A study protocol) Contemporary Clinical Trials 40 (2015) 47-53.

Under Peer Review

Negus, J.J., Cawthorne, D., Clark, R., March, L.M., Parker, D.A. (2014) Validity and reliability of the Nintendo Wii Fit Stillness score for assessment of standing balance in healthy adults. (Submitted to Gait and Posture.)

Invited International Presentations

Negus, J.J. (2014) TKR-POWER. Asia Pacific Orthopaedic Association – Biennial Meeting. Pattaya, Thailand - Australian Young Ambassador

Negus, J.J. (2014) TKR-POWER – 1 year results. Australian Orthopaedic Association - Annual Scientific Meeting. Melbourne, VIC - AORA prize winner

Presentations and Posters

Negus, J.J. (2013) PODIUM. TKR-POWER Study. Australian Orthopaedic Association (AOA)– Annual Scientific Meeting. Darwin, NT

Negus, J.J. (2013) PODIUM. TKR-POWER Study. Australian Orthopaedic Registrars Association (AORA) – Annual Scientific Meeting. Darwin, NT

Negus, J.J. (2011) POSTER. Is the Wii Fit for purpose? British Orthopaedic Research Society – Cambridge, UK

Negus, J.J. (2011) PODIUM. Is the Wii Fit for purpose? Validating Nintendo's chosen measure of balance for clinical use. Australian & New Zealand Orthopaedic Research Society. Brisbane, QLD

Negus, J.J. (2011) PODIUM. Is the Wii Fit for purpose? Validating Nintendo's chosen measure of balance for clinical use. Australian Orthopaedic Registrars Association – Annual Scientific Meeting. Mt Buller, VIC

Awards and Scholarships

Royal Australasian College of Surgeons foundation for surgery - PhD scholarship 2010

Australian Orthopaedic Registrars Association – Registrar Prize 2011

Australian Orthopaedic Registrars Association - Allan Frederick Dwyer Prize 2013

Australian Young Ambassador - Asia Pacific Orthopaedic Association meeting, Pattaya, Thailand 2014

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ABS – Absolute value

ACL – Anterior cruciate ligament

ACORN - Arthroplasty Clinical Outcomes Registry

ADLs –Activities of daily living

AKSCRS - American Knee Society Clinical Rating Score

AKSS - American Knee Society Score

ANOVA – Analysis of variance

ANCOVA – Analysis of Covariance

AOANJRR - Australian Orthopaedic Association National Joint Replacement Registry

AP – Anteroposterior

BLISS _ Bellamy Low Intensity Symptom State-attainment

CAS – Computer Attitude Scale

cm – Centimetres

CoG - Centre of gravity

CONSORT – Consolidated Standards of Reporting Trials

CoP – Centre of pressure

DSEC - Double-leg Stance with Eyes Closed and feet together

DSEO - Double-leg Stance with Eyes Open and feet apart

Ft – Feet

GP – General Practitioner

GUG – Get Up-and-go

HADS – Hospital Anxiety and Depression Score

Hz – hertz

ICC – Intraclass correlation coefficient

IMMPACT - Initiative on Methods, Measurement and Pain Assessment in Clinical Trials

JPS – Joint Position Sense

Kg – Kilograms

KOS – Knee Outcome Survey

LoC – Locus of control

M – Metres

Max – Maximum

MCAS – Maximal Clinically Acceptable State

MCID – Minimum Clinically Important Difference

MCII – Minimum Clinically Important Improvement

ML – Mediolateral

mm – Millimetres

NHP - Nottingham Health Profile

NSAIDs – Non-Steroidal Anti-Inflammatory Drugs

NSW – New South Wales

NU – Normalised units

OA - Osteoarthritis

OARSI - Osteoarthritis Research Society International

OKS – Oxford Knee Score

OMERACT - Outcome Measures in Rheumatology Clinical Trials

PASS – Patient Acceptable Symptom State

PROMs - Patient Recorded Outcome Measures

RCT – Randomized Controlled Trial

ROM – Range of motion

SCT – Stair-climbing Test

S.D. – Standard deviation

S.E. – Standard error

SEM – Standard error of measurement

SF – Short Form

SIP - Sickness Impact Profile

SORI – Sydney Orthopaedic Research Institute

SPSS – Statistical Package for the Social Sciences

SQRT - Square root

TKR – Total knee replacement

TKR – POWER – Patient Enhanced Outcomes in Wii Enhanced Rehabilitation following Total Knee

Replacement

THR - Total hip replacement

UG – Up and Go

USEC   Unilateral-leg Stance with Eyes Closed

USEO   Unilateral-leg Stance with Eyes Open

VAS – Visual Analogue Scale

WBB – Wii Balance Board

WOMAC – Western Ontario McMaster Universities Index of Osteoarthritis

CHAPTER 1 - Introduction

Total knee replacement (TKR) is a successful operation as measured by functional, objective, clinical outcomes and quality-adjusted life years. It relieves the pain of patients suffering from severe disease of the knee caused by osteoarthritis in over 95% of cases. Although it leads to satisfaction levels of 80% this compares poorly with total hip replacement's satisfaction levels of around 90%. There is therefore considerable clinical and research interest in the gap in satisfaction levels between the two procedures.

Rehabilitation is a large part of the recovery from a TKR operation. The type, amount, and duration of rehabilitation vary from country to country as well as within countries and even between surgeons in the same hospital facility. There has been a lack of evidence to support choice of appropriate rehabilitative therapy for TKR patients and quality evidence has only begun to appear in the last five years.

Adherence to any rehabilitation protocol is a known problem with rates of 50% non-adherence reported. It should be self-evident that not only do we need to know the most effective exercises to prescribe for these patients but we also need to prescribe them in a manner which will encourage ongoing adherence to the protocols.

The Nintendo Wii-Fit (Nintendo of America, Redmond, WA, USA) software with associated balance board, designed to work with the Wii console, was released for sale to the Australian public in 2008. The upgraded Wii-Fit 'plus' software was then released in 2009. It has proved to be a very popular game with over 22 million units sold worldwide up to March 2012. From the very early days, clinicians and researchers saw the potential for its use as a therapeutic tool and began to use it for fitness purposes as well as various rehabilitative uses.

The advantages of the Wii-Fit include: real-time feedback on balance, progress updates, increasing difficulty of exercise, variety, and automatic data recording. These features can benefit both the subject using it and the clinician or researcher directing the rehabilitation period.

To date, there has been no registered, randomized clinical trial using the Wii as a rehabilitation tool in the TKR population. The objectives of this thesis are to evaluate the Wii-Fit as a balance testing tool and a rehabilitation adjunct for patients to use at home.

The specific aims include the following:

- systematically review the literature for current methods used for rehabilitating patients following knee surgery, focusing on TKR
- assess the validity and reliability of the Wii-Fit for measuring standing balance in healthy volunteers
- evaluate the association of balance and function in patients with severe osteoarthritis of the knee
- evaluate the ability of elderly patients to cope with technology such as the Nintendo Wii
- evaluate functional outcomes when the Nintendo Wii is used at home as part of a rehabilitation protocol following TKR surgery
- evaluate adherence to the rehabilitation protocols with the Nintendo Wii compared with a routine protocol

Chapter 2 gives an in-depth background of TKR surgery, rehabilitation, outcome measures and current uses of the Nintendo Wii.

Chapter 3 describes the methodology of the various studies.

Chapter 4 details the pilot studies with subsections on the validation of the Wii balance measure, the association of balance with function in those with knee osteoarthritis, and the use of technology in the elderly.

Chapter 5 details the main randomized trial and the associated study of adherence.

Chapter 6 summarizes the results and discusses their relevance.

CHAPTER 2 - Background And Literature Review

2.1 CONCEPTS IN TOTAL KNEE REPLACEMENT

2.1.1 THE SURGERY

Total knee replacement (TKR) surgery is performed for patients with end-stage disease which is

causing them severe pain or functional disability. There should be radiographic evidence of significant

arthritis and the patient should have failed non-operative measures such as regular analgesia, the use

of walking aids, lifestyle modification, weight loss, and exercise. [1]

The number of people who have a TKR is increasing year on year in the developed world. In Australia,

the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) reported

42,202 surgeries in 2012. This is an increase of 4.1% from 2011 and 92.4% from 2003 (see Figure 1).

Figure SK3: Trends in Usage of Knee Replacement by Procedure Year
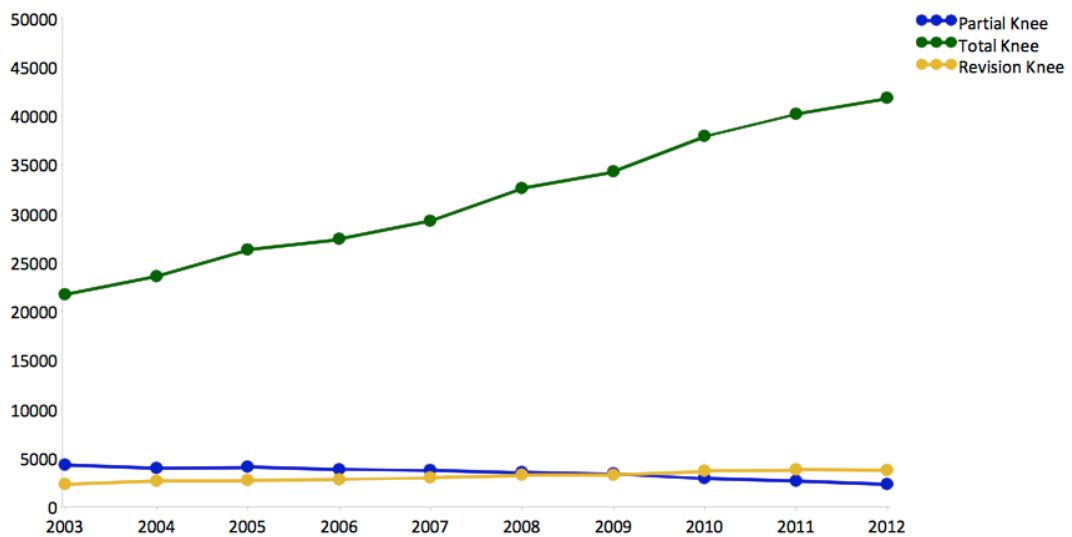
Figure 1 – Number of knee replacements in Australia 2003–2012 [2]

The prosthesis is a combination of a metallic femur resurfacing made of titanium or cobalt-chrome and a metallic tibial tray with a polyethylene insert. The biomechanics of the replacement do not match the native knee. The implants have been developed over the years to mimic some of the properties of the native knee, such as femoral rollback. This is aimed at improving function, as measured by metrics such as the degree of flexion. There are hundreds of prostheses available worldwide although there are only 34 TKR combinations with over 350 procedures and ten-year data in the Australian Joint Registry. The cumulative percentage revision at 12 years for primary TKR undertaken for osteoarthritis is 6.5%.

The surgery routinely takes between one and two hours. Patients remain as hospital inpatients for three to five days before being discharged. They can be sent home or released into the care of a rehabilitation specialist as either an inpatient or an outpatient. The patient will have ongoing postoperative pain which can be severe in the first week for between 6 and 12 weeks. The knee remains swollen for weeks and can feel warmer than the other knee for months.

2.1.2 PATIENT SELECTION

Osteoarthritis (OA) is the most common diagnosis (97.4%) in the AOANJRR for patients who have TKR surgery.[2] This is followed by rheumatoid arthritis, other inflammatory arthropathies, and osteonecrosis (see Table 1). Of those with OA, approximately 80% have medial compartment disease, with the loss of medial cartilage and even bone, giving them a varus or 'bow legged' deformity. Five to ten per cent develop lateral compartment OA resulting in a valgus or 'knock-kneed' deformity. Although patellofemoral joint OA is common in association with medial or lateral compartment OA, a small percentage of patients have isolated patellofemoral joint OA.

Primary Diagnosis	Number	Percent
Osteoarthritis	342574	97.4
Rheumatoid Arthritis	5648	1.6
Other Inflammatory Arthritis	1715	0.5
Osteonecrosis	1181	0.3
Tumour	439	0.1
Fracture	184	0.1
Chondrocalcinosis	17	0.0
Osteochondritis Dissecans	4	0.0
Other	113	0.0
TOTAL	351875	100.0

Table 1 – Primary diagnosis of primary TKR. [2]

The majority of osteoarthritis of the knee is primary or idiopathic. Secondary OA can be caused by

trauma, congenital deformities (such as axial or rotatory deformities), or rheumatoid arthritis. The

established risk factors for developing osteoarthritis are age, genetic predisposition, obesity, previous

knee injury or surgery, osteoarthritis at other sites, and female gender. More controversial risk factors

include physical activity, smoking, oestrogen and vitamin D deficiency. [3-6]

The demographics of TKR patients in the 2013 AOANJRR indicate that 56% are female and 40% of are

aged between 65 and 74. In recent years, there has been a trend towards younger patients being

operated on with the proportion of 55-64-year-olds (28%) overtaking the 75-84-year-old group (22%)

for the first time in the registry. Fewer than 5% are over 85 and 7% are younger than 55 (See Figure

2).
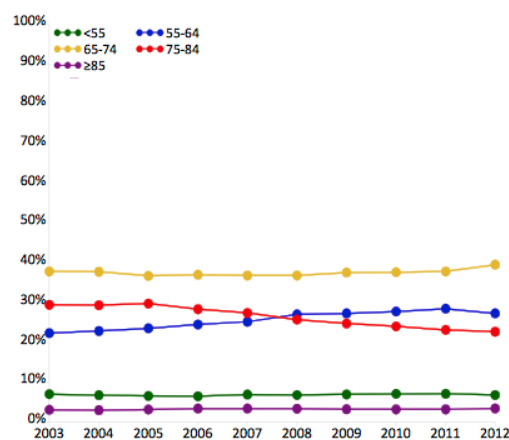
Figure 2 – Primary total knee replacement by age in Australia, 2003–2012. [2]

2.1.3 THE UNSATISFIED 20%

It is widely reported that whereas approximately 90% of patients who have total hip replacement surgery are satisfied, the satisfaction rate for total knee replacement is only 80%.[7-11] In addition, more total hip replacement (THR) patients would choose to have the surgery again, given the choice.[12, 13] Greater deficits in pain and function exist after TKR compared with THR.[14] These consistent findings have been a source of much conjecture and researchers have tried to establish the causes and provide a remedy.

The perceived success of TKR surgery relies on a complex mix of patient, surgical, and rehabilitation factors. Success is also defined by the outcome which is being measured. Surgeons have always been interested in the longevity of implants, as measured by revision rates. Initial scoring systems focused on objective measures such as alignment and range of motion. This made sense, as these outcomes are easy to measure and quantify. The concern with objective outcomes measured by the surgeon is that they may not reflect the level of patient function or satisfaction. It would be difficult to class a TKR surgery as a success because it is well aligned, has a good range of motion and is not wearing out if the patient is in pain and totally dissatisfied.

In recent years, researchers have focused on this issue by measuring patient reported outcome measures (PROMS). These outcomes are subjective and are administered in the form of questionnaires, completed by the patient. Surgeons or their research staff usually collect these scores from the patients. The Joint Registry of England and Wales now includes PROMS as part of its follow-up of TKRs. This is in conjunction with the online data collection program, 'My Clinical Outcomes', which will be described later (www.myclinicaloutcomes.co.uk). In Australia, the Arthroplasty Clinical Outcomes Registry (ACORN) are now collecting PROMS data, which are being analysed alongside the revision data of the AOANJRR. They have just published their first annual report for 2013.[15]

Revision is seen as a 'hard' end-point, making it relatively easy to measure. However, TKRs should

only be revised for a specific and reversible cause. Without this, the knee may not be revised even though the patient is unhappy. The patient with a revised knee may well be more satisfied after their revision, but the registry does not give us that information. The PROMS data give clinicians an opportunity to quantitatively assess a patient's perceived level of function and satisfaction and compare their patient population with the national average.

The approach to the unsatisfied 20% has to be multimodal, with continued research into preoperative expectations, perioperative management including analgesia, surgical technique, the kinematics of prosthesis design, and, importantly, postoperative rehabilitation.

2.1.4 THE NINTENDO WII

The Nintendo Wii[TM] with Wii-Fit[TM] software is a popular, motion-controlled gaming system used across a broad age range and available to the general public. The Wii-Fit software uses the Wii balance board as a control device via adjustments in the participants' balance and proportions of body weight through each leg. It is easy to use and can deliver a varied exercise routine, with a focus on balance.

The Wii has been used to rehabilitate patients with neurological injuries [16-19] but there have been few quality studies describing the use of the Wii to rehabilitate orthopaedic patients. [20, 21] A feasibility study using the Wii as an adjunct to physiotherapy concluded that it has potential for patients following TKR. The study found no difference between the groups for a range of outcome measures including pain rating, subjective balance reporting, patient reported functional outcomes and length of outpatient rehabilitation. Their post hoc power analyses calculated that 'the sample size of 50 patients did not provide sufficient power for the majority of the outcomes assessed.' They calculated that a sample size of 136 in two evenly sized groups would be powered appropriately for one of their main outcome measures, the lower extremity functional scale. The study used the Wii in the physiotherapy rooms during a supervised session, not as a home-based device. [21] Although published

23

after the completion of the Patient Outcomes using Wii Enhanced Rehabilitation after TKR trial (TKR-POWER) this paper by Fung et al was useful in justifying certain decisions made in the design of that study (see chapter 5.1).

2.2 CONCEPTS OF TKR REHABILITATION

2.2.1 INTRODUCTION TO REHABILITATION

The World Health Organization's definition of rehabilitation is a progressive, dynamic, goal-oriented, and often time-limited process, which enables an individual with impairment to identify and reach his/her optimal mental, physical, cognitive, and/or social functional level. [22] The aim is to maximize the functional capacity of the patient with a given pathological condition by taking into account their physical, psychosocial, and social factors.

The knee is a joint which allows us to stand upright with a straight leg, walk, and ascend stairs, but also sit in a chair. In order to do this it needs to be well aligned in relation to the mechanical axis of the leg, powered by surrounding muscles with appropriate levels of proprioception while giving the patient little or no pain.

With a mechanical replacement in place, it is vital that the patient rehabilitates, allowing the soft tissues to recover from the surgery while maintaining and even improving the strength and coordination of the supporting musculature. This can be even more challenging for the TKR population because of their advanced years.

TKR surgery is a highly cost-effective treatment for severe symptomatic osteoarthritis[23-26] but the costs and cost-effectiveness of rehabilitation are less extensively studied.[27] There is also a lack of evidence-based clinical guidelines for physiotherapy rehabilitation after TKR surgery. With annual increases in the number of TKR surgeries performed, the costs and effectiveness of different modes of

rehabilitation need to be investigated to maximize functional recovery at the best cost to the

healthcare system.

2.2.2   METHODS OF REHABILITATION

2.2.2A PREOPERATIVE PHASE

Rehabilitation after a TKR involves a number of stages. The first stage starts even before the

operation with patient education. It can include options for optimizing health and strength, what to

expect from the hospital stay, and what type of rehabilitation is planned. It is important to make

patients aware that the success of their knee surgery is heavily influenced by their efforts in the

recovery phase. The knee will be painful and difficult to move; without significant efforts on their

part, they may be left worse off than before the operation. This education phase should start in the

surgeon's rooms when the case is booked and continuing education sessions are often run by the

treating hospital.

2.2.2B PERIOPERATIVE AND EARLY REHABILITATION PHASE (WEEK 0 TO WEEK 6)

A 2011 study found the mean length of stay following TKR surgery to be 5.6 days (s.d. 2.1). [28] This

figure can vary significantly according to the hospital where the surgery is performed According to

www.myhospitals.gov.au the average length of stay at Royal North Shore public hospital in Sydney

following TKR surgery was 7.1 days. [29]

The goals of rehabilitation during this in-patient period are to prevent the hazards of bed-rest (e.g.

deep vein thrombosis, pressure ulcers) and to control pain. The focus is then on assisting the patients

to gain an adequate and functional range of motion including strengthening of knee musculature.

The physical therapists concentrate on these goals with a focus on gaining functional independence with their activities of daily living including independent ambulation to enable discharge. A functional range of motion is a key requirement following TKR surgery. A paper by Laubenthal et al in 1972 is most commonly quoted in the literature to demonstrate that patients require 67 degrees during the gait balance phase, 83 degrees for stepping up stairs, 93 degrees to rise from a chair, 100 degrees for stepping down and at least 105 degrees in order to be able to rise from a toilet seat of normal height.[30] If the patient has bilateral disease or replacements then one of the knees needs to flex to at least 105 degrees to allow this basic level of function. It is worth considering the number of patients who will undergoing simultaneous or staged bilateral TKR surgery necessitating a functional range in at least one of the replaced knees. The AOA National Joint Replacement Registry for 2014 reports on 85,473 bilateral TKR procedures out of a total of 396,472 primary TKRs. Of these, 18,146 were simultaneous bilateral surgery (18.6% of all bilateral TKR surgery in the registry). 9,731 patients (10% of patients who have had both knees replaced with primary TKRs) had the second surgery within 6 months of the first and a further 36,081 between 6 months and 3 years after the first. This means that 11.6% of all patients entered into the Australian National Joint Replacement Registry for a primary TKR on one side had a primary TKR to the other side within 3 years.[31]

When investigating methods of improving rehabilitation, it has been common practice to take exercises and protocols which have been used in age-matched subjects with or without arthritis and transpose these into a post-surgical TKR population. The theory is that if declining function or arthritic symptoms are treated by these exercises, then they may have similar success in the postoperative TKR population. This would seem a reasonable starting-point and, indeed, studies have shown beneficial effects from these exercises, especially in terms of reducing pain and improving function.[32-35] However, a prosthetic knee is very different from a native knee. The surgery involves the release and excision of ligaments which affect the proprioception of the joint and can realign a limb which has been functioning out of alignment for many years.[36]   This is one reason why improving the balance and proprioception of the operated leg is likely to be an important focus for rehabilitation.

The type of exercise which should be prescribed is not clear. A recent evaluation of RCTs and systematic reviews found only one relevant systematic review[37] that addressed the issue TKR rehabilitation. Their findings regarding the type of exercises that should be included in a rehabilitation program were indeterminate. [38] Traditional exercise programmes for patients following a TKR concentrate on regaining a range of motion, strengthening musculature, and building endurance into those muscles. Studies have also tested the effectiveness of functional exercise versus traditional and demonstrated some good results. However, these are often short-lived, not lasting beyond the 3-4-month postoperative point.[32, 37, 39]

2.2.2c POST-ACUTE REHABILITATION (WEEK 6 ONWARDS)

The timing of rehabilitation is also under review in the literature.[37, 40, 41] The traditional approach has been to concentrate efforts into the first six weeks and then discharge the patient from formal rehabilitation. Current research is investigating the idea that delaying the more intense phases of a rehabilitation protocol to start at the six-week point may allow a patient the chance to address their swelling and pain.[40] Once the knee is less acute, the intense muscle building and functional exercises may have more effect, as the patient will be more comfortable. This also has significant implications for the cost of current practice, especially for those patients prescribed physiotherapy either as an inpatient or an outpatient.

2.2.3 LOCATION AND DURATION OF REHABILITATION

The location and duration of rehabilitation in Australia is driven by a number of factors. Rehabilitation team protocols, hospital policy, insurance status, and patient preference all play their part. However, surgeon preference can be a significant factor. There are a variety of opinions on the role of rehabilitation following joint replacement surgery ranging from dogmatic protocols to individually tailored programmes.

The location of post-acute rehabilitation varies between and even within countries. Lingard et al showed that in Australia and the United Kingdom, this stage tends to be conducted via outpatient-based services, in contrast to the inpatient rehabilitation prescribed in the USA.[42] This has changed over time with the USA now aiming for a reduced hospital stay and more discharges to home driven largely by cost. In terms of Australian management, this study looked at one public and one private hospital so is not representative. The majority of TKR procedures in Australia are performed in the private sector, over 30,000 a year, and 46% of these in New South Wales and 32% nationally received inpatient rehabilitation in 2011-12.[2] Even outpatient services vary widely between one-to-one physiotherapy and a more clinic-style approach. Each of these rehabilitation options has a cost implication, and so knowledge of the effectiveness of each is essential, especially in a public health system, if the cost-effectiveness is to be calculated.[43-46] It is also important to know the results of not rehabilitating the patients adequately from a clinical and financial perspective.

As part of this thesis, an online questionnaire was sent to the members of the Australian Arthroplasty Society and the Australian Knee Society in February 2012, with respect to their views on rehabilitation following TKR surgery. The questionnaire assessed the number of arthroplasty surgeries which each respondent performed each year to establish the nature of their orthopaedic practice. The questions related to what their current practice involved in the acute, post-acute and long-term phases of rehabilitation, and what level of importance they placed on each of these phases. Data were also collected on where this rehabilitation took place in their practice and how important the location was. See questions below:

1. What is your area of speciality? - 17 responses

    a. Knees only   9 (53%)

    b. Hips and knees   5 (29%)

    c. Lower limb   2 (12%)

    d. Hip only   1 (6%)

2. How many TKRs do you perform a year? – 17 responses

    a. Unilateral TKR

        i. 0   1 (6%)

        ii. 1-50   1 (6%)

        iii. 51-100   3 (18%)

        iv. 101-250   10 (58%)

        v. 251+   2 (12%)

    b. Simultaneous Bilateral TKR

        i. 0   3 (18%)

        ii. 1-50   11 (64%)

        iii. 51-100   3 (18%)

        iv. 101-250   0 (0%)

        v. 251+   0 (0%)

3. What is your preferred choice of post operative destination for the majority of your TKR patients?

    a. Home   7 (41%)

    b. Home with out patient physiotherapy   6 (35%)

    c. Inpatient rehabilitation   4 (24%)

4. How long do you routinely send your patients to inpatient rehabilitation? – 4 responses

    a. <1 week   0 (0%)

    b. 1-2 weeks   4 (100%)

    c. > 2 weeks   0 (0%)

    d. No preference   0 (0%)

5. How important do you think post operative inpatient rehabilitation is to the outcome for your patients? – 17 responses

    a. Essential   5 (30%)

    b. Important   6 (35%)

    c. Not important   6 (35%)

d. No view   0 (0%)

6. How long do you routinely advise your patients attend out patient rehabilitation?

  a. <6 weeks   3 (18%)

  b. 6-12 weeks   9 (52%)

  c. > 12 weeks   4 (24%)

  d. no view   1 (6%)

7. How important do you think post operative rehabilitation is for the outcome for your patients?

  a. 0-2 weeks

    i. Essential   7 (41%)

    ii. Important   9 (53%)

    iii. Not important   1 (6%)

    iv. No view   0 (0%)

  b. 2-8 weeks

    i. Essential   8 (47%)

    ii. Important   7 (41%)

    iii. Not important   2 (12%)

    iv. No view   0 (0%)

  c. 8 weeks onwards

    i. Essential   4 (24%)

    ii. Important   6 (35%)

    iii. Not important   7 (41%)

    iv. No view   0 (0%)

8. How important do you think post operative home exercise is to the outcome for your patients?

  a. 0-2 weeks

    i. Essential   9 (53%)

	ii.	Important	6 (35%)
	iii.	Not important	1 (6%)
	iv.	No view	1 (6%)

b. 2-8 weeks

	i.	Essential	9 (53%)
	ii.	Important	8 (47%)
	iii.	Not important	0 (0%)
	iv.	No view	0 (0%)

c. 8 weeks onwards

	i.	Essential	8 (47%)
	ii.	Important	7 (41%)
	iii.	Not important	2 (12%)
	iv.	No view	0 (0%)

9. Do you feel the current TKR rehabilitation practice is ideal? – 16 responses

a.	YES	7 (44%)
b.	NO	9 (56%)

The survey invitation was sent by email to the members by the secretaries of each society. The survey was designed and conducted on 'Survey Monkey'. From the 78 questionnaires sent out, there were 17 respondents (22%), of whom 9 were purely knee surgeons and 8 performed more than 100 TKRs a year. One respondent was a hip surgeon who performed no TKR surgery.

The results demonstrated that these knee and lower limb surgeons had no single preference or protocol for how they rehabilitated their patients. While 41% preferred to discharge their TKR patients straight home from hospital with no formal rehabilitation, the remainder chose some form of postoperative rehabilitation. This was consistent with how important they felt postoperative inpatient rehabilitation to be, which ranged from essential through important to not important in three equal

groups. The majority recommended that the duration of outpatient rehabilitation should be 6-12 weeks.

Only those who indicated that they sent their patients to in-patient rehabilitation were then questioned on the length of that rehabilitation (n=4). The 41% (n=7) of surgeons who indicated that their preferred post operative destination was home, did still respond to the remaining questions.

More surgeons felt that the first two weeks of outpatient rehabilitation were essential (41%) compared with the period from eight weeks onwards (24%). Equally, only 6% felt the first two weeks were not important although they attached less importance to phases which followed (two to eight weeks 12%, eight to twelve weeks 41%).

Around half the surgeons (47%) thought that home exercise was essential from discharge to beyond eight weeks. Only 6% felt that these exercises were not important. Beyond 12 weeks, 89% thought that home exercises were essential or important.

To the question, 'Do you feel that current TKR rehabilitation practice is ideal?', 56% answered 'No', compared with only 25% for total hip replacement rehabilitation.

This survey was limited by a low response rate but did give an indication of how some high volume knee surgeons approached post-operative rehabilitation for TKR surgery. On questioning several of the members, this survey came at a time of particularly high survey traffic and there was a sense of fatigue for online questionnaires.

In summary, there is a disparate view as to where patients should be discharged for their rehabilitation with no consensus on the importance of inpatient rehabilitation. This is mirrored in the literature (see chapter 2 section 2.3.2.3)., The majority of respondents felt out patient rehabilitation to be important or essential up to and even beyond 8 weeks (59%) and home exercise to be essential

(89%) even beyond 8 weeks. This shows both that there are knee surgeons who value and promote rehabilitation for their patients but that there are also surgeons who feel it has little role. The numbers involved do not allow for further conclusions to be drawn about the relative size of these groups in Australia.

2.2.4 CURRENT PROTOCOLS

The various approaches to postoperative TKR rehabilitation have led to significant variation in the protocols used in different practices. Most institutions have a protocol that can be found with a quick internet search, and they are broadly similar. There is limited evidence for protocols as a whole. The evidence base for TKR physiotherapy rehabilitation has tended to concentrate on specific questions such as timing of interventions and specific types of interventions such as proprioceptive training.[37, 39]

Most protocols target swelling, range of motion (ROM), strength, patella mobilization techniques as well as balance and proprioception to some degree. Current practice in Australia involves many different protocols.[47] It would seem opportune to examine the evidence, to assess the optimal start time, duration, type of exercise, and the location of postoperative TKR rehabilitation. It may be that we do not even have the appropriate evidence yet, leaving room for further research into this critical area of patient management.

The final report of the National Institute of Health consensus meeting in 2003 on TKRs concluded,

> 'there is a notable lack of consensus regarding which medical and rehabilitative peri-operative practices should be employed, mostly because of the lack of well-designed studies testing the efficacy and effectiveness of such practices… The use of rehabilitation services is perhaps the most understudied aspect of the peri-operative management of TKR patients' and that 'no evidence-based guidelines exist for promoting or limiting post-TKR physical activity.'[48]

This has yet to be updated, even though it is over 10 years ago.

A Cochrane review in 2008 looking at multidisciplinary rehabilitation programmes following joint replacement at the hip and knee in chronic arthropathy failed to find the evidence to suggest an

optimum 'dose' of therapy.[49] Their recommendation was that the evidence provided only modest support for TKR patients being assessed for their need for an appropriate rehabilitation intervention.

Westby et al synthesized a consensus statement in 2014 across two expert panels of clinicians, researchers and patients from Canada and the USA.[50] There were increasing levels of agreement for two thirds of the key statements for TKR rehabilitation. However, there was a lack of consensus with widely ranging opinions from strongly agree to strongly disagree, regarding the level of direct supervision needed. There were also disparate views on the subjects of timing, dose, format and specific interventions. This was partly due to strongly entrenched local practice regardless of the evidence supporting cost savings associated with group programs versus individual supervision for example. They state in their discussion that 'the literature contributes little to our understanding of optimal timing,' and 'lack of clear consensus in the present study is consistent with the literature, indicating optimal exercise dose and progression is unknown and there remains a need to identify the "active ingredients" of effective rehabilitation interventions for both THR and TKR.'

2.3 LITERATURE REVIEW ON THE REHABILITATION OF PATIENTS AFTER TKR FOR KNEE OSTEOARTHRITIS AND THE OUTCOME MEASURES USED

A literature review evaluated the current state of patient rehabilitation following a total knee replacement. The objectives were to:

- review outcome measures that have been utilized in the evaluation of rehabilitation post TKR

- evaluate current types of rehabilitation programs including the location and duration of rehabilitation

- assess the extent to which novel technology such as the Nintendo Wii has been used in rehabilitation to date

2.3.1 METHODS

A literature search was performed in Medline, EMbase, Cochrane Musculoskeletal Group Trials Registry and Cochrane Central Register of Controlled Trials up to October 2014. Inclusion criteria were articles reporting all methods of rehabilitation following TKR surgery, articles reporting any rehabilitation using the Nintendo Wii and articles relating to relevant outcome measures. The search was limited to humans and publications in English.

The strategy shown in Appendix i was for the Medline search and demonstrates the medical subject headings (MESH) and keywords used. The bibliographies of the relevant articles were searched for further articles. The results are set out according to:

- subjective outcome measures

- objective outcome measures

- location, timing, and type of rehabilitation including the Wii

2.3.2   Results

2.3.2.1   SUBJECTIVE OUTCOMES

Patient-reported outcome measures (PROMS) are mostly questionnaire-based subjective outcome measures. They are commonly expressed as a score. The questions can be divided into categories, such as functional ability, perceived pain, or emotions. There are many outcome scores which have been validated for use in various diseases and operations. PROMS give clinicians data on the outcome of an intervention from a patient's perspective. This is as valuable as objective outcomes which often provide clinicians with data which are of clinical relevance but not important to the patient. PROMS are quick to complete and the data can be collated and analysed easily. They have been adopted by individual hospitals and clinicians for many years but their use in health systems is still uncommon and largely restricted to England, Sweden and the US. [51]

They suffer a number of disadvantages, however. There is a trade-off between providing enough detail to classify all patients accurately and too many questions which are time-consuming for patient and clinician, limiting completion rates.[52] This can lead to important areas relevant to the patient such as driving, interpersonal relationships and community life.[53]

Subjective questionnaires may be prone to bias. [54] A patient may put down a more favourable score either because of coercion (intended or otherwise) by a researcher or surgeon or from a sense of loyalty to their clinician. The score can be influenced by other factors such as psychological states like depression or a preoperative expectation of outcome.

The same absolute value of outcome score for a particular test often represents a different result for different patients. This can be due to levels of preoperative function, expectation levels or pain tolerance. Change scores represent improvement on an individual level, resulting in a more meaningful clinical result.

A selection of relevant studies and systematic reviews were evaluated In order to focus this literature review towards particular outcome measures. In the systematic review by Minns-Lowe and colleagues in 2007, there were only six different trials identified which met the inclusion criteria which looked at effectiveness of physiotherapy exercise after TKR. [37] All the selected studies were asking the same clinical question but they chose to measure a wide variety of different outcomes. This highlights both the range of options in terms of outcome measures for this subject and the lack of a consensus on the most appropriate outcomes to select. The trials in this review included questionnaire-based scores such as the Oxford Knee Score (OKS), the American Knee Society Score (AKSS), the Western Ontario and McMaster Universities (WOMAC) Index of Osteoarthritis, the Bartlett patellar score, and the SF-36, and SF-12 scores.

Dunbar and colleagues looked at the appropriateness of a number of outcome scores for knee arthroplasty. They selected 3,600 random patients from the Swedish registry with a diagnosis of primary OA.[52] They split these patients into 12 groups of 300, each being sent a combination of one general health questionnaire (Nottingham Health Profile (NHP), SF-12, SF-36, Sickness Impact Profile (SIP)) and one disease/site-specific questionnaire (Algofunctional Index for the Knee (Lequesne), Oxford Knee Score (OKS), and WOMAC).

The mean age of the patients at surgery was 71.0 (range = 55-90 years). The mean follow-up time was 6.7 years (range = 1-23 years). Seventy per cent of the sample was female and 95% had not undergone revision surgery. Eighty-five per cent responded to the mail-out within four weeks.

There was a significantly higher response rate for SF-12, SF-36, and NHP than for SIP (p<0.001) .There was no difference in the response rate for the disease/site-specific questionnaires. The SF-12 had the highest percentage of completed questionnaires returned at 75.4% (p<0.001). The SF-36 had a significantly lower efficiency of completion at 63% (p<0.001). Of the disease/site-specific questionnaires, the OKS had the highest percentage of completed questionnaires (85%) followed by

WOMAC (83%), and the Lequesne (79%). The completion efficiencies for these questionnaires were all significantly different (p<0.001).

Of the general health questionnaires, the SIP required the most time to complete (23 minutes) and the SF-12 the least (7.7 minutes). Of the disease/site-specific questionnaires, the WOMAC required the most time (12 minutes), followed by the OKS (10 minutes), and the Lequesne (8 minutes). The differences between all the times for general health and disease/site-specific questionnaires were significant (ANOVA p<0.0001) .

Patients reported a significantly greater frequency of requiring assistance to complete the SF-36 than the others (28.7% p=0.005). The disease/site-specific questionnaires had similar frequencies for requiring assistance.

The SF-12 had no floor effect whereas the SF-36 did (17.1%) but this floor was lower than for the other two general health questionnaires. The WOMAC had the highest floor effect of the disease/site-specific questionnaires (18.3%) with the OKS scoring much lower at 6.8%.
The SF-12 ranked best overall for the general health questionnaires and the Oxford Knee Score ranked best overall for the disease/site-specific questionnaires when the mean individual ranks for each parameter were considered.

The studies and reviews detailed above guided the search strategy and therefore the focus of the following review.

2.3.2.1A   WOMAC

This section reviews the Western Ontario and McMaster (WOMAC) Osteoarthritis Index as a potential primary outcome measure evaluating patients following TKR surgery for osteoarthritis of the knee.

WOMAC is a patient-administered, multi-dimensional, health status questionnaire. It was designed to be a purpose-built, disease-specific, and high-performance instrument for evaluative research in clinical trials looking at osteoarthritis. It was developed within the context of a double-blind, randomized, controlled, parallel trial of two non-steroidal anti-inflammatory drugs. [55]

The index consists of three subscales and a summative total score. It can be scored on a visual analogue scale (VAS), a numeric rating scale or a Likert scale. The VAS is scored from a mark made by the patient, which crosses a 10cm long line. The line is marked at each end by the extremes of the response options: i.e. no pain at one end and extreme pain at the other. The distance along this line to the patient's mark is then measured and turned into a score and each question has a score range of zero to 100. The 24 questions in the WOMAC questionnaire give a total score range of up to 2,400. The Likert scale offers five choices for each question: responses are none, mild, moderate, severe, and extreme. These responses score from zero to four points respectively. The pain subscale has five questions with a score range of zero to 20, the stiffness subscale has two questions with a score range of zero to eight and the function subscale has 17 questions with a score range of zero to 68. In all scales, a score of zero indicates the best functional outcome. The total score has a range of zero to 96. On average, the patients take five to ten minutes to answer the questions.

The response criteria model for expressing data from the OsteoArthritis Research Society International (OARSI) and Outcome Measures in Rheumatology Group (OMERACT) requires the units to be normalized on a scale of zero to 100, 100 being the worst possible outcome. This is to allow comparison with other scores in the literature if they are transformed in a similar manner.

WOMAC is recognized in Osteoarthritis Research Society International (OARSI) Guidelines for clinical trials, by the FDA, European Medicines Agency, Outcome Measures in Rheumatology Clinical Trials (OMERACT),[56] and Initiative on Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT) Guidelines.[57, 58]

The WOMAC index has been amongst the most commonly used patient-reported outcome in the rheumatology literature with Angst et al stating that It "is the most widely used and best proven condition-specific health assessment instrument for patients with OA of the lower extremities to measure condition-specific dimensions as well as overall state of health".[59] While it is a valid, reliable, and responsive measure of patient-reported outcome in osteoarthritis,[60][61] it has been further validated for measuring outcomes following knee and hip arthroplasty.[62] In this validation study, Bellamy and colleagues selected 30 patients with primary OA requiring a total joint replacement (14 for TKR & 16 for THR). Their mean age was 68.3 years (range 54-83 years). They needed to be ambulatory, unrestricted in their functional capacity by any associated condition, and not have undergone prior replacement surgery on the joint under study. Trained examiners assessed them the day before surgery, at six weeks, and then at six months postoperatively. The assessment included the WOMAC OA index and a global assessment. Both scores were completed by the interviewer and the patient. The patients were given two versions of the WOMAC. One was on Likert scales and the other on 10 cm horizontal visual analogue scales. Secondary outcome measures were chosen to test the construct validity of the five WOMAC dimensions. These measures were the modified Doyle index for hip and knee, the Lequesne index, the Bradburn index of well-being, and the social component of the McMaster Health Index Questionnaire. The tertiary outcome measures for the TKR patients were 50 foot walking time and total range of motion. These surgical outcomes were selected to measure the relative efficacy of the final WOMAC question battery.

The pain, stiffness and physical function components of the WOMAC questionnaire all showed statistical significance for responsiveness at six weeks and six months. The reliability of Cronbach's alpha was greater than 0.8 in all cases and on both scales, at all three timepoints. The authors point to a previous publication conferring face and content validity. [63] Validity was demonstrated by correlations between the appropriate WOMAC dimension and the relevant component of the comparative measures. The authors withdrew the social and emotional components of the score because of poor correlation with other indices such as the MHIQ social component. The authors concluded that the WOMAC offered the advantages of superior efficiency over existing indices,

potentially reducing sample size requirements in clinical studies which used the WOMAC as the

primary outcome measure. They also state that it is more patient-relevant than other unidimensional

measures as it probes patient-relevant outcomes.

Dunbar et al found that even though it takes a few more minutes to complete, the response rate is

similar to the Oxford Knee Score (OKS). Although the floor rate appears to be higher, the effect size of

the WOMAC leads to a much more realistic sample size of up to 200 patients for a clinical trial

compared with the OKS requirement of nearer 1,000. [64, 65]

Bellamy and colleagues also developed a survey questionnaire to assess normative values for the

WOMAC index in the general population. They captured demographic information and the function

subscale of the WOMAC index.[66] They removed the attribution statement to OA so that the questions

were relevant to a population without arthritis. The sample size was calculated at 24,000 with a

return rate of 30%. They stratified the subjects into six male and female age groups from 30 to >80

years old. They received 5,560 (23%) questionnaires back. Importantly for the age groups of 70 to 74

and 75 to 79 which represent the mean age of TKR patients, the mean WOMAC function scores were

2.08 (s.d. 2.20), and 2.11 (s.d. 2.26) respectively. The original scores were on a range of zero to 170

and were normalized to a range of zero to10, 10 being the worst function. Therefore, these

normalized scores equate to a score of 14.14 (s.d. 14.96) and 14.35 (s.d. 15.37) on a standard 0-96

scale WOMAC with function scored from zero to 68.

Fortin and colleagues split a cohort of TKR patients into low or high preoperative WOMAC using the

median score as a cut-off. [13] Those patients who scored a lower WOMAC prior to surgery made the

most improvements afterwards but they did not reach the same functional level as those with higher

preoperative scores. This difference remained at two years.[67] A regression analysis found the baseline

physical function subscale of the WOMAC and SF-36 as well as the pain subscale of the WOMAC to be

the most important predictors of the six-month scores. This showed that the single best predictor of

pain and function at six months after THR or TKR is the subject's baseline pain and function.

Furthermore, advanced functional loss because of OA of the knee was associated with worse outcome at six months.

Minimum Clinically Important Difference and related measures

Although the large effect size for patients with knee arthroplasty has been well documented, there is a lack of estimation of smaller effects. It is these smaller effects and their clinical relevance, in which we are interested, when comparing two rehabilitation protocols.[63, 68-70]

The minimum detectable change (MDC) is the smallest increment of change possible on that instrument. It represents the amount needed to be observed before it is considered above the bounds of measurement error for that instrument in that application. This is calculated as 1.96* (SQRT(2))*SEM with SEM calculated as (S.D. of baseline scores) * SQRT (1-ICC test retest).[71] This gives a 95% confidence interval around zero which represents no change.

We can assess the response of individuals to treatment by the degree of improvement in the health state to be measured, by analysis of the proportion of patients who respond with a given degree of improvement or by the proportion of patients who achieve an acceptably low level of undesirable sentinel symptoms or signs. This is also referred to as state attainment. The use of the Minimum Clinically Important Difference (MCID) is an example of this. Other examples of similar measures include minimal perceptible improvement (MPI) and minimum clinically important improvement (MCII). While this review will quote the specific measure for each relevant study, they can all be considered similar to MCID as a measure of clinically relevant change

MCID was first defined by Jaeschke as "the smallest difference in score, in the domain of interest which patients perceive as beneficial and which would mandate in the absence of troublesome side effects an excessive cost, a change in the patient's management".[72] Since then the definition of this concept has varied.

Beaton and colleagues describe a taxonomy of MCID with six categories.[71, 73, 74] The categories include differences between people at one point in time, within-person differences over time, and relative changes within a group compared with a control group.

Angst and colleagues investigated the MCID for WOMAC in patients with osteoarthritis of the knee or hip. [59] They specified MCIDs for both differences of improvement and worsening:

WOMAC

      global = - 0.82 points for improvement and 0.96 for worsening

      function = -0.80 for improving and 1.03 for worsening

      stiffness = -1.01 for improving and 0.29 for worsening

(on a 10-point scale)

They found the MCID for improvement from the baseline to be between 17 (s.d. 39%) and 22% (s.d. 35%) for the subscales and 18% (s.d. 37%) for the total score. As in previous literature, they found a regression to the mean effect, the patients with better preoperative WOMAC scores showing lower improvement than those with worse preoperative WOMAC scores. Those below the median baseline WOMAC score showed improved MCID of 20% (s.d. 44%) and those above the median had an MCID of 17% (s.d. 24%). The regression to the mean was more apparent in the worsening scores, below median patients showing an MCID of 35% (s.d. 66%) and those above scoring 8% (s.d. 31%). They also looked at effect size and found an effect size of 0.39 for pain and 0.37 for function, which compared well with the results of Kazis and colleagues and Deyo and colleagues. [75, 76]

Minimum Clinically Important Improvement (MCII) is another method intended to quantify which changes are clinically relevant. It is defined as the 75th percentile of the distribution of change in WOMAC for patients who considered they had a slightly or moderately important improvement during the study. [77]

Tubach and colleagues investigated MCII in patients with hip or knee OA.[64] The 603 patients with

knee OA had a baseline WOMAC function score (Normalised to a range of 0-100 NU) of 42.8 (s.d.

16.1). The MCII for these patients' WOMAC function scores was: absolute change of 9.1 NU

improvement (95% CI -10.5 to -7.5); and the relative change was 26% improvement (95%CI -28.6% to

-23.3%). When they stratified the patients into low, intermediate, and high baseline scores, they

found that the absolute and relative change scores were higher in those with higher baseline WOMAC

function scores (worse function). This means that those with higher baseline scores needed a larger

change to consider themselves clinically improved. The estimates of MCII did not vary across age,

disease duration tertiles or gender.

Norman and colleagues looked at MCID across 33 published studies and found that MCID was

approximately half of a standard deviation for all but six studies (mean 0.495 S.D. = 0.1555).[78, 79]

Ehrich and colleagues investigated the minimal perceptible difference for WOMAC in patients with OA

of the knee or hip.[80] They treated each of three groups with one of two NSAIDS or a placebo for six

weeks. The MPCI was defined as the difference in the mean change from baseline in WOMAC

(100mm normalized Visual Analogue Scale) between patients who had recorded either no response

or only a mild response to therapy on a Likert scale. MCPI was determined to be 9.7, 9.3 and 10.0 for

WOMAC pain, stiffness, and function subscales respectively. They concluded that mean changes of

between 9 and 12 points on a 100mm normalized scale for WOMAC subscales were perceptible

changes to patients with hip and knee OA.

Response levels are important in analysis. Bellamy and colleagues performed a secondary analysis of

data from a previously conducted randomized controlled trial (RCT).[81] They split subjects into groups

depending on whether their WOMAC scores had reduced by at least 20%, at least 50%, or at least

70%. This method changes the comparison from a between-group continuous data analysis to a

responder criteria-based analysis which can categorize individual patients according to levels of

improvement. This method allows for an assessment of clinical significance. Although it should not be the sole assessment of improvement, it is a helpful tool for adding to an analysis of WOMAC scores.

The Low Intensity Symptom State-attainment (BLISS) Index was proposed by Bellamy and colleagues. [82, 83] They used data collected in a previously reported health outcomes trial. That trial evaluated visco-supplementation in patients with symptomatic knee OA. The pre-defined threshold levels of the WOMAC subscales were ≤5, ≤10, ≤15, ≤20, and ≤25 on a 0-100 normalized unit (NU) scale. They chose 35 NU as the baseline pain intensity requirement for inclusion in the BLISS study. They randomized 255 patients with 127 in the treatment arm and 128 in the control arm. In their analysis, they made many assessments of BLISS scores in different situations such as the number of days required to reach a BLISS score. Importantly, they found that if only a single BLISS threshold had been posited a priori, then the uncorrected analyses indicated that a threshold of ≤15 NU would have resulted in uniformly significant results, irrespective of the WOMAC subscale or specific state-attainment definition used. If all subscales and definitions were included then the Bonferroni-corrected analyses ($p<0.01$) indicated that a threshold of ≤25 NU would be required. The WOMAC stiffness scale appears to be an especially sensitive measure of BLISS states. The analyses based on patients with a BLISS response at any time during the study observed a statistically significant between-group difference in WOMAC stiffness, function, and total score at ≤10 NU consistent with the previously published value for the pain subscale.

Change scores can be defined using absolute or relative differences in outcome scores. The absolute difference is 'final value – baseline value'. This is a measure of the change in value between two timepoints. Using the relative difference is one method of accounting for any differences in the baseline measurement. This can be important as baseline scores can affect outcomes independently of intervention. The relative difference is '(final value – baseline value) / baseline value'.

MCID is an alluring concept but one which has been described as 'elusive as the crock of gold at the end of a rainbow'.[84] For many reasons, the use of a single value which represents an important change may prove impossible to agree upon.

2.3.2.1B   OXFORD KNEE SCORE

This is a patient-administered, disease-, and joint-specific questionnaire developed by Dawson and colleagues in Oxford, UK. [85] It was first published in the JBJS in 1998. They interviewed 20 patients in their outpatients' clinic who were under consideration for a TKR. A questionnaire was drafted from their initial responses. The questionnaire was tested on 20 new patients who were also given a second copy to fill in the next day and return in the post. Modifications were based on additional comments from these patients. This process was repeated for two further sets of 20 patients to produce the final questionnaire.

The final version contains 12 questions, each with five categories of response, scored from one to five. The single final score ranges from 12 (fewest difficulties) to 60 (most difficulties). Unlike the WOMAC questionnaire, it asks about the patient in the last month compared to the WOMAC asking about the current situation.

It was tested prospectively on 117 consecutive patients with an average age of 73 years (46-89), who were booked for a unilateral primary TKR. They were scored preoperatively in clinic and again at six months by mail with a new OKS questionnaire, an SF-36, and a few extra questions. Some 86% had advanced primary OA, 8% had secondary OA to fracture or osteonecrosis, 8% had inflammatory arthritis, and 7% had diagnoses representing a variety of diseases such as gout or Paget's disease.

The authors found the questionnaire had a high test-retest value of 0.92. There was moderate correlation with both components of the American Knee Society Score (AKSS) and parts of the SF-36,

with content particularly related to physical function and pain. In terms of responsiveness, there was a very substantial improvement reported at six-month follow-up. Effect sizes were larger than for any of the individual parts of the SF-36 and the change scores were significantly greater for patients who retrospectively reported the greatest improvement in their condition. It proved sensitive to the improvements obtained by TKR. They also reported a high completion rate.

Murray and colleagues suggested some amendments to the original scoring system.[86] Along with a couple of wording changes to aid clarity, the scores are now generated from responses graded from zero to four, giving a possible range of up to 48, 48 being the best function. They do not recommend transforming the score to a 0-100 scale, as they feel it could lead to confusion.

The OKS was developed by recording changes in the first six months after TKR. Murray and colleagues state that most of the improvement in function and knee scores takes place in the first year so it is not unreasonable to assess the outcome at one year. It is of questionable benefit beyond one year. Murray and colleagues summarized the data from a large, multicentre RCT of patients who had TKR surgery. [87] They used the OKS, SF-12, and the EuroQol as their principal outcome measures as well as intraoperative and postoperative complications. These data showed that absolute scores tend to decrease with age so a score less than the maximum 48 may be normal for an elderly patient. [86]

Garratt and colleagues performed a systematic review of the literature concerned with patient-assessed instruments for the knee. They found 31 articles relating to 16 patient-assessed instruments, which they assessed for reliability and validity. [88] The OKS has demonstrated acceptable evidence for internal consistency reliability (0.45-0.83) [85] as well as test-retest reliability (0.92).[89] It has been assessed for face and content validity. Garratt and colleagues recommended that a general health questionnaire should be used as well, such as SF-12 or SF-36.

Bremner-Smith and colleagues randomly selected volunteers from elderly persons' clubs, churches and voluntary hospital workers. Volunteers were excluded if they had any history of hip, knee or spine

disorder, or dementia. The 100 remaining volunteers were scored for the International Knee Score, the Bristol Knee Score, and the American Knee Society Score. The scores were obtained for both knees and a mean calculated. The scores were then normalized to obtain a percentage. [90] The scores were also examined in the presence or absence of a co-existing major medical condition. These were cardiovascular, respiratory, malignancy, visual disorder, or Menière's disease. They also divided the OKS into pain and function to assess for age-related changes.

They found a significant negative correlation between knee score and increasing age ($p<0.001$). They found the presence of a major co-existing medical condition was associated with a significantly lower knee score ($p<0.001$). The most significant negative correlations for both age and co-existing medical conditions were found in the function component of the score. This study found no significant relationship between score and socio-economic status. [90]

Murray and colleagues also reiterate that, with the OKS, if one or two questions are left unanswered, it is reasonable to enter the mean values representing all their other responses in order to fill the gaps. If more than two questions are unanswered, the overall score should not be calculated. If a patient indicates two answers for one question, the worst response should be adopted. [86] Bremner-Smith suggests that the statistical analysis needs to be non-parametric because of skewed data. [90] However, the Oxford Group has found parametric statistics to be satisfactory in most analyses and using a change score is less likely to be problematic as change scores tend to be more normally distributed. [86]

The MCID tends to be half of the standard deviation of change in other patient-reported outcomes. In joint replacement studies, the standard deviation of the Oxford change scores tend to be about ten but dips as low as six. Therefore the expected MCID is between three and five. A recent study identified an MCID for the OKS as five.[91]

Harcourt and colleagues [92] looked at the specificity of the OKS with special focus on the part played by co-existent hip or spinal pathology. They contacted 346 patients on the waiting list for hip or lumbar spine surgery, asking them to complete the OKS questionnaire. They also filled out a questionnaire to determine the possibility of co-existing or previous knee pathology.

They compared their results with a group of patients who had undergone TKR and completed an Oxford Knee Score, two years after their procedure. They received 194 complete responses and divided them into three groups. Group A were the 104 patients awaiting hip surgery who denied any knee symptoms. Group B were the 37 patients awaiting spinal surgery who denied any knee symptoms. Group C were the 53 patients awaiting either surgery who admitted to knee symptoms. Group D was the comparator TKR group.

They found that a patient awaiting hip or spinal surgery had a significantly lower Oxford Knee Score than those patients who were two years on from their TKR, regardless of whether the preoperative patients had previous knee symptoms. Analysis of groups A to C found no significant difference between the means, suggesting that knee pathology does not significantly affect the score in patients awaiting hip or knee surgery.

In a pilot study comparing post-discharge physiotherapy with usual care, Minns-Lowe and colleagues calculated that measuring OKS as the primary outcome at 12 months would need 521 patients per arm to achieve 90% power and a 0.05 level of significance. This was based on a minimum clinical difference of 2.5 on the OKS. [65]

Overall, the Oxford Knee score has been shown to be a valid, responsive disease/site-specific instrument which is of use in a clinical trial for the first year of assessment following total knee replacement. There were some issues with a couple of questions, which have been addressed by the original authors, and a change to the way the final score is presented has reduced confusion about the direction of benefit. It is suggested for use in combination with a general health questionnaire.

When one analyses the results, it is important to bear in mind that the functional component of the score is likely to be inversely proportional to the age of the patient and to be negatively associated with any pre-existing medical conditions as well as any hip or spinal pathology. Finally, the Minns-Lowe pilot study suggests that over 1000 subjects would be needed to adequately power a study looking at a rehabilitation intervention in a TKR population using the Oxford score as a primary outcome.[65]

2.3.2.1c   *SF-12/36*

It is well recognized in the literature that studies looking at outcomes from TKR surgery use a general health questionnaire as well as specific joint- or disease-specific questionnaires. The 12-item Short Form (SF) Health survey and SF-36 Health Survey are two of the most well-known and widely used general health questionnaires in the orthopaedic literature. In my proposed studies, they are important in certifying that the intervention and control groups are comparable in terms of their general physical and mental health before their operations.

The SF-36 has satisfactory reliability in patients with OA of the knee and was more responsive than OA-specific instruments in a group of outpatients. [93] There is evidence for SF-36 responsiveness following physical therapy for knee impairments [94] and total knee arthroplasty. [95, 96] It has been recommended that the SF-36 is used in conjunction with the OKS or the WOMAC for assessing the outcomes of TKR. [97] The WOMAC was found to be more responsive than the SF-36 in patients undergoing knee surgery [98] where the SF-36 was more responsive than the WOMAC in patients with OA attending a rheumatology clinic. [93]

The SF-12 HSS was developed using normative data for the SF-36 in the United States but there have been local equivalence studies finding it suitable for use in Australia.[99] All 12 questions were from the SF-36. These include 2 questions concerning physical functioning; 2 questions on role limitations because of physical health problems; 1 question on bodily pain; 1 question on general health; 1 question on vitality; 1 question on social functioning; 2 questions on role limitations because of

51

emotional problems; and 2 questions on general mental health. Like the SF-36 it is not age or disease specific but has been recommended in preference to SF-36 amongst other scores for patients who have undergone TKR, and it takes less time to complete. [52]

The SF-12 is weighted and summed using a scoring algorithm to provide both physical composite scale (PCS) and mental health composite scale (MCS) from the 12 questions with a range from 0-100 for each score. A score of zero represents the lowest level of health and one hundred represents the highest. The algorithms have been developed to give a mean score of 50 with a standard deviation of 10. The data obtained with the SF-12 has been developed, tested and validated by Quality Metric Incorporated.[100]

The scores change with age with the MCS increasing with age and the PCS declining. This means that it is important to compare scores within a study to the scores of age-matched cohort. This comparison gives an age-specific mean difference score. Each individual score has a cut-off point which is a 95% confidence interval calculated at 1.96 time the standard error of the measurement which for PCS is +/- 6.97 and MCS is +/- 6.24.[99] If the individual's difference score falls within this range then their health is considered to be average for their age.

2.3.2.1D   VISUAL ANALOGUE SCALE FOR PAIN – VAS

The Visual Analogue Scale (VAS) for pain is a subjective outcome which is easy to use across a variety of clinical settings. It can be a global measure of pain or an activity based measure of pain, such as in the WOMAC, depending on the specific wording of the question asked. It can also be analysed with parametric statistics.[101]

The patient is presented with a 100mm long line drawn horizontally across a sheet of paper. The left end of the line is marked with a zero and the description 'No pain'. The right end is marked with '100'

and the description 'Extreme pain'. The patient is instructed to place a mark perpendicular to this

horizontal line to indicate the level of their pain according to a specific question. The question used in

this study is 'How much pain are you feeling in the operated knee on an average day

within the last week?' The distance from the zero end is then measured to obtain a score in

millimetres.

Using this score, moderate pain is defined as a score of more than 30mm.[102]

Pain is a very important outcome for patients following TKR surgery. Although the WOMAC index

includes a pain subscale, this additional measurement can add more information to the analysis.

2.3.2.1F *ADHERENCE*

Studies suggest that non-adherence to physiotherapy is similar in proportion to drug therapies, where up to half of all patients do not follow recommended drug regimens. [103-106]

Positive incentives, such as improved function, can encourage adherence to a prescribed physiotherapy protocol. The incentive can also be an avoidance of a negative outcome such as not recovering to a level enabling participation in a favourite hobby or pastime. However, it is well noted that adherence to exercise regimens drops off over time.[107]

There have been efforts to reduce the proportion of non-adherent patients in medication-based trials, [103, 104] but to little effect. Sociological research suggests that patients use their own beliefs about drug taking, their personal experiences, and the information they have at hand to make 'reasoned and rational' decisions about taking their medication.[106, 108]

Campbell and colleagues explored adherence with physiotherapy from the patients' perspectives in a study which was nested inside an RCT designed to test the effectiveness of a complex physiotherapy intervention in reducing knee pain and mobility restriction associated with osteoarthritis of the knee.[109] They used qualitative research methods to find the levels of and reasons for non-adherence to physical therapies in 87 subjects with symptomatic patellofemoral osteoarthritis. The results of the RCT showed that five months after the start of the treatment, there was a small decrease in pain and a significant increase in the strength of the quadriceps muscle in the index knee. However, after one year, there were no significant differences in the outcome levels, most of which had returned to pre-treatment levels.

Adherence was assessed by the trial physiotherapist and then a group of patients were selected from the intervention arm with a mix of adherence and demographic variables. Twenty informants were interviewed in depth after their eight weeks of physiotherapy and eight participants identified as mostly adherent were interviewed after the one-year point.

Of this initial group of 20 patients at the 2-5-month stage, eight were fully adherent, and the rest only partially. The adherence of the physiotherapist's assessment had a high degree of concordance with the patients' accounts. They tended to continue with the exercises which they found easiest or felt were giving the most benefit. They also demonstrated two phases of adherence: initially when still attending physiotherapy sessions and a later phase.

The patients described an obligation towards the physiotherapist and a desire not to let her down as important reasons for high levels of initial adherence. There were also reasons of reciprocity and altruism given such as wanting to help the researchers and wanting to contribute to something worthwhile or future knowledge. This must be borne in mind when adherence is investigated. Continued adherence was more complicated and revolved around the interplay between the condition (symptoms of pain and stiffness in the knee), the perceived effectiveness of the intervention (exercises and taping), and "motivation" (the reasoning behind adherence and non-adherence).[109] Campbell and colleagues devised a model based around the "motivation to comply" (see Figure 3).

This was associated with the following themes.

- Attitudes towards exercise

The successful attitude towards exercise seemed to be that of those who could accommodate the exercise in their regular routine. This was more important for them than a positive disposition to exercise. However, when interviewed, patients also admitted to stopping the exercises because of a mixture of difficulties in doing them alone and a perceived lack of progression in the improvement of their symptoms. Competition with a partner was suggested.

- Perceived severity of the knee symptoms

Those patients with the most severe pain and/or loss of mobility were most likely to continue to exercise. They were encouraged by the reduction of their symptoms. In contrast, those with co-

morbidities, those who felt that they had less severe disease than others or those who just had a stoic attitude to knee symptoms lost the motivation to comply.

- Ideas about the cause of arthritis

There was a more resigned attitude in those who saw arthritis as caused by immutable factors such as age, obesity, and 'wear and tear'. They found it difficult to believe that the intervention would have any effect and this reduced the resolve to comply. On the other hand, those who felt that there was no cure for arthritis and there were things they could do to minimize its impact, including physiotherapy, were most likely to be continuing compliers.

- Perceived effectiveness of the intervention

There was a close relation between high levels of continued compliance and the perception that the physiotherapy intervention was effective. Those who noticed an improvement in their knee symptoms were much more likely to comply than those who did not.
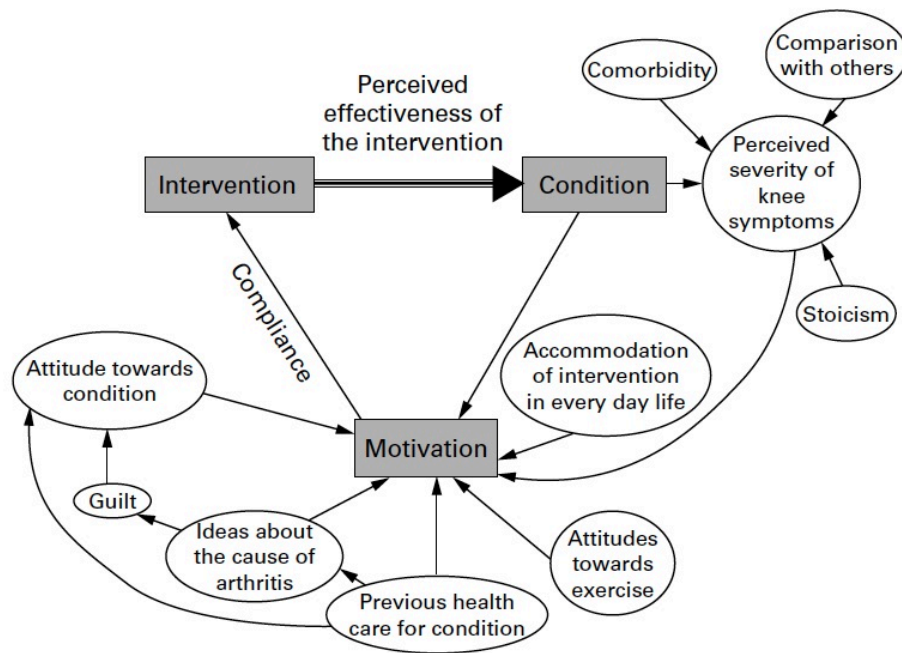
Figure 3 – Model of continued compliance (from 'Why don't patients do their exercise')[109]

The paper concludes that all informants were adherent to some extent initially, usually citing loyalty to the therapist or an altruistic desire to help in the research. However, long-term adherence was more complex. Although most understood the need for performing the exercises regularly, many only partially complied with what was easiest to accommodate into their usual daily routine or with the exercises which they felt gave most benefit. The similarities to drug regimen adherence were found to be that symptoms had to interfere with life sufficiently to warrant the treatment, and that the intervention needed to be perceived as effective while fitting into daily life. [106]

Therefore, Campbell and colleagues provided further evidence that the decision to cease therapy is often rational and reasoned. This needs to be taken into account in any future trials involving exercise intervention.

A further interesting point raised by Campbell and colleagues was the timing of the measurement of outcomes in an exercise intervention study. Should the measurement take place at the end of the supervised intervention period when adherence is maximal, nine weeks in their case, or at a later time when there is a split in the group between those who have incorporated the exercise into their daily life and those who have not?

As there was such an effect of adherence when patients were still seeing the physiotherapist, continued surveillance or follow-up may be necessary to maintain enthusiasm, education, and adherence.

The authors concluded that the patients needed to perceive the symptoms as interfering with life sufficiently to require treatment and that an intervention needed to be perceived to be effective and suitable for incorporation into everyday life.

This is an important observation. If the patient deems that the symptoms cease to warrant the extra effort, then the exercises will not be done. adherence is driven by the perceived reward from the point of view of the patient.

Another large incentive seems to be doing well to please the therapists and the surgeon. Arthroplasty patients are clearly informed during the preoperative and acute phases or rehabilitation that their recovery is largely down to their own efforts. They must do their exercises regularly in their hospital bed, even if they are painful. The physiotherapist judges their progress on a daily basis and the range of motion of the knee is a very visible marker of progress and therefore a proxy for effort made by the patient. The surgical team reviews the patient's range of motion on their daily ward round. If this is not satisfactory, they will ask the treating physiotherapist for reasons why. The patient does not want to be seen to have not done the work in front of the surgeon. The patient often has a clear goal of attaining 90 degrees of flexion before discharge.

Following discharge, most patients are seen regularly at a rehabilitation centre, either as an inpatient or an outpatient. They are faced with clear markers of progression and targets to aim for. They know they have a six-week appointment with the surgeon and they will again assess their range of motion as well as their general function. There are clear incentives to improve and to do this through physical therapy exercises.

Following the six-week check, there is often little further interaction with healthcare professionals until the one-year point. The outpatient physiotherapy finishes around this time and there is no further monitoring. [109] There is a significant decrease in supervision, new goal-setting, and assessment of progress.

The important factor here has to be the level of function which the patient is aiming for.
In 2002, Frost and colleagues looked at 47 patients after TKR and randomized them into functional or traditional exercise groups.[32] They lost nearly 50% at one-year follow up and these patients were

assessed as having low motivation during inpatient rehabilitation. This only tells us that patients with low motivation in the immediate inpatient phase of their postoperative recovery tended to be less inclined to continue in a research project. The majority were lost to follow-up as they refused to continue or were just not contactable.

Adherence to a rehabilitation protocol is a significant issue in clinical trials and in clinical practice. The strategies employed to improve adherence must take into account the patients' beliefs about exercise and its association with their recovery from surgery. The exercise has to be incorporated into their daily routine and they must feel that it is effective and perceive a benefit to them personally. An element of oversight by the treating clinician via the automatic recording of quantity and regularity of exercises performed could also help. Finally, it is important to recognize the level of function which each patient is trying to return to, and if necessary change their expectations so that they are aware of the possibility of gaining a higher level of function than they previously believed possible, with the appropriate volume of rehabilitation.

2.3.2.1G   SATISFACTION

Following a TKR, most patients show improvement in pain, stiffness, self-reported function and quality of life. Only 75 to 89% of patients are satisfied with their outcome, which is a worse statistic than that for patients having a total hip replacement.[8, 9, 14, 110-116] The greatest improvement is reported in the first six months.[117, 118] Studies have shown the factors which contribute to this satisfaction include self-reported quality of life variables, preoperative mental functioning, expectations, postoperative pain, and joint stiffness.[7, 9, 10, 114, 116]

It is vital to know the nature and strength of these relationships so that they can be addressed before the operation. It is possible to use methods to educate patients regarding their expectations before the operation to maximize their postoperative satisfaction. There is a conflict between measuring expectations in a quantitative or a qualitative manner.

Westby and Backman explored the experiences of patients and health professionals in Canada through focus groups. The stakeholders of interest were TKR (and THR) patients, allied health professionals, physicians and orthopaedic surgeons. They uncovered six major themes:

1. Let's talk – poor inter-professional as well as patient-provider communication.

2. Expect the unexpected – patients felt inadequately prepared for certain challenges – pain intensity and management, sleep disturbances, psychological issues, and unrealistic activity expectations. Advice on whom to go to for more pain meds.

3. It's attitude which counts – the importance of patients remaining motivated.

4. It takes all kind of support – peer, spousal and family support.

5. Barriers to recovery – patient, provider and system level factors.

6. Back to normal – the most important view of patients was to return to 'Normal' after surgery, over and above being pain-free and mobile. This normalcy also included the desire for a consistent method of measuring outcome.

The overriding tenet of the whole paper was that the health professionals wanted more integrated communication throughout the whole process and the patients wanted to be treated as a whole with inter-professional communication about their own progress. [119]

An interesting suggestion was the use of centralized information in the form of a communication document, which stayed with the patient. Eight years on, the obvious suggestion is a personalized website with accessibility for smart phones and tablets. [119]

Mahomed and colleagues [120] used a quantitative methodology to show that expectations played a 'central and dominant' role [121] in influencing satisfaction, although the nature of this relationship remains unclear. [122-124]

Woolhead and colleagues, in a letter to the editor of the *Journal of Rheumatology*, explained that although patients have been shown to be able to describe their expectations in quantitative studies using questionnaires, rating scales, or open-ended questions, qualitative studies using semi-structured interviews found the opposite. [125, 126] They described how patients tended only to be able to describe their hopes and fears, hope being either idealistic or pragmatic. These feelings reflected an optimistic view of the outcome and the probability of achieving this. They described how hopes are emotionally based whereas expectations tend to be based on rational thought and logical reasoning. Therefore, the patients in their qualitative study were not able to forecast their 'expectations' but only theorize their hopes and fears. The patients could, however, retrospectively describe their expectations after the operation. As many realities are unanticipated, individuals may not know what to expect. Woolhead and colleagues suggest that for this reason, expectations cannot be used as a starting-point on which to base an assessment of satisfaction level. [121] They postulate a public positive expression of outcome which matches the socially desired view of the operation as a success and a private expression of outcome which reflects ongoing issues such as pain.

Mahomed and colleagues responded to this letter by disagreeing on two major points. First, they pointed out that retrospective recall of preoperative expectations is open to recall bias. Second, they argued that Woolhead and colleagues were confusing satisfaction and the WOMAC score as a marker of functional outcome, a well-validated measure. [121]

Vissers and colleagues looked at two further factors which they felt were relevant and potentially modifiable – the functional capacity of daily activity and actual daily activity. [127] As these factors have been demonstrated to be related to quality of life and mental health, [128, 129] the authors hypothesized that TKR patients with a higher functional capacity and actual daily activity level would be satisfied with their results more often. Patient satisfaction was measured with the question 'How satisfied are you with the results of the surgery?' and responses were graded on a five-point Likert scale. The functional capacity of daily activity was assessed by three tests. Walking was assessed with the six-minute walk test, a timed five-step stair-climbing assessment, and a timed five sit-to-stand movement assessment. They measured actual daily activity using an activity monitor for 48 hours, six weeks

before the operation, and six months afterwards. Expectations and fulfilled expectations were assessed by questioning the patients preoperatively on expectations of pain after surgery, the limitations of daily activity, and the overall success of the surgery. They were then questioned six months after surgery as to whether these expectations had been fulfilled. Joint strength was measured with a hand-held dynamometer, and the patients completed the WOMAC, Hospital Anxiety and Depression Score (HADS) and the SF-36 questionnaires.

The satisfaction levels were 50% very satisfied, 22.7% moderately satisfied, 11.4 % neutral, 9.1% moderately dissatisfied, and 6.8% very dissatisfied. They found that whereas preoperative functional capacity and actual daily activity were not related to patient satisfaction, a higher mental function on SF-36 and HADS was. Only the walking test showed a relationship with patient postoperative satisfaction.

Therefore it was concluded that functional capacity and actual daily activity do not contribute to patient satisfaction. The patients who were satisfied more often were those who experienced less pain and fulfilled their expectations relating to pain.[127] As the authors only followed the patients up for six months, it was unclear what the results would be at one year. The question which remains is whether patients have a low MCS score preoperatively because they are suffering pain and functional disability or whether a preoperative psychological intervention aimed at reducing distress could improve their satisfaction following surgery.

It appears that the greatest improvement in satisfaction occurs in the first six months. This could mean that this time period is the one to target or that the period beyond six months is relatively neglected by healthcare professionals and satisfaction gains can be made here. Pain seems to be more important than function in the postoperative satisfaction ratings given by patients. Whatever the cause and time period, patients want a unified approach from the healthcare professional team with good communication and empathy.

2.3.2.1E   COMPUTER ABILITY AND ACCESSIBILITY OUTCOMES

In preparation for allocating a group of arthroplasty patients to rehabilitation with the Nintendo Wii-Fit, it was important to know if they possessed the ability to use technology and had a willingness to do so. The main trial required all patients with access to the internet and a working email address to fill in their questionnaires online in response to a prompting email. It also required the intervention group to be able to use a Nintendo Wii at home. One of the concerns raised about the trial in the set-up phase was that the age of the patients undergoing a total knee replacement meant they would not be technically competent enough to meet either of these two requirements.

In 1989 Davis and colleagues said that "Understanding why people accept or reject computers has proven to be one of the most challenging issues in information systems research'.[130] Studies have shown that attitudes play a key role in predicting the acceptance of the user as well as their satisfaction with the use of computers.[131, 132] Further studies tried to determine the effects of attitudes and beliefs on individuals' use of computers.[133-135]

A formal assessment of the technology available to the patients, their competency with it, and their willingness to use it would be useful for planning the trial and as a baseline demographic measure. There are a number of outcome measures for computer attitudes and abilities, including the Computer Attitude Scale (CAS). Unfortunately, many of them either concentrate on specific areas of computing which are irrelevant to this study such as setting up a database and writing computer code, or they were out of date, asking questions about using floppy discs and mainframes. Looking at over 30 instruments designed to assess attitudes towards computers, Shaft and colleagues found that only four had been assessed for stability over time and three were specific to an educational setting.[136] With the rapid pace of change in computers and digital technology, this is even more of an issue ten years on. For example, demographic data from 1988 indicated only 1% of adults over 65 years old use a computer. [137] More recent data from the UK show that whereas less than 10% of over 65-year-olds used a computer daily in 2006, by 2013 the number was 37%.[138]

The CAS was validated in 1985.[139] It incorporates the three subscales of computer liking, computer confidence, and computer anxiety. It was initially targeted towards secondary students but two further papers in 1986 applied it to teachers and the general population. However, the sampled groups were still students and secondary teachers. [140, 141]

The Computer Attitude Questionnaire was validated in 1993. It is a self-reported instrument using a Likert-type scale. It was designed for use with middle school children aged between 10 and 12 years old.[142] It has many specific references to teachers and homework, which makes it unsuitable for adults.

The Attitudes Towards Computer Usage Scale was developed in 1987 by Popovich and colleagues.[143] It comprised 20 items and a seven-point Likert scale. It has some questions relating to video cassette recorders and typewriters which date it considerably.

The validation studies show that the majority of these studies took their sample population from student teachers, undergraduates, or college students. These may well not be appropriate for the age group in arthroplasty studies.

The Attitudes Towards Computers Questionnaire (1992) was designed for the elderly and tested 101 adults in a community home between the ages of 57 and 87 (mean 75.1 s.d. 5.93). [144] It examined whether the attitudes of older adults became more positive as they gained computer experience. It demonstrated that older adults' computer attitudes are modifiable and that direct experience with computers is an effective means of change. It also found that it was important for the elderly subjects to experience success when operating the computer as they felt more comfortable.

The Nintendo Wii software has been designed to be simple to use and provides regular feedback and advice on both the use of the device and the exercises. The simple advice relating to navigating the

software to find the relevant exercises and tests is likely to add to the feelings of comfort and achievement as the subjects use it for the first time.

It has been known for some time that attitudes drive behaviour and that these attitudes can be changed through experience. [103, 145-148] In the case of attitudes towards computers, this can be through direct use of a computer or indirectly, through the observation of others using one. This indirect observation now pervades the ubiquitous nature of technology in our daily life from smart phones to smart televisions. The media make everyone very aware of available technology such as gaming devices and tablet devices. In the case of the Nintendo Wii, an extensive advertising campaign targeted the elderly, using more mature actors or depicting games played with grandchildren.

The instruments described all contained too much detail or did not ask about aspects of computing which are relevant to this thesis. The subjects of this thesis were going to be subjected to a large volume of questionnaires and so it was felt that this one needed to be concise. This was the driver of the decision to create a simple questionnaire focused on the specifics relevant to the TKR-POWER study. The questions assess the ability to access the Internet at home, access to an email account, shop on the internet as a proxy for general internet ability, and the awareness and use of gaming consoles. It was named the Computer, Internet & Gaming Ability for Rehabilitation score (CIGAR).

2.3.2.2   Objective Outcomes

Objective outcomes are assessed by the clinician and those specific to the knee include range of motion, alignment and stability. These values can be combined into an objective score such as the objective part of the American Knee Society Score. There are also more functional assessments including quadriceps strength, timed walks, stair climbs, and the TUG. [149, 150]

The majority of patients who have a TKR attain good-to-excellent ratings on self-assessment questionnaires. [151] This improvement in scores is not matched by improvement in performance-based functional scores. Only 50% of subjects in the rehabilitation trial of Moffet and colleagues reached the level of the age-matched controls at one-year post-TKR surgery.[39] Walsh and colleagues found patients with TKRs had long-term deficits in walking speed and stair-climbing ability, with women affected more. They took twice as long as controls to ascend and descend a flight of 10 stairs. [152]

The advantage of objective scores is that they measure a quantifiable variable. The results can be compared with those of other surgeons and patients as well as across research papers. They give an absolute measure as to the result of an operation such as a total knee replacement. The main disadvantage is that they do not always correlate well with patient-reported function or satisfaction.

2.3.2.2A   AMERICAN KNEE SOCIETY SCORE

In 1989, Insall and colleagues published the rationale behind the Knee Clinical Rating System, which has become known as the American Knee Society Score (AKSS).[153] They separated the functional rating from the objective knee rating. The knee was rated in three areas – pain, stability, and range of motion. They dealt with flexion contracture, extensor lag, and malalignment as deductions from the score. A well-aligned knee with no pain, flexion > 125 degrees, and negligible AP and ML instability scored 100 points. The patient function considered walking distance and stair-climbing with

deductions for walking aids. The maximum score for function was 100 if the patient could walk unlimited distances and go up and down stairs normally.

The objective knee rating is a detailed examination of the surgical outcome. It is useful to be able to stratify patients with this rating so if any functional deficit is present a surgical cause can quickly be investigated or discounted.

The functional rating, however, is limited in its scope. Many patients now demand more from their knee than merely walking and going up and down stairs. This rating has a ceiling effect which limits its use in studies looking to improve high-end function for active, motivated patient groups.

Bremner-Smith assessed the Bristol Knee Score, OKS and AKSS for demographic-related changes in the knee scores in a normal population.[90] The AKSS suffers the same significant reduction with age and co-existing morbidities as other scores. The functional component was the major component of change, which was the same for all scores.

In summary, the objective rating of the AKSS is a useful tool with which to assess the surgical outcome of a knee post-TKR but the functional rating is limited compared with other available outcome assessments. The range of motion of the knee is a common parameter which is measured and analysed separately in the knee literature.

2.3.2.2B   BALANCE

Humans need to maintain and control their balance over a small base of support to be able to stand and walk. This balance control has many inputs including the brain, the musculoskeletal system, the visual system, and the inner ear. Balance control can be affected by pathologies, medications, and the general functional decline of ageing.[154]

The risk of falls increases with age, with approximately one in three people over the age of 65 years old living in the community falling at least once in the course of a year. This number increases if the individual is living in an institution. [155, 156] These falls are of importance as they can decrease confidence, reduce independence and also place a burden on the healthcare system as 50% of the falls lead to injury. Of these injuries, 5 to 10% are fractures.[155, 157] When the injury is a fractured neck of the femur, it is worth bearing in mind that the one-year mortality rate can be as high as 36%.[158]

Risk factors can be intrinsic or extrinsic. [159] The risk factors most associated with falls include muscle weakness, gait deficits, balance deficits, history of previous falls, use of assistive devices, arthritis, depression, and older age. The risk of falls increases with an increased number of risk factors. [155, 160] Of these, postural instability is an independent risk factor and it increases with ageing. [161-163]

The majority of patients with a TKR in Australia in 2013 were aged between 65 and 84.[2] It is understood that balance ability decreases with age and the rate of falls increases. [155, 164, 165] The decline is multifactorial in nature and is affected by musculoskeletal factors such as deteriorating joints and lower limb strength as well as deteriorating visual ability to assess hazards and a slower reaction time to an unexpected change in circumstances such as a slip. A good sense of balance is important for the patient's confidence in activities of daily living as well as sporting activities. It is therefore an essential part of the rehabilitation programme if the aim is to maximize patients' postoperative function. A large body of literature deals with assessing and training the elderly with a risk or history of falls regardless of musculoskeletal surgery.

A link has been shown between poor balance and the risk of falling. The priority in terms of assessing patients who have fallen is trying to work out their risk of falling again. Consequently, preventative measures can be directed towards those who need them.

There is evidence demonstrating the importance of assessing both the nature of the fall itself and the patient's objective balance ability. Overstall and colleagues observed that patients who had fallen as a result of a trip or slip during activity had normal sway for their age. This differed from those whose falls had occurred in the absence of a hazard.[164] Wild and colleagues performed a study on elderly people who had fallen at home. They found that those subjects who had impaired balance, as determined by the response to pressure over the sternum, had a much greater tendency to further falls. They also had much higher mortality in the year following the index fall than had subjects whose balance was apparently normal using this test (Chi squared 17.52, $p < 0.01$).[166]

Interest in the assessment and management of balance deficits has led to the development of a number of balance-measuring instruments.

To assess the physical parameters associated with fall risk, there are observational performance tests such as the Berg Balance Scale. This has been shown to provide valuable information but suffers from ceiling effects and limited precision in detecting small changes in performance.[167, 168] Other tests include the Tinetti Balance scale and the Timed Up-and-Go test.[169-173]

Physical parameters can also be assessed by looking at performance measurements and using equipment. These assessments can look at static or dynamic balance. The equipment is often a force platform, measuring the centre of pressure (CoP). These are considered the gold standard of balance measurement and have identified important outcome measures too subtle to detect with a subjective scale.[174 175]

However, the use of force platforms to assess balance is limited by cost, the need for technical expertise, and their lack of transportability.

It is important to clarify the instrument used to measure balance in any assessment and establish the variable it is trying to measure. Standing balance measures can look at symmetry or steadiness. Symmetry relates to the ability to distribute weight evenly between the two feet whereas steadiness is the ability to keep the body as still as possible.[176]

Centre of pressure describes the location of the ground reaction force on the surface of a balance plate. Its position can be described by co-ordinates along the medial-lateral axis (x-axis) and antero-posterior axis (y-axis) of the force platform. Gait analysis force platforms can also measure the vertical component (z-axis) of the ground reaction force. With the z-axis, the co-ordinate described represents the centre of gravity. The centre of pressure moves in response to the motion of the centre of gravity (CoG), thereby correcting the sway of the body. In order to achieve this, the motion of the CoP must be greater than the CoG.[177, 178]

A number of assessments of standing balance have been calculated from the raw CoP data used in the literature[176, 179, 180] (see Table 2). They each measure a different aspect of the CoP movement such as the path length of the line traced between each successive CoP coordinate or the size of an ellipse covering 95% of the data points.

Table 2 – Measures of standing balance

Measure	Formulae	Key
Path length[180]	$= \text{Sum}(\text{SQRT}(((\text{ABS}(X^{n+1}-X^{n})))^2 + ((\text{ABS}(Y^{n+1}-Y^{n})))^2$	ABS – magnitude without a sign
Distance from start	$= \text{Sum}(\text{SQRT}((X^{n}-X^{1})^2 + (Y^{n}-Y^{1})^2)$	$X^{n}=n^{th}$ x-axis value $X^{1}=1^{st}$ x-axis value
95% ellipse	$5.99 = \{[X(t)-uX]/sX\}^2 + \{[Y(t)-uY]/sY\}^2$	u = mean s = st dev t = timepoint
Sum of x-axis & y-axis amplitudes	$= \text{ABS}(\text{MAX}(X)-\text{MIN}(X)) + \text{ABS}(\text{MAX}(Y)-\text{MIN}(Y))$	ABS – magnitude without a sign

There is no consensus in the literature as to which measure is better for any given population or clinical situation. It is important to work out which aspect of balance the trial is focused on and match it to the characteristics of a chosen measure.

Voluntary postural sway movements represent a simple approach to examining deficits in postural control which may contribute to falls. The three main categories of voluntary postural sway tasks include maximum voluntary leans, continuous steady-state voluntary sway, and rapidly initiated voluntary sway movements. [181] Maximum static lean amplitudes have been found to be weakly associated with or non-predictive of falls in healthy older adults. Reactive postural sway movements significantly differentiate between younger and older men and between young adults, older adults at low risk of falls, and older adults at high risk of falls. For rapid orthogonal switches of voluntary postural sway between the AP and ML directions, ageing, and elevated fall risk among the elderly result in a slower reaction time, reduced COP response amplitudes, and altered postural coordination relations.

There is a significant increase in postural sway under single-task conditions in older subjects compared with younger ones.[182] There is also a significant increase in postural sway in both age groups while performing dual-task tests. The study showed that postural adjustments require cognitive processing and more attention. The cognitive task interfered with the ability to maintain postural stability because of resource competition. Teasdale and Simoneau showed that as the sensory information decreased, the postural task became increasingly difficult for older subjects and required more of their attention capacity.[183] These findings suggest that balance control is not only a motor output as a discrete entity but is also concerned with motor behaviour as an integrative outcome of perceptual, cognitive, motor, and sensory processes. The notion that the restoration of postural control is partly based on cognitive processes which are not directly accessible in a motor assessment procedure using simple tasks has clear clinical implications.[184]

In a systematic review, Piirtola and colleagues looked at prospective follow-up studies using force platform equipment to measure postural balance as a change in CoP.[175] The study had to analyse the association between balance and risk of falling in people aged 65 years or older. They found that out of the nine studies which fulfilled their criteria, the most common outcome variables measured included mean or maximal amplitude, speed of the medio-lateral (ML) and antero-posterior (AP) movement of the CoP, mean ML, and AP CoP displacement (in mm or cm).

Predictive sway parameters for falls were higher ML sway amplitude with eyes open and closed, higher AP speed with eyes open, and higher root-mean-square values for ML CoP displacement with eyes closed.[185, 186] Boulgarides and colleagues found the predictive parameter for falls was higher mean speed of CoP displacement while standing on a firm surface with eyes closed.[187] Bergland and colleagues found a higher amplitude of ML CoP movement was predictive of indoor falls but not of outdoor falls.[188]

For recurrent fallers, the values of the amount of ML sway (cm) with eyes open [189] and sway area ellipse [186] were higher compared with non-fallers or those falling only once.[175]

It must be taken into account, however, that these predictors come from a review of only nine studies of which only five were able to provide evidence of the predictive value of force platform measures for falls. [175]

In all but two of the studies, the subjects were community-dwelling older people who were relatively healthy. Piirtola and colleagues surmised that as subjects become frailer and have a substantial risk of falling, the balance tests could be more sensitive as the overall importance of intrinsic risk factors increases. This theory is supported by the results of Bergland and colleagues who found an association between force platform measures and indoor falls but not outdoor falls.[188]

Sihvonen and colleagues conducted an RCT and found that targeted balance training could reduce falls risk during a one-year period and the CoP measures demonstrated a reduced speed and area of CoP movement.[190, 191]

Poor results in force platform tests are associated with difficulties in activities of daily living.[192] The ability to assess standing balance can provide important information which ranges from predicting falls in the elderly to assessing post-operative recovery after orthopaedic surgery. [193-196] However, the evidence for prediction falls from standing balance measures is mixed and unclear.

There is a loss of functional fitness associated with the decline of the ability to produce force. Therefore it can be postulated that the maintenance of muscular strength is crucial for older adults' maintenance of their independence. [197-199] The decline of lower body strength in older individuals can vary from 14 to 16% per decade.[200] We know that this can be reversed, to a degree, by strength and aerobic training. The optimum timeframe is long-term, non-stop. However, most studies look at periods of between10 and 12 weeks. Patients recovering from an operation are even more likely to have declined in strength. This occurs before the operation because of the reduced mobility caused by painful arthritis, and after the operation because of the pain of the procedure. It is fair to assume that any activity following an operation is likely to increase the patients' strength and functional

ability. However, the protocols used are important as the lower limb pain which patients experience following a TKR gives them an extra disincentive to exercise.

Balance plays an important role in the overall functional outcome following a TKR. Many studies have demonstrated the decrease in proprioception of the postoperative TKR knee. This manifests as a reduced ability to detect joint position and motion, delayed muscle latency, altered amplitude of muscle activity, and decreased postural control.[36, 201-203] This suggests that targeting exercises in the rehabilitation period towards the balance and proprioceptive system could improve functional outcome.

Howe and colleagues undertook a Cochrane review in 2007 looking at exercise interventions designed to improve balance in older people living in the community or in institutional care. Statistically significant improvements in balance were observed across a variety of outcome measures for exercise interventions compared with controls. The exercise interventions were heterogeneous, occurring mainly in gyms or clinics with supervision. The duration ranged from four weeks to twelve months and was typically for an hour, three times a week. Gait, balance, co-ordination and functional tasks exercise interventions showed statistically significant improvements compared with control in measures such as AP stability during stance with eyes open, single leg stance with eyes open and the Berg Balance Score. General physical activity demonstrated statistically significant improvements in Timed Up-and-Go and functional reach testing. [204]

Piva and colleagues performed a pilot study in 2010 to investigate the feasibility and safety of an exercise programme using challenging balance tasks for TKR patients. [205] A minimum of two months from TKR was required and no recent history of falling or a current walking distance of less than 100ft (30.48m). The patients received 12 sessions of supervised exercise over six weeks with the functional training programme based on Moffet and colleagues and a balance programme based on the protocol of Fitzgerald and colleagues. [39, 206] At the end of the supervised programme, participants were asked

to continue their exercises at home twice a week for four months. Adherence was recorded in an exercise log. They were called once a month to encourage adherence. The study personnel who performed all outcome assessments were blinded to group allocation. Outcome data were collected at baseline, after the supervised programme (two months), and after the four-month home programme (six months). They collected the WOMAC score, the Lower Extremity Function Scale, exercise diaries, self-selected gait speed, a timed chair rise test, and single leg stance time. Forty-three patients were randomized of the 76 who were assessed for eligibility and 35 completed the study (20% drop-out). Adherence to the supervised sessions was 100% and to home sessions was 67% in the functional group and 64% in the balance group. Importantly the balance exercises did not increase pain or stiffness via the WOMAC subscales compared with the Functional group.

Schwartz and colleagues looked at balance outcomes preoperatively and at one year post-TKR and found that, although balance improved overall, those in whom balance improved more had better functional outcomes. [195] Webster and colleagues showed in an observational study that a group of patients post-TKR showed improved functional outcomes if they had improved balance confidence. [207]

A fall can have serious consequences for a patient. The risk of falling can be due to a deficit in the patient's balance and proprioception following TKR surgery. Levinger et al compared a group of patient shaving TKR surgery with an age-matched control group performing a battery of tests as part of a physiological assessment to quantify falls risk. [194] They found that both groups had a low risk of falling at 4 months post surgery and were not significantly different. However, there were individual components of the risk assessment where the surgical group had significantly worse proprioception and quadriceps strength both before and after surgery. The surgical group also had a significantly greater fear of falling compared to the controls.

Swinkels et al conducted a longitudinal study of falls in 99 TKR patients, recruited preoperatively.[208] There were 24 patients who had fallen in the three months prior to surgery and 46% of those

continued to fall post-operatively. They found that the odds of a pre-operative faller becoming a faller in the first post-operative year are almost eight times those of a non-faller (adjusted OR 7.75, 95% CI 1.7-35.7, p=0.008). The pre-operative WOMAC scores did not significantly affect the odds of falling in the first post-operative year. While balance confidence improved significantly post-operatively in the non-fallers, this was not seen in the fallers.

Balance is an obvious function for targeted rehabilitation. It is an important outcome to measure as the link between balance, function and risk of falling is investigated further in this population. However, it is highly complicated and without a clear consensus regarding assessment in the elderly patient population. This means that it is an outcome of interest but is secondary to an outcome which can be measured and understood more clearly.

2.3.2.2c   STRENGTH MEASURES

Muscle weakness through loss of muscle mass and overall strength has been well documented in ageing as well as its effect on functional performance.[209-213]   Specific weakness of the quadriceps femoris muscle has been implicated in the development and progression of joint degeneration [214, 215] and has been reported to be the strongest single predictor of functional limitations in patients with knee OA.[216] To add to this problem, quadriceps strength loss of 60% has been demonstrated in the acute postoperative period.[217] It is rare for such strength to be restored to normal levels and the majority of evidence points to substantial weakness in long-term assessments.[218, 219]

Therefore, quadriceps muscle strength represents an outcome measure which has a direct clinical relationship with the patient's postoperative function and it can be measured preoperatively. Were it to have a predictive value, it could be used to direct exercise therapy preoperatively.

Reduced quadriceps muscle strength is frequently observed in orthopaedic patients with knee pathology. The quadriceps are affected by the pain inhibition of osteoarthritis and post- surgical pain and relative immobilization. It has been suggested that reduced quadriceps muscle strength is a surrogate marker for the progression of osteoarthritis in some patients as well as one of the most

important predictors of functional disability. The strength of the quadriceps can be related to both subjective knee scores and functional test outcomes following TKR. [220]

Muscle strength and power can be measured in different ways.

- The maximum force generated by an isometric contraction

- The maximum load which can be lifted once

- The peak torque during an isokinetic concentric eccentric contraction

There is little agreement in the literature on the ideal method for measuring orthopaedic patients. Isometric strength assessment is easy, quick, and reliable and is therefore well suited to measurements in clinical rooms. Isokinetic testing can take a long time and is much more expensive. Usually the large dynamometer is housed a reasonable distance from where the patients are usually assessed by the surgeon, making it impractical on a large scale.

Isometric activity is rare in everyday life but the measure has a strong predictive relationship with functional capacity.

Maffiuletti describes the isometric evaluation of knee function using a handheld dynamometer.[220] This method allows for quick and cheap assessment of maximal isometric contraction of quadriceps and hamstrings. This is preferable to transporting every patient to the lab for a formal strength assessment which can take a long time to set up, let alone perform. The inter-rater reliability is easy to calculate thanks to the limited data processing and standardized calibration procedures.

Observational studies have demonstrated that lower limb strength is markedly reduced in the operated limb compared with both the unaffected limb [221] and age-matched control subjects.[219, 222] Strength and pre-existing gait deficits persist for up to 2 years after surgery.[219, 222-224]

There is a loss of functional fitness associated with the decline of the ability to produce force. Therefore it can be postulated that the maintenance of muscular strength is crucial for older adults to maintain their independence in their capacity to perform ADLs independently. [197-199] Strength

declines by as much as 16% per decade[200] but in an older population, this decline can be slowed by aerobic and strength training over 10-12 weeks.

The TKR patient population is even more likely to suffer a reduction in strength preoperatively because of the reduced mobility caused by their painful arthritis. Postoperatively, they experience disuse and pain inhibition of their knee muscle groups.

Strength is dependent on the integration of muscular and neurological systems. Isokinetic performance used to measure strength is the optimal integration of the neuromuscular properties to generate force. The potential for strength training to improve physical performance in older adults is dependent on the inverse relationship between strength (kg) and fatigue (sec). A weaker person is incapable of sustaining the contraction which allows performance in the most efficient manner. Time required to complete a task can also be a measure of balance and coordination. [225]

Lower levels of preoperative quadriceps muscle strength and lower levels of self-perceived functional ability predict decreased function between 6 and 24 months.[226-228]
The strength of the non-operated leg has been shown to be an important factor in overall function post-TKR. Zeni and Snyder-Mackler looked at 155 patients with primary unilateral TKRs for OA with minimal pain in the contralateral knee. These patients underwent a standard rehabilitation protocol, assessed at six weeks, one year, and two years. [229] Quad strength of the non-operated limb in the early postoperative period significantly improved the predictive ability of their model for TUG, stair-climbing test (SCT) and activities of daily living (ADLs) subscale of the Knee Outcome Survey (KOS) score. KOS-pain subscale and strength of the operated leg did not predict outcome scores.

Mizner and colleagues looked at the predictive value of strength in 40 subjects with OA scheduled for unilateral TKR.[227] They were assessed two weeks preoperatively and one year postoperatively via self-report questionnaires (SF-36 and KOS-ADLS), ROM, functional assessment with TUG and SCT, and measurement of quadriceps strength. Isometric quadriceps strength was measured in a seated

dynamometer with hips flexed to 90˚ and the knee flexed to 75˚. The authors used a burst

superimposition technique for a maximal voluntary contraction reading, normalized to BMI.

There was a significant improvement in bodily pain, extension ROM, SF-36 PCS, KOS-ADLs, SCT, and

TUG between the preoperative assessment and the one-year assessment analysed with paired T-

tests. However, there was no significant change in quadriceps strength, knee flexion ROM, or SF-36

MCS. The regression analysis using predictive models showed no significant correlations between any

of the independent variables (preoperative involved quad strength, preoperative bodily pain score,

age, and preoperative knee flexion.) Preoperative quad strength was a significant predictor of TUG

and SCT score, accounting for an R squared of 41% and 54% respectively (p<0.001). However,

quadriceps strength did not significantly affect the predictive power for SF-36 PCS or KOS-ADLs. The

greatest predictors of postoperative questionnaire scores are preoperative questionnaire scores.[13, 67,

228] Therefore, although knee ROM and bodily pain are seen as important markers of outcome, neither

was predictive for postoperative function in these tests. It is quadriceps strength which accounts for

the R squared rising from 0.1 to 0.5 on the SCT.

In conclusion, although self-reported measures improved one year after surgery, the improvement

was not matched by that of the objective functional tests. This is consistent with other studies.[13, 112,

152] This study has limitations in that it has a higher proportion of males than is usual in the TKR

population and used a generic bodily pain measure, not a knee-specific pain measure. The sample size

was small but based on an analysis of 19 subjects with preoperative quad strength and one-year

functional assessments as the outcomes of interest.

One conclusion of this study is that a potential side-effect of delaying surgery is the development of

substantial weakness of the quadriceps. Quadriceps strength decreases over time in people with

progressive knee OA, over and above the age-related changes.[230] This can limit the ultimate

functional success of the surgery.

Valtonen and colleagues found that knee extensor strength deficits persisted ten months after TKR, suggesting that stairs could be limiting without this being addressed. [231]

Muscle strength is vital in recovery from TKR surgery. The quadriceps muscles extend the knee and are essential for all aspects of effective gait and function such as rising from a chair. Strength can be assessed as a direct power measurement of individual muscle groups or as a composite function such as rising from a chair.

2.3.2.2D   TIMED UP-AND-GO

The test assesses lower body strength and power as well as serving as a simple and objective measure of dynamic balance and mobility. This test started as the Get-Up-and-Go test (GUG). In 1986, Mathias and colleagues studied a mixture of 40 inpatients, outpatients and day patients of medical and geriatric departments of a large urban hospital in the UK. [232] They ranged in age from 52 to 94 (mean 73.8 years) and there was an equal number of men and women. All subjects were believed to suffer from some degree of balance disturbance although the authors do not elaborate further on how this conclusion was reached. A straight-backed, high-seat office chair with arm rests was placed three metres from a wall in a video studio. Subjects sat comfortably in the chair and then were asked to rise, to stand still momentarily, to walk towards the wall, to turn around without touching the wall, to walk back to the chair, and to turn around and sit down. After a trial run, subjects performed the test while two remote-controlled video cameras focused on them. They then stood on a force platform with feet part and eyes open for 30 seconds while a recording was made of body sway. The mean overall body sway in the AP and lateral planes was used as a measure of sway. Subjects were then asked to walk a short distance on a walkway while automatic recordings were made of gait speed, step length stride width, frequency of stepping, and the ratio of double support time to stride time. Video tapes of the tests were arranged randomly in blocks of ten and each block was viewed by medical professionals who scored performance on the following scale: 1 = normal, 2 = very slightly abnormal, 3 = mildly abnormal, 4 = moderately abnormal, 5 = severely abnormal. Normal meant that

the subject gave no evidence of being at risk of falling during the test or at any other time. Severely abnormal meant that the subject appeared at risk of falling during the test performance. Intermediate grades related to the presence of undue slowness, hesitancy, abnormal movements of the trunk or upper limbs, staggering, and stumbling as indicators of the possibility of falls in less favourable circumstances. The authors measured observer variation between the senior doctors, junior doctors, medical students, physiotherapists, occupational therapists, and laboratory technicians who observed the tests.

The observer variation analysis showed that doctors and medical students scored subjects more than half a point higher on average, grading a perceived abnormality as more serious than the physiotherapists or occupational therapists. The doctors' results also demonstrated a larger standard deviation than those of the physiotherapists or occupational therapists. The laboratory technicians, who were not trained clinical observers, scored very similarly to doctors. Although the difference between physiotherapists' scores and doctors' scores was significantly different, the Kendall coefficient of concordance confirmed that agreement within each group was higher than would be expected by chance.

The Pearson correlation coefficient between score and sway path was significant at 0.50 (p<0.001). Virtually all subjects with scores <2 corresponding to very slight abnormality, had a sway path of < 18 m/s, which is approximately the upper limit of normal for elderly people in their laboratory. However, many patients with a sway path of between 15 and 20 mm/s had high scores on the GUG test. Many of these were stroke patients whose gait appeared grossly abnormal but who in fact had good balance. This would suggest that observers were scoring gait, not balance. No patient with a sway path of >25 mm/s had a score of <2.

The correlation between score and gait speed was significant at 0.75 (p<0.001). Subjects were given no instructions about the speed at which they were to walk on the walkway. The authors felt that the results should reflect their feelings of comfort. All patients with a gait speed >1 mm/s had scores of

two or under. All subjects with a score of two or more had a gait speed of <0.8 m/s. This could be a simple correlation between the observers' scoring based on gait speed or it could be because of a subject feeling insecure and walking slowly.

When sway path was plotted against gait speed in 32 subjects, the Pearson correlation coefficient was -0.482 (p<0.01). All subjects with a gait speed of >1 m/s had a sway path of <20 mm/s and all subjects with a gait speed of <0.8 m/s seemed to form two groups, those with a sway path ranging from 25 to 50 mm/s and those with a sway path of between 10 and 20 mm/s. This could represent a group of patients walking slowly because of balance disorders and a group walking slowly because of other causes such as cardiorespiratory disease. Spearman rank correlation showed that the relationship between score & speed and score & sway, respectively was significant (p<0.001).

The authors concluded that the GUG test is almost as reliable as a measure of balance provided that other factors that may affect gait speed are allowed for by a trained observer and are not allowed to distract the observer from the estimation of the balance function in action. They confidently recommend the GUG test as a simple practical guide to balance function. It should be regularly assessed during rehabilitation and communicated between hospital staff, as a score of three or more indicates a risk of falling. [232]

Podsiadlo and Richardson developed the timed version of the test (TUG) in 1991 as a simple objective measure of dynamic balance and mobility. It assesses lower body strength and power and can discriminate between older adults who are functionally independent and those who are dependent.[233]

The timed test requires the subject to stand from a seated position, walk 3 m (9.84 feet) as quickly and as safely as possible to a line marked on the floor, turn around and walk back to the chair, and turn again and sit down, in the shortest time possible. The time is measured to the nearest tenth of a second and it has a high test-retest reliability in older women (R=0.90; 95% CI 0.83-0.95). [234, 235]

To discriminate between functionally independent and dependent older adults, participants who could complete the test in fewer than 20 seconds were considered to be independent in transfer tasks associated with basic activities of daily living. They also scored in the upper third of the Berg Balance Scale. Participants requiring longer than 30 seconds tended to be much more functionally dependent. Most required assistance with basic transfers. Only one patient in this group could climb a set of stairs independently and none were able to go outside without requiring some type of assistance.

In 1999, Rikli and Jones found that the distance needed to perform the test was impractical for home administration and that there was often confusion with regard to whether the patient had to cross the line fully with one foot or with both. They reduced the overall distance required to perform the test to 8 ft (2.44 m) and replaced the line with a cone to signal the turnaround point. They reported the test-retest reliability to be very high at 0.95. [234] They called this shorter test the Up-and-Go Test (UG test).

Zeni and colleagues found TUG and stair-climbing to be better predictors than subjective outcome scoring with KOS-ADLs with more than twice the R squared value. [229]

Gunter and colleagues found that the TUG was the best predictor of faller status compared with other clinical tests of mobility and laboratory assessments of strength and power. It was highly sensitive as 98% of 109 fallers were accurately classified, but had considerably lower specificity as only 31% of 48 non-fallers were correctly classified. [236] Shumway-Cook and colleagues were able to correctly classify 87% of the 15 recurrent fallers and 87% of the 15 non-fallers who participated in their study. Their sample size was smaller and their criteria for fallers were more stringent. They also put the cut-off point for discriminating fallers from non-fallers at 13.5 seconds. This was considerably shorter than the original Podsiadlo value of 30 seconds. The authors attributed this to the fact that the earlier study had included many elderly patients with known neurological diseases. [237]

The scores of the UG test can be used directly to compare the performance of an older adult with others of the same age and gender and still discriminate between physically independent and dependent older women. Rose and colleagues investigated the ability of the test to predict fallers.[238] They tested 134 adults (112 women) aged between 60 and 90 years old who lived independently in the community and could walk 50 feet without any assistive devices. They tested them with the TUG and the UG tests as well as the other elements of the Senior Fitness Test, Berg Balance Scale, a 50-foot walk, and five additional balance measurements using computerized standing balance assessment. Falls were defined as an event resulting in an individual unintentionally coming to rest on the ground, floor, or other lower level and not as the result of a major intrinsic event or overwhelming hazard. They excluded adults who had fallen on only one occasion in the previous year or had a medical reason for the fall. This was in an attempt to maximize the identification of true fallers as compared with non-fallers.

Their results suggested that there were significant performance differences between the two groups (fallers and non-fallers) for both the TUG and the UG tests. The mean scores for the TUG were 8.2 seconds and 12.3 seconds for the non-faller and faller groups respectively. For the UG test these results were 7.1 seconds and 10.6 seconds respectively. Logistic regression analyses showed that those adults requiring 8.5 seconds or longer to complete the UG test were classified as fallers, with an overall prediction rate of 82%. The specificity was 86%, with 61 of 71 non- fallers correctly identified. The sensitivity was lower at 78%, with 49 of the 63 fallers correctly identified. The authors pointed out that the results from the TUG are at odds with Shumway-Cook and colleagues' findings. [237] The cut-off score of 13.5 seconds established in the latter's paper of 2000 gave a prediction rate of 64% instead of the reported 90%. Only 19 of the 63 fallers in this study by Rose and colleagues required longer than 13.5 seconds to complete the TUG test, yielding a sensitivity of 30%. The cut-off which they found to be most accurate was 10 seconds with a sensitivity of 71%, specificity of 89% and predictive value of 80%.[238] There are a number of potential reasons for this: the Shumway-Cook study

had a smaller sample size (30 vs. 130), frailer patients, an older falls group compared with the non-fallers group, and inclusion criteria necessitating two falls or more within the last six months. Therefore, the UG test would seem to be a useful test to predict fallers in a community sample of elderly adults as it demonstrates significantly different results for fallers and non-fallers. The cut-off score for the two groups is 8.5 seconds.

Balance and agility are broadly studied neuromuscular components of functional fitness in older adults, because of their relation to the ability to perform ADLs. [239] The dynamic balance and agility as tested by the eight-foot UG test were shown by Hallage and colleagues and Shigematsu and colleagues to improve by around 20%. [235, 240]

The Up-and-Go test is a useful functional outcome which can be easily and quickly tested in a relatively small space. It measures and quantifies a function of everyday living for most patients, which makes it relevant for the overall outcome of TKR surgery.

2.3.2.3   OVERVIEW OF REHABILITATION STUDIES POST TKR BY LOCATION, TIMING OF DELIVERY AND INTERVENTION TYPE

The review by Minns-Lowe and colleagues examined the effectiveness of physiotherapy after total knee replacement. [37] The review included studies comparing location of physiotherapy, types of physiotherapy intervention such as functional exercises or group sessions versus individual supervision. The overall result was a pooled standardized effect size at three to four months of 0.33 (95% CI: 0.07, 0.58) demonstrating a small effect in favour of physiotherapy. However, at 12 months the effect size was -0.07(-0.28, 0.14), demonstrating no effect. The quality of life, pooled, weighted, mean difference in score after three months was 1.70 (95% CI: -1.0, 4.3), but after 12 months it was 0.03 (95% CI: -0.2, 0.25).[37] These results merited further investigation of different aspects of physiotherapy interventions commencing in the first few months after a TKR.

2.3.2.3A   LOCATION

The provision of postoperative TKR rehabilitation differs around the world. Rehabilitation in some form is recommended in the first six weeks.[47, 241, 242] This can be provided to both inpatients and outpatients. The outpatient provision can be one-to-one, group-based, or a home visit.

Post-acute rehabilitation in the United States is mostly provided as an inpatient or as home visits. In contrast, in Australia and the UK it tends to be conducted via outpatient services. The reasons behind the specific rehabilitation prescribed for a patient can be as much to do with anecdote, financial incentives, and sociopolitical influences as scientific evidence supporting a particular choice.[42]

A 2002 study based on three public hospitals in Victoria, Australia, found that 56% of patients had achieved functional independence enough to be discharged home. However, only 36% of TKR patients were actually discharged directly home from their acute hospital stay, which lasted on average 6.5 days.[243] This is longer than current averages, especially in the US where 3 - 3.5 days is a more common average length of stay.[244]

Naylor and colleagues surveyed 95 Australian hospitals of the 270 registered with the Australian

Orthopaedic Association National Joint Replacement Registry for performing primary TKR as of

November 2004.[47] They collected data from 60 acute hospitals and 5 rehabilitation providers and the

respondents were orthopaedic physiotherapists and department managers. Routine referral for

outpatient or community-based physiotherapy was commonly reported (73%, 95% CI 61-83%).

Routine referral for inpatient rehabilitation was uncommon (3%). A majority were offered or referred

to a primary programme with the predominant mode of service being delivered via outpatient

programmes (71%). One-to-one treatments (63%) outstripped supervised classes (23%) and

monitored home exercise programmes (9%) as the primary programme. Twelve per cent were not

routinely referred on for further physiotherapy as it was deemed unnecessary by either the unit or by

the treating surgeons. The mean frequency of one-to-one treatments was 2.1 (0.5-12) inpatient

sessions per week compared with a mean of two (range 1-3) for supervised group or home-based

therapy. The average outpatient session lasts 31 to 45 minutes in 71% of cases. Of the 52 centres

which did refer patients on to further programmes, the mean reported duration of post-acute

rehabilitation was 5.6 weeks (mode 4 weeks, range 1 to 18 weeks).

Artz and colleagues studied the provision of physiotherapy following TKR in the UK via a telephone

survey of 22 high-volume public orthopaedic centres. They found that eleven centres provided group

physiotherapy to patients after discharge and five offered one-on-one outpatient therapy.[241]

A telephone survey of 24 major orthopaedic centres in England and Wales asked the physiotherapy

departments about the current rehabilitation practice after TKRs. Seventy per cent were referred for

outpatient physiotherapy, 48% for group therapy, and 22% to one-on-one after discharge. The

centres which did not routinely refer patients for outpatient physiotherapy offered a variety of

options from referral if clinically indicated to a needs assessment at two weeks post-discharge.[241]

In Canada, 2001 data for Ontario showed that 42% of patients received inpatient rehabilitation and 30% had home care services on discharge.[245] Compare this with a teaching hospital in British Columbia, Canada where fewer than 5% of primary, non-complex cases are discharged to inpatient rehabilitation or other hospital settings.[245] There is a wide variation in the USA, where it is common in some cities for patients to receive two weeks of inpatient rehabilitation followed by outpatient rehabilitation as needed, if appropriately insured.[119] Data from France in 2001 showed that 44% of TKR patients were discharged into inpatient rehabilitation.[246]

Many studies have shown no significant difference between rehabilitation results when comparing locations such as outpatient or home-based therapy. They are detailed below.

Ko and colleagues [247] studied TKR patients in an RCT in three treatment arms lasting six weeks, randomized at two weeks post-op. All patients received the same initial home exercise programme. Exercises were progressed by the therapist for centre-based programmes (one-to-one and group) and monitored home programme participants were given fortnightly progression instructions in their exercise booklet. Outcomes were assessed up to one year post-surgery. The primary outcome was the OKS. Secondary outcomes included WOMAC pain and function subscales. They showed no significant difference between the three groups at any timepoint, up to one year after surgery.

Kramer and colleagues [248] randomly assigned 160 patients to either one-to-one clinic-based or home-based rehabilitation post TKR from the second week to the twelfth week post-surgery. They found no difference in any of the nine criteria measured including the American Knee Society Clinical Rating Scale, WOMAC total or function score, 30-second stair test, ROM, or SF-36.

Moffet and colleagues observed small but significant differences in mobility eight months after surgery in the group of one-to-one exercises versus usual care. The intensive group attended twelve 60-90 minute sessions in two months supervised by physiotherapists. The control group had usual care, which did include supervised home visits in some cases. The treatment effect size of the

WOMAC total score at four months between the intervention and control groups was 7.9 (95% CI 2.7 – 13.1, p=0.004).

Madsen and colleagues found no difference between group education sessions and individual home-based therapy in 80 patients after TKR. [249]

Mitchell and colleagues found no difference between group-based and home visit physiotherapy. [250] They did observe that the supervised home visit physiotherapy sessions were significantly more expensive than the hospital outpatient sessions (mean difference = −£136.5, 95% CI -£160 to -£113; p=0.001).

Mahomed and colleagues looked at home-based rehabilitation versus clinic-based physiotherapy. They randomized 234 patients, and empowered their study to detect an effect size of 0.5 with a type-I error of 5% and 95% power. [43] They found no significant difference between the two groups in their WOMAC or SF-36 scores.

The paper by Rajan and colleagues is titled 'No need for outpatient physiotherapy following Total Knee Arthroplasty: a randomized trial of 120 patients.' However the authors based the conclusion in their title on their study finding no difference post operative ROM, between patients given outpatient physiotherapy or no physiotherapy at all. While ROM is an important outcome of TKR surgery it is difficult to justify dismissing the need for physiotherapy based solely upon this single outcome measure. [251]

Current practice with regard to location of rehabilitation varies across the world, even within a local area in Australia. The evidence suggests that there is no advantage of one-to-one or group-based physiotherapy compared with home-based therapy using standard protocols. Moffet and colleagues demonstrated that increasing the intensity of prescribed exercise in a one-to-one setting produces a

small effect. It could be hypothesized that the intensity of the exercise may be the crucial factor rather than the setting or level of supervision and this has potential cost implications. [250]

2.3.2.3B   *TIMING OF REHABILITATION*

Current practice has been described already. It begins with the acute stage before an inpatient is discharged to either an inpatient or an outpatient rehabilitation programme. It lasts for a varying amount of time but often up to six weeks after the operation. Beyond six weeks, the rehabilitation is generally home-based and with little or no further supervision. The main question regarding timing is when it is most beneficial to begin the intensive phase of post-acute rehabilitation. Traditionally this intensive therapy continues from discharge from the acute hospital stay. It is hypothesized that delaying this phase for six to eight weeks may prove more beneficial.

Vuorenmaa and colleagues found that an exercise programme started at two months after TKR surgery had a positive effect on maximal walking speed and flexion strength but there was no difference in WOMAC improvement compared with a group receiving usual care with no additional guidance beyond the two-month point. [252]

Valtonen and colleagues randomized patients from a population-based sample who had had TKR surgery between four and eighteen months earlier. The intervention was a small class, a progressive aquatic resistance training class prescribed twice a week for 12 weeks. The intervention group showed superior quadriceps strength, walking speed, and stair-climbing ascent time compared with a control group receiving no intervention. There was no change in self-reported WOMAC score. [253]

These studies suggest that delaying the intensive phase of rehabilitation could be more beneficial than current practice in terms of functional gains, even if the patients do not report WOMAC as beneficial.

LaStayo and colleagues compared a traditional progressive resistance exercise programme with an eccentric resistance training programme in individuals one to four years after a TKR in a randomized trial. [254] Both groups trained for 30 minutes three times a week for 12 weeks. Both groups showed improved TUG and stair ascent at the end of training. The eccentric group also demonstrated additional gains in quadriceps strength, six-metre walking distance, and stair descent time. This study demonstrates that gains can be made in function, even well beyond the traditional time period of recovery from surgery.

The MARKER study was set up to examine the long-term effectiveness of starting intensive rehabilitation six weeks after TKR surgery rather than on discharge from hospital. The hypothesis is that patients are unlikely to be able to exercise at an intensity which would contribute significantly to muscle strength in the first few weeks after TKR. This could be because of swelling, pain, or even anaemia. [40]

Moffet and colleagues began an intensive exercise intervention two months post-surgery lasting for two months. They observed a significant difference in self-reported WOMAC score in favour of the intervention group, as described above. (treatment effect at four months post-op WOMAC total 7.9 (95% CI 2.7 – 13.1, p=0.004). [39]

The evidence suggests that a variety of exercise programmes, if started at a minimum of six weeks from surgery, can demonstrate a significant difference in outcomes favouring intervention compared with usual care. The objectively measured functional outcomes are more commonly improved over subjectively self-reported measures such as WOMAC.

2.3.2.3C   TYPES OF EXERCISE

In the survey of current Australian rehabilitation practice by Naylor and colleagues involving inpatient physiotherapy, the responses were multifactorial, with all citing more than one criterion. Gait retraining and exercise prescription were the only interventions almost universally cited. Nearly half

cited CPM as well as transfer practice, cryotherapy, and hydrotherapy. All respondents specified independent and safe mobility as a criterion for discharge, with the criteria of a specified range of motion only cited by 40% of respondents. [47]

Of the studies in a 2007 review, Frost and colleagues found no significant differences between groups when comparing a traditional exercise group with a functional exercise group.[32] They performed an RCT of unilateral arthroplasty patients performing home-based exercises in a functional exercise group versus a traditional exercise group. They assessed walking ability, leg extensor power, knee flexion, and pain for the year following the operation. The intervention for the 24 patients in the traditional exercise group was to continue with the exercises they had been taught in hospital. The 23 patients in the functional exercise group were taught three functional exercises aimed at increasing general activity and function. Interestingly, 50% of patients did not attend their one-year follow-up appointment. There were no significant differences between the groups in terms of their selected outcomes at any follow-up – 3, 6, and 12 months. However, post hoc analysis of required sample sizes ranged from 100 per group for a significant difference at 12 months in walking speed to 545 per group for pain during walking. The authors concluded that the differences between groups were likely to be of moderate size for any trial comparing exercise regimens of similar intensity. This trial was subject to a type II error because of its small size. This is common in clinical research because of lack of funds.

Moffet and colleagues looked at intensive versus standard rehabilitation. [39] They found a significant difference at four months in walking distance for the group receiving functional rehabilitation versus usual care (treatment effect 22.5 metres (95% CI 1.1, 44.0, p=0.040)) but this difference was not present at one year (treatment effect 26.4 metres (95% CI; -1.3, 54.0, p=0.061)).

Mockford and Beverland found a significant difference in range of motion favouring the intervention of outpatient physiotherapy versus controls receiving usual care.[255] However, this was work was an abstract for presentation at the Irish Orthopaedic Association. The article published in 2008 demonstrated a significant difference in range of motion at three months but none at one year. This

would have affected the outcome of Minns-Lowe's meta-analysis and may have changed her conclusions. [37]

Piva and colleagues ran a pilot study investigating the feasibility of applying a balance exercise programme in patients with a recent TKR and whether the addition of balance training could improve physical function. The results suggested balance may be beneficial but importantly, in this elderly population, the training was well tolerated and adherence was high. [256]

2.3.2.3d   *SUMMARY*

There is a large variation in the type, location, duration and intensity of rehabilitation offered to patients after a knee replacement. They are determined by surgeon preference, rehabilitation facility, funding source, geographical variations and patient preference. The literature supports patients receiving rehabilitation following TKR surgery although there is no strong evidence for one type of exercise over another. The literature is providing more evidence that most patients do not need individually supervised rehabilitation and that group or home based rehabilitation is equally effective but more cost efficient. There is a move towards investigating a delay to the start of intensive rehabilitation to beyond the six week mark. Should the delay prove beneficial, there could be a significant cost saving on the reduction of routine inpatient or intensive early outpatient therapy. Functional training may have benefits over traditional therapy.[39] This has not been shown in other studies[32] but the large effect size in Moffet's paper may have been due to the delay in starting intensive therapy until 2 months after surgery.

There have been limitations in the literature to date. There was little in the way of quality evidence prior to 2007 with few studies and those that available had small sample sizes or contained significant bias in their design. The topic seemed of little interest to the wider clinical and research community. The more recent studies of Mahomed et al, Ko et al and Moffet et al for instance demonstrate the elements of good study design limiting bias and they were appropriate powered to answer their research questions. [39, 43, 247]

While the recent literature has shown that the location of rehabilitation and level of supervision don't significantly affect patients' outcomes, there is clearly a need for further studies in the other important area of rehabilitation following TKR surgery. The optimal time to start intensive rehabilitation, the dose of exercise given and the type of exercises prescribed are all questions yet to be answered. The relationship between balance, proprioception, falls and functional outcome with any particular rehabilitation program is yet to be established in this population.

2.3.3   The Wii Balance Board And The Wii-Fit

The Wii-Fit has been used for training and rehabilitation but its various balance tests have received little scrutiny in the literature and none of them has been validated. In order for the Wii-Fit to be used by clinicians, notwithstanding the lack of regulatory approval, the measures need validation. The potential outcome measures that can be collected from the Wii and WBB include various non-validated balance measures from the Wii console, validated balance measures when the WBB is connected to a laptop running custom software, data on adherence including time spent on the machine exercising and which exercises were performed and to what level.

Validity and reliability

The Wii-Fit software runs on the Wii console and the Wii balance board (WBB) transmits centre of pressure data to the console via a Bluetooth signal. Clark and colleagues validated the WBB as a force platform capable of assessing standing balance in a healthy population.[180] Holmes and colleagues then validated its use in a population of patients with Parkinson's disease.[19] They both used the WBB connected to a laptop via Bluetooth, running a custom software program, making it impractical outside a laboratory setting. Pagnacco and colleagues challenged the findings of Clark, making a number of points mainly centred around the construction quality of the WBB. They stated that it should be treated as a toy rather than a piece of scientific equipment.[257] It is worth noting that Pagnacco is the designer of a posturographic device for assessing standing balance which costs considerably more than the WBB.

Negus and colleagues presented data on the Wii Stillness measure, part of the Wii-Fit software.[258] With a WBB connected to a Wii console running Wii-Fit and a laptop running the custom software used by Clark, the Wii Stillness test was tested for validity and reliability in healthy volunteers. The software measured the sum of the maximum amplitude or range in the x and the y-axis, which was then transformed into a percentage score. This measure was valid and reliable in three of the four

stances tested. As regards the unilateral stance with eyes closed, too many of the subjects were unable to maintain the stance for the full 30 seconds. When a subject failed a trial in this way, there was no centre of pressure data from the Wii to analyse.

Jorgensen and colleagues have the only published data on the Stillness test.[259] They found them to be reliable in a test-retest format in community dwelling older adults (ICC 0.86 (95% CI 0.74 – 0.93)). They also found that the Stillness test had moderate to excellent concurrent validity compared with selected force plate measures. (i.e. CoP velocity & CoP confidence ellipse). This was in contrast to the concurrent validity for the selected measures of Negus and colleagues.[258] The measures used to test for concurrent validity need to be investigated further.

Wikstrom found that there were poor correlations between all 96 comparisons of Wii-Fit activity outcomes and established balance outcomes (r<0.50). He also found intraclass correlation coefficients from 0.39 to 0.80 for intrasession reliability and intersession reliability (ICC = 0.29-0.74) to be poor for 11 out of 12 scores.[260]

Feasibility

Several authors have investigated the feasibility of using the Wii in an older population. Bieryla and Dold randomized 12 subjects into a group training on the Wii and a group continuing their usual activities.[18] The average age was 81.5 years (s.d. 5.5) and they were all healthy subjects living independently in a senior community facility. The intervention group attended one-on-one supervised training sessions three times a week, for three weeks, with each session lasting 30 minutes. The chosen exercises were a combination of yoga, aerobics, and balance games. There was no variation of the exercise choice between or within subjects. The control group continued with their usual daily activities. All subjects had balance assessments pre-intervention and then one week and one month after the end of the intervention. The Wii group had significantly increased Berg Balance Scale scores

at the one-month post intervention assessment. The Wii as a tool was generally well received by the intervention group and many expressed enjoyment during the study.

Acceptability

Laver and colleagues examined the acceptability of the Wii among an older population using a discrete choice experiment design.[261] They included 21 patients with a mean age of 85.4 years (s.d. 4.7) from a geriatric rehabilitation hospital. The patients were questioned about their preferences for different therapy options by means of six choice sets. These included questions about using the Wii or traditional gym-based physiotherapy after they were shown pictures of elderly people using both. The other sets were related to intensity, cost and recovery. They were then prescribed Wii-Fit-based therapy for 25 minutes per day five days a week for the duration of their stay. They were assessed again one month after discharge from hospital.

The mode of therapy offered did not significantly influence the subjects' choice of therapy programme at the initial assessment. The patients showed an aversion to longer and more demanding therapy sessions and favoured a longer recovery time. At the post-intervention assessment, the subjects preferred the therapy which did not use the Wii. The interviewer administering the discrete choice experiments reported that the subjects often commented "that they perceived that the conventional therapy was more effective than therapy using the Wii-Fit". The authors concluded that this evidence is at odds with the widely held belief that the elderly would respond well and value the Wii once they had been exposed to it. They raise some interesting points about the nature of the games not being specific to this age group and that games more specifically targeted at the elderly may have a greater effect.[16, 262, 263] They also acknowledge that were there more evidence that this intervention worked, the subjects might have been more willing to embrace it.

Williams and colleagues studied the acceptability of using the Wii-Fit in 15 community dwelling subjects over the age of 70 who had fallen at least once in the previous 12 months.[264] They were compared with a control group recruited from a local falls clinic. However, of the 13 potential recruits, 7 were excluded or dropped out. The subjects were assessed on a number of balance scales, an attitude to falls intervention scale, and monitored attendance. The Wii group was also interviewed to assess the acceptability of the Wii-Fit. The intervention lasted 12 weeks during which time the Wii group attended one-on-one supervised sessions twice a week. The control group attended a local falls group. The average age was 76.6 years (s.d. 5). Attendance was good and 80% of the participants attended 75% or more of the sessions. There was no significant difference in the attitude scale between the two groups. Although a statistically significant (p=0.02) improvement in the Berg Balance Score was recorded in the four-week assessment of the Wii group, but not the control group, it was not present at week 12. The authors do not provide any comparative analysis between the two groups as the control group of six subjects was deemed 'frailer' at baseline.

All the intervention group found the Wii-Fit to be enjoyable and acceptable and 92% expressed a strong desire to exercise with the Wii-Fit in the future with no preference for exercising on their own/with company or supervised/unsupervised. The main finding of this study is that the Wii appears to be acceptable and enjoyable to subjects in their seventies and may be equivalent to a falls group as an intervention. The subjects also indicated that they would be willing to exercise with the Wii in their own homes.

Prosperini and colleagues trialed Wii training at home in a pilot study of patients with multiple sclerosis.[265] The important finding in relation to any study of home based treatment was that there were five adverse events of joint pain but no-one had to retire from the study and there were no serious adverse events.

Clinical trials

Nitz and colleagues investigated eight healthy women in a pilot study of convenience.[266] All subjects needed to have access to a Wii-Fit. Their mean age was 46.6 years (s.d. 9.9) and they were healthy. The intervention involved two sessions lasting 30 minutes every week for 10 weeks. The exercises were of the subjects' choice and they kept a log. The overall adherence measured by sessions completed was 70%. The subjects showed an increase in function (TUG $p=0.09$), strength ($p<0.02$), and unilateral stance balance ($p<0.05$) outcomes. More importantly, the authors used the results of this small study to calculate a sample size to power a study at 90% with a significance of 0.05 to demonstrate clinical improvement in TUG and six-minute walk test. They calculated a total sample size of 61, allowing for a 15% drop-out rate.

Baltaci and colleagues randomized 30 men with anterior cruciate ligament (ACL) reconstructions into a Wii group and a conventional ACL rehab group.[267] The Wii group exercised at the physiotherapist centre, spending 15 minutes in each of four games over a one-hour session, three times a week. The control group had the same number of sessions following a conventional protocol. The intervention period lasted three months following surgery. The subjects were assessed by evaluations of coordination, proprioception, response time, and modified star excursion balance test. The authors found no significant difference between the two groups in any of the outcomes assessed at the first, eighth, or twelfth week following surgery. This study suggests that use of the Wii-Fit in this patient population was equivalent to conventional physiotherapy despite the small sample size. The power analysis gave a sample size of 15 per group and was calculated to find a 1.1cm difference in a modified star excursion balance test with power of 80% and significance of 0.05. This suggests a large effect size and there is no mention of the effect size of the other outcomes. The small sample risks a type II error in these other outcomes.

Fung and colleagues carried out a randomized pilot study in patients who had had TKR surgery.[21] They recruited 50 subjects with an average age of 68 years (s.d.11), a recruitment rate of 58%. All subjects

were outpatients of a rehabilitation hospital who had had a unilateral TKR. Each group attended the rehabilitation centre twice a week for a one-hour conventional session followed by 15 minutes of lower extremity exercises in the control group and 15 minutes of Wii-Fit exercises in the intervention group. A variety of objective, functional, and subjective outcomes were measured on the first visit and then every two weeks until discharge. There was no difference between the groups in any of the outcomes. The effect size for the self-perceived lower extremity function scale was 0.50. The authors calculated that a total sample size of 136 would be needed to demonstrate a significant difference although they do not mention the power or significance level of this calculation. The authors also state that during the satisfaction survey, although both groups reported enjoyment, the Wii group often stated that the Wii was fun and that they were glad they were in the Wii group.

A good summary of 'Wii-search' is provided by Goble and colleagues.[268] They conclude that , 'Multiple studies now show improvement in balance control following reasonable periods of training with the Wii Fit games and custom design applications'. However, 'Greater emphasis must still be placed on implementing more powerful, randomized control designs with larger sample populations to test questions of interest'.

2.3.4   DISCUSSION

This literature review revealed an increasing level of interest in rehabilitation following TKR surgery since the systematic review of Minns-Lowe and colleagues in 2007 which demonstrated a paucity of high-quality investigation to that point. This development mirrors the increasing interest of the surgical community in patient-related outcome measures. A number of studies have also considered the Wii balance board and the Wii-Fit in musculoskeletal and elderly populations. The WBB has been found to be valid and reliable as a force platform but there is little work on the validation of the Wii-Fit software or any of its measures. No high-quality, randomized trial adequately powered to compare the Wii-Fit with routine care after a total knee replacement is as yet available.

2.3.4.1 CHOICE OF OUTCOMES

It is clear from the literature that, in order to assess fully the outcome of a TKR, the clinician must use a variety of outcome measures to cover general health and joint specific function and pathology. Subjective, objective, and functional measures are needed to give a complete picture of the patient and their knee.

The WOMAC has been validated for use in the TKR population and has a minimum clinical difference which gives a practical sample size for a clinical study. The SF-12 covers the patient's general health status. The Hospital Anxiety and Depression Score can be used in addition to the locus of control and the SF-12 mental subscale to quantify patients' mental state. A satisfaction score should be used although there is little consensus over the ideal validated score.

It must be factored into the analysis that studies have shown that preoperative self-report questionnaires consistently predict postoperative questionnaire scores. [13, 67, 228, 269, 270]

Objective assessment of the American Knee Society Score gives information on the range of motion and alignment of the knee. It is a consistent finding that the strength, especially in terms of the quadriceps, of both the operated and non-operated leg affects outcome after a TKR. Because of time and money constraints, as well as applicability in the real world, a hand-held dynamometer can provide adequate information on isometric force production. The UG test can be easily tested in a clinic scenario to provide a functional assessment. Finally, the majority of training on the Wii-Fit involves an element of balance work so standing balance on both legs and one leg is assessed.

There is no suitable questionnaire relating to abilities with computer technology and the internet in this age group. A questionnaire was designed to produce the required data and it was called the **Computers, Internet and Gaming in Rehabilitation (CIGAR – see chapter 3).** Finally, adherence to the prescribed exercises should be recorded to inform the results.

2.3.4.2 *Location, Timing And Type Of Rehabilitation Intervention*

The evidence is becoming clearer that rehabilitation after a TKR does not need to be supervised and does not need to happen in an outpatient setting. If exercise at home is appropriately directed, there should be no negative effects.

Current practice in Australia is varied but tends more than many other countries to apply supervised rehabilitation through the first six weeks after a TKR. However, gains are likely to be limited in this initial period because of postoperative oedema, pain and low haemoglobin levels which can persist for the first six weeks after surgery. This reinforces the need to investigate the efficacy of a cost-effective method of rehabilitation beyond six weeks to achieve physiological and lasting change. As this coincides with a surgical outpatient appointment in the majority of cases, it also allows for clinical assessment of patient suitability for any intensity increase.

The effect of any physical intervention is related to the amount of exercise. Although intensity can be increased, it can be difficult for the patient and poorly tolerated in an elderly population recovering from an operation. Therefore, an intervention with low-to-moderate intensity, performed daily, is preferable. To enable daily intervention, this needs to be a home-based intervention to control costs.

There is no clear evidence that any one type of intervention is more effective than another. A combination of strength, balance, and aerobic and functional exercises would seem to be the ideal.

2.3.4.3 *The Wii*

The WBB has been validated as a force platform for assessing standing balance, but only when connected to a laptop running software which measures its Bluetooth transmitted CoP data.[180]

As training tools in the musculoskeletal population, the Wii balance board and Wii-Fit software have shown promise in a number of clinical trials although most are pilot studies. The Wii is designed as a games console for the home and its relatively low cost makes it ideal for the study of exercise at home.

Some of the exercises in the Wii-Fit software are similar to standard rehabilitation exercises which allow a protocol to be developed which mimics the basics of routine care. One of the advantages of the Wii-Fit is that there are many different exercises offering variety and an ability to progress as the subject improves their performance. A combination of yoga, strength, and aerobic and balance exercises should offer the necessary variety.

The intensity of the exercises can be tailored to the patient, allowing for an increase in difficulty and intensity as the patient progresses with their rehab. This is likely to maximize the gains from the time spent exercising without intimidating those patients at an early stage in the protocol. It is also likely to increase their boredom threshold as they are continually challenged to perform new and more difficult exercises.

The Wii console records the amount of time spent by each subject in performing the exercises. This enables data collection on adherence during the intervention period, in addition to a self-reported patient diary.

CHAPTER 3 - METHODS

3.1 Objectives Of The Thesis

1. Is the Wii-Fit Stillness score a valid and reliable measure of standing balance?

2. Which preoperative factors affect outcomes in patients with a TKR, with a particular focus on balance?

3. Does using the Wii-Fit improve outcomes after TKR?

3.2 RECRUITMENT OF HEALTHY SUBJECTS FOR WII VALIDATION STUDY

(Chapter 4.1)

All healthy subjects were recruited via a recruitment poster at the Kolling Institute of Medical Research, North Shore Hospital campus, St Leonard's, NSW or via the staff and students of the Sydney Orthopaedic Research Institute (S.O.R.I.), Chatswood, NSW. They were screened via interview for co-morbidities which could affect balance such as inner-ear disease or Parkinson's disease.

All subjects were fully informed of the requirements of the test including the need for repeat measurements at a later date. They all gave verbal consent, with study approval and requirements according to the local ethics committee (Harbour Human Research Ethics Committee) and the University of Sydney Ethics Committee.

3.3 Recruitment Of Patients for Wii Validation Study and Randomised Controlled Trial

(Chapters 4.2-3 & 5.1-2)

All patients were recruited from the rooms of Dr David Parker, Dr Myles Coolican and Dr Brett Fritsch in Chatswood, NSW. They all had a routine clinical appointment with one of the surgeons after a GP referral. If they were then put onto a waiting list for a primary TKR, they were directed to Sydney Orthopaedic Research Institute (S.O.R.I.) next door, which was usual protocol. They were then screened for suitability against the inclusion and exclusion criteria. All subjects gave oral and written consent, with study approval and requirements according to the local ethics committee (Harbour Human Research Ethics Committee) and the University of Sydney Ethics Committee.

3.4 Clinical Assessment And Data Collection

In Chapter 4.1, basic demographic data not limited to age, gender and BMI were collected from the subjects at recruitment. During the study, balance data were collected.

Chapters 4.2-3 and 5.1-2 provide more detailed demographic data collected from the patients including co-morbidities and medications. The other baseline measurements included objectively measured outcomes such as muscle strength and range of motion, multiple patient-reported outcome measures such as the Western Ontario and McMaster (WOMAC) index of osteoarthritis, and other functional measurements such as the 'Timed Up-and-Go'.

Patients were seen either at one timepoint (cross-sectional) or at a number of timepoints from the baseline preoperative assessment and then over the course of one year after the operation (longitudinal).

3.5 PATIENT-REPORTED OUTCOME MEASURES

Patient-reported outcome measure (PROMs) data were collected in the form of questionnaires. A number of collection methods were used. Whenever possible, these data were collected via a website designed for PROMs, www.myclinicaloutcomes.co.uk.It is used by a number of NHS trusts in the UK for patient data collection and so has high levels of data security and confidentiality appropriate to clinical data. It has been shown to be comparable to paper-based collection of outcome scores. [271] The website was not always used, as some patients did not have access to email or the internet and some had just forgotten or not managed to complete their questionnaires by the time they attended the research institute for their assessment. In these circumstances, they completed the PROMs on a computer in the institute using the 'Socrates' database system used by the institute. This is an Orthopaedic outcome specific database preprogrammed with many of the commonly used outcome scores.[272] If they were unable to use a computer, a research staff member asked them the questions and filled in the answers for them. A small number of scores were completed on paper either because of technical issues or because people needed the forms to be posted to them at home. For the VAS pain and satisfaction scores, there were no computer options so they were all completed on paper and manually added to a database.

3.5.1   WOMAC

The Western Ontario and McMaster (WOMAC) index of osteoarthritis is a questionnaire which was developed for osteoarthritis but has been validated for use in arthroplasty of the knee.[55, 62] It can be scored as three subscales and a total score. The questionnaire can be completed on a Likert scale with five options scored from zero to four as follows: none, mild, moderate, severe, extreme. Each question can also be answered using a visual analogue score. The Likert scale was used and the five questions in the pain subscale gave it a range of scores from 0 to 20, the two questions in the stiffness

subscale scores from 0 to 8, and the 17 questions in the functional subscale from 0 to 68, giving a total score of 96, with 96 being the worst score.

The score is often normalized to a range of 0 to 100 to allow easier comparison with other scores in the literature.

The WOMAC data were collected on a computer when possible using www.myclinicaloutcomes.co.uk or the Socrates database. Otherwise they were completed on paper.

See Chapter 2.3.2.1a and appendix (ii)

3.5.2   OXFORD KNEE SCORE

The Oxford Knee Score (OKS) is a patient-administered, disease- and joint-specific questionnaire developed in Oxford in the UK.[85] It has 12 questions, each with five categories scored from zero to four, giving a range of scores from 0 to 48 with a score of 48 indicating the best function.

The score can be normalized to a scale of 0 to 100 but this is not recommended because of confusion over previous changes to the scoring method. [86]

The OKS data were collected on a computer when possible using www.myclinicaloutcomes.co.uk or the Socrates database. Otherwise they were completed on paper.

See Chapter 2.3.2.1b and appendix (iii)

3.5.3 SF-12

The SF-12 is a general health questionnaire. It is recommended that a general health questionnaire is used in conjunction with a disease-specific questionnaire such as WOMAC for assessing the outcomes of TKR surgery.[97] It is a questionnaire consisting of 12 questions. The score is divided into a physical component and a mental component. The scoring algorithm is complicated and computed automatically in a preprogrammed database. It is based on a population mean score of 50 with a standard deviation of 10. Zero is the worst score and 100 the best in both the mental and physical domains.

The SF-12 data were collected on a computer when possible using www.myclinicaloutcomes.co.uk or the Socrates database. Otherwise, they were completed on paper.

See Chapter 2.3.2.1c and appendix (iv)

3.5.4 VISUAL ANALOGUE SCALE FOR PAIN

The Visual Analogue Scale (VAS) for pain is a simple one-page score using a horizontal line of a set length with '0 – No pain' marked at one end and '10 – Extreme pain' at the other. Patients were asked a standard question –
'How much pain are you feeling in the operated knee on an average day within the last week.'

The VAS pain data were collected on paper for all patients.

See Chapter 2.3.2.1d and appendix (v)

3.5.5 SATISFACTION ASSESSMENT

There was no validated satisfaction questionnaire appropriate for the patient group in the TKR-POWER study. Therefore satisfaction was assessed with three self-reported questions.

1) How satisfied are you with your operated knee?

   (four-option Likert score – very dissatisfied, dissatisfied, satisfied, very satisfied)

2) How satisfied are you with your rehabilitation?

   (four-option Likert score – very dissatisfied, dissatisfied, satisfied, very satisfied)

3) Does the operated knee feel normal?

   (two options – Yes/No)

The Satisfaction assessment data were collected on paper for all patients.

See Chapter 2.3.2.1g and appendix (vi)

3.5.6 CIGAR - COMPUTER, INTERNET AND GAMING ACCESS FOR REHABILITATION

The Computer, Internet and Gaming Access for Rehabilitation (CIGAR) questionnaire was devised for this study. It is not validated and has not been used before. It has eight question stems and the final two stems are multiple choice answers.

The CIGAR questionnaire has not been tested or validated and its reliability, internal consistency, and stability over time are not clear. It was designed purely as a descriptive tool for the population in the following studies and not for comparison with other populations or detailed analysis.

The questionnaire does not have an overall score but generates descriptive data on the use and levels of comfort associated with using various forms of computer technology.

The CIGAR data were collected on a computer when possible using www.myclinicaloutcomes.co.uk. Otherwise, they were completed on paper.

See Chapter 2.3.2.1e and appendix (vii)

3.6 Objective measures

3.61 AMERICAN KNEE SOCIETY CLINICAL RATING SCALE (AKSCRS)

Objective data from the American Knee Society Clinical Rating Scale (AKCRS) were collected for each patient. However, in all the data analyses, only the range of motion components was used.

The centre of a goniometer was positioned over the lateral knee joint space with the arms aligned with the lateral midline of the femur using the greater trochanter as a reference and the lateral midline of the fibula using the head of the fibula and lateral malleolus as a reference (see Figure 4). Knee extension was measured with the leg flat on the bed while the patient was supine. Knee flexion was measured as an active movement while the patient was supine.[273]

Figure 4 – Measuring knee flexion using a goniometer

3.62   MUSCLE STRENGTH

Maximal isometric knee flexion and extension strength are assessed with a handheld dynamometer

(see Figure 5). The subject is seated for quadriceps with the knee sitting between 2 and 5cm from the

edge of the bed and flexed to 90 degrees. Hamstrings are tested with the subject prone and the knee

flexed to 60 degrees. The maximum from three attempts is used for the analysis. [274]

The results are in kilograms.

Figure 5 – Testing knee extension (quadriceps) strength using a handheld dynamometer

3.7 FUNCTIONAL MEASURES

3.7.1 TUG

This is the shortened version of the Timed Up-and-Go test. It is the time it takes the participant to get up from a chair, walk 2.44m, turn around a cone, and return to the chair to sit down. The subjects had one practice attempt and then two recorded attempts, with the mean score being used for analysis according to standard methods.[234]

The result is recorded in seconds to one-tenth of a second.

3.7.2 BALANCE

In the validation studies, the Biodex Balance SD (Biodex Medical Systems 1999) platform was used (see Figure 6). This is an integrated force platform and computer. It assesses balance with options to customize the type and duration of test. It also has the capacity to unlock the base, changing the stability of the platform in which the subject is standing. The data are stored in the built-in computer and raw data can be exported as a .csv file for analysis on a standard computer.

The Wii and Wii balance board were used for balance testing and training (see Figure 7). In the validation studies for balance testing, the Wii balance board was connected to either the Wii console or a laptop for different tests. When connected to the Wii console, data were collected by the Wii-Fit software Stillness program. The results were given as a percentage. When connected to the laptop, the data were collected via the Labview software. The Labview software produces a series of x, y coordinates, at a specified frequency, representing the centre of pressure (CoP) measured. Our data were sampled at 40 Hz and filtered using an eighth order Butterworth filter with a low-pass cut-off frequency of 12 Hz as per Clark and colleagues.[180] The Stillness test in the Wii-Fit software generates a

percentage score. The data filtering and cut-off levels used by the Wii-Fit software were not available

to the author and are not readily supplied by Nintendo for commercial reasons.

Figure 6 – The Biodex SD Balance platform with integrated screen

Figure 7 – Nintendo Wii console with handheld controller

Figure 8 – A patient exercising on the Wii with balance board

For balance training, the patients in the intervention group all had a Wii console running Wii-Fit software and balance board at home (see Figure 8). They had it connected to their television in a convenient room. The Wii console recorded data on their usage, levels of attainment and balance testing scores throughout the intervention period.

3.8   ADHERENCE

Adherence was recorded by the patients in self-reported diaries and by the Wii consoles which saved data on each exercise as it was completed. The Wii group received a separate diary for the Wii exercises. Both groups received a set of instructions and a diary for the routine 'control' exercises. The diaries produced data on the number of each exercise type performed each day throughout the intervention period.

The Wii consoles recorded the time spent exercising each day and the category of exercise (yoga, strength, balance, or aerobic). It also recorded high scores and when increasing levels of difficulty had been attained.

For the Wii group diaries - see appendix (viii)

For the control exercise instructions and diary - see appendix (ix)

3.9   STATISTICAL ANALYSIS

Statistical analyses were performed using SPSS version 21 or 22 (SPSS Chicago, IL) and SAS (SAS, Cary, NC). Variables were presented either as numbers (%) or means with standard deviations. For non-parametric datasets medians with interquartile ranges were presented

3.9.1   CONTINUOUS DATA

Normality was tested by assessing both the Shapiro-Wilk test and the skewness and kurtosis of the data.

Comparison of continuous data between two groups was made with the Student's t-test – independent and paired where appropriate.

Comparison of the main outcome measure in chapter 5.1 was made with the ANOVA and ANCOVA tests.

3.9.2   CATEGORICAL DATA

Categorical data were compared using the Chi-squared test or, if the frequency of occurrence was small, the Fisher's exact test.

3.9.3   STATISTICS FOR ASSESSMENT OF VALIDITY AND RELIABILITY

A two-way, random-effects, single measure (median of the three trials) intraclass correlation coefficients ($ICC_{(2,1)}$) model was used to assess reliability. Point estimates of the ICCs were interpreted as follows: excellent (0.75-1), modest (0.4-0.74), or poor (0-0.39).[275]

Comparison between two measuring devices was analysed by plotting the difference in percentage score between the two devices against the mean results. This resulted in a Bland-Altman plot (also

known as a Tukey mean difference plot). The upper and lower limits are 95% limits of agreement and are used to evaluate the smallest detectable difference.

Standard error of measurement (SEM) and minimum detectable change (MDC) values were calculated to assess the concurrent validity between device outputs, as well as the within-device test-retest reliability and measurement error over two testing sessions. The MDC, which is based on the reliable change index score, was calculated using the equations reported previously by Jacobson and Truax.[276] It is expressed as the percentage test-retest change in score, required to find a significant difference at an alpha level of 0.05 based on the Day 1 mean value.

3.9.4   REGRESSION ANALYSIS

The correlation between two continuous variables was analysed by simple linear regression to obtain a Pearson's correlation coefficient (r) (see chapter 4.2).

Multiple regression models were constructed for covariates considered clinically significant from the Pearson's correlation coefficient values and with p values <0.2 (see chapter 4.2). Checks were made for outliers, linearity, homoscedasticity, independence of errors, and normality of residuals.

Given that the TKR-POWER trial (see chapter 5.1) collected data at multiple timepoints (preoperative, week 6, week 18 and 1 year assessments) general Equation Estimate models were constructed.

CHAPTER 4 – Pilot Studies

4.1   VALIDITY AND RELIABILITY OF THE WII STILLNESS SCORE IN HEALTHY SUBJECTS

This study has been submitted to 'Gait and Posture' and is currently under review.

4.1.1   INTRODUCTION

Humans need to maintain and control their balance over a small base of support to be able to stand and walk. This balance control has many inputs including from the brain, the musculoskeletal system, vision, and the inner ear. Balance control can be affected by pathologies, medications, and the general functional decline of ageing.[154]

Postural instability increases with age and is an independent risk factor for falls.[161-163] The risk of falls increases with age, and approximately one in three people over the age of 65 years old and living in the community falls at least once over the course of a year. This number increases if the individual is living in an institution.[155, 156]

Poor balance also affects the ability to mobilize and carry out the activities of daily living.[277] This is even worse in patients with osteoarthritis of the knee whose balance is worse than that of their age-matched controls.[164, 278-281]

Interest in the assessment and management of balance deficits has led to the development of a number of balance measuring instruments. A subjective assessment such as the Berg Balance Scale is commonly used. This has been shown to provide valuable information but suffers from ceiling effects and limited precision in detecting small changes in performance.[167, 168] Objective measurements can be made from centre of pressure (COP) measurements on a force platform. These are considered the gold standard of balance measurement and have identified important outcome measures too subtle to detect with a subjective scale. [174 175] (outcome measures discussed in detail in chapters 2 and 3)

However, the use of force platforms to assess balance is limited by cost, the need for technical expertise and the lack of transportability.

The re-education of balance functions is included in most lower limb rehabilitation protocols but is still relatively neglected compared with strength and range of motion exercises. [195] The focus of many rehabilitation trials is still range of motion and strength.[40] The review of Minns-Lowe and colleagues on rehabilitation for patients who had a TKR for knee OA did not include a specific balance search term. [37]

The ability to assess standing balance can provide important information which ranges from predicting falls in the elderly to assessing postoperative recovery after orthopaedic surgery. [193-196] Significant improvements in pain, function, and proprioception have been reported following TKR surgery.[202, 203] In these patients, the proprioception of the operated leg improves compared with the contralateral leg also affected by osteoarthritis.[36, 201, 282, 283]

However, there have been few studies on the risk of falls after TKR surgery. Swinkels and colleagues found a 24% fall rate in the last preoperative quarter in a cohort of patients undergoing a primary TKR. Forty-six per cent of preoperative fallers continued to fall postoperatively, giving them an eightfold increase in the risk of postoperative falling. [208] Although across the whole cohort, patients improved their function and balance confidence up to one year after surgery, the preoperative fallers

did not show the improvements in function or balance confidence between 3, 9 and 12 months of the non-fallers. They found that the odds of a pre-operative faller becoming a faller in the first post-operative year are almost eight times those of a non-faller (adjusted OR 7.75, 95% CI 1.7-35.7, p=0.008).

Levinger and colleagues found that at four months post-TKR surgery, patients had decreased knee extension strength and proprioception when tested as part of a falls risk assessment compared to age-matched controls. [194] The most effective fall prevention strategies target high-risk populations or individuals.[284, 285]

Given these findings of fall risk after TKR it would seem reasonable to investigate the role of balance and include more balance-related exercises in postoperative rehabilitation exercise regimens.

It is important to clarify the instrument used to measure balance in any assessment and establish the variable it is trying to measure. Standing balance measures can look at symmetry or steadiness. Symmetry relates to the ability to distribute weight evenly between the two feet whereas steadiness is the ability to keep the body as still as possible.[176]

A force platform measures the centre of pressure (CoP) and describes the coordinate on the force platform in the x- and y-axis of the movement around the centre of gravity of the subject. The centre of gravity is located above the CoP point in the z-axis. During movements, the CoP moves in response to motion of the centre of gravity. This CoP measurement can be used to produce a value to describe the subject's balance. There are many valid measures (see Table 4) such as the length of the path traced by the CoP in a set time period or the 'spread' (amplitude) of the coordinates or variability (standard deviation) of this spread. They can be used to measure balance while the subject maintains a static pose, in response to a stimulus designed to disrupt their balance (a perturbation) or during an active balance task such as shifting their centre of gravity in order to move their CoP as shown on a

screen to different locations to hit certain targets. These methods all measure subtly different aspects of balance.

The Biodex balance SD (BIODEX, USA) (described in detail in chapter 3.7.2) is a force platform which has a standing platform which can be unlocked from a static position into an unlocked unstable state. It has an inbuilt console which provides instruction and feedback as well as collecting and storing balance data at the end of each trial. It is well recognized as a force platform for data collection and balance training.[286, 287]

The Nintendo Wii balance board (WBB) (Nintendo, Kyoto, Japan) is a reliable and valid tool for assessing standing balance.[180] It is a gaming system designed for use in the home and consists of a small console plugged into the TV and a wireless force platform the size of a large set of bathroom scales. As it is a relatively cheap and portable system, this could prove very useful for clinicians and rehabilitation specialists. [19] However, the software used in many of the studies on its use is custom designed and not universally available.[19, 180, 288-290] The WBB is designed to work with the Wii console and Wii-Fit software, which contain both balance tests and training games. It has already been integrated into rehabilitation protocols of neurological patients with balance defects in one study and has been shown to have the potential to enhance motivation levels in another. [17, 180]

The Wii-Fit software contains a number of tests which give the user a variety of scores such as a percentage score or a 'Wii-Fit Age'. These scores suggest a quantification of the user's standing balance, with the option to chart improvements or deteriorations in that ability, but it is not known how the software generates these scores from the WBB raw data. A number of clinicians have started to investigate patients' standing balance based on these scores. [291 18] Until they are validated, any data from the Wii-Fit software should be used with caution in a clinical setting.

The WBB has been shown to have the potential to provide clinicians with an inexpensive, portable, and valid alternative to a laboratory grade force platform to assess standing balance. This has been

demonstrated by Clark and colleagues, who validated it against a laboratory grade force platform in healthy volunteers[180] and Holmes and colleagues who investigated its use in patients with Parkinson's disease.[19]

However, the use of the WBB as a force platform to measure standing balance has been questioned on technical grounds. The WBB demonstrates more electronic noise than clinically approved platforms as the cables are not shielded and the electronics are not designed for noise minimization and the channels are not simultaneously sampled. Pagnacco and colleagues found the precision of the CoP coordinates to be approximately 0.5mm, which is much less than in other devices and force platforms.[257]

The aim of this study was to investigate the reliability and validity of the Stillness score generated by a Wii console running the Wii-Fit software in terms of assessing standing balance. The Stillness scores from the Wii-Fit were compared with data from a WBB connected to a laptop running custom software during a variety of balance tests performed on a WBB.

4.1.2 METHODS

Reliability

The first step was to test the reliability of the Wii-Fit software using a test-retest method.

Participants – Reliability

Twenty-four injury-free individuals were tested on two occasions within four weeks of each other (mean 9.5 days (sd 5.5)) and at least 24 hours apart. There were 8 males and 16 females with a mean age of 28.4 years (s.d. 9.1) and mean BMI of 23.3 (s.d. 4.4). No participant reported any major symptoms, history of pathology, or recent injury in the back or lower limb. There was also no history of medication use or neurological disease history which could have influenced standing balance.

The local Human Research Ethics Committee approved the study and all participants provided informed consent.

Procedures - Reliability

On each of the two assessment dates, participants performed standing balance tasks on a WBB placed on the floor 1.5 metres from a wall. A cross was marked on the wall at eye level to provide a consistent focal point. The real-time output and results were displayed on a screen in front of the assessor but were not visible to the subjects. All sound was muted so they received no visual or auditory feedback during testing.

In each assessment, the subjects were instructed to take up one of the four stances and attempt to complete three trials for each. This was repeated for all four stances. These stances were chosen to mirror the stances studied by Clark and colleagues when they validated the Wii balance board as a force platform.[180] The WBB was connected via its Bluetooth connection to either the Wii console (referred to from now on as 'Wii') measuring the Stillness score or a laptop running custom software (referred to as 'Laptop") to measure CoP coordinates. The Stillness test is a measure of CoP movement over 30 seconds which results in a percentage score.

Once the participants had completed all the trials in each of the four stances and the data were collected on one device, the WBB was connected to the second data collection device. They then repeated their trials in the same four stances. The order of stances and testing devices was randomly assigned for each participant, but remained consistent between testing sessions.

The four stances:

DSEO   Double-leg Stance with Eyes Open and feet apart.

DSEC   Double-leg Stance with Eyes Closed and feet together

USEO   Unilateral-leg Stance with Eyes Open

USEC   Unilateral-leg Stance with Eyes Closed

The subjects were instructed to take a comfortable stance for the DSEO and the distance between their heels was measured and kept consistent between tests and sessions for that subject. All unilateral stances were performed with the subject standing on their dominant leg.

During each trial, the participants were instructed to keep their hands by their sides and to remain as still as possible for the duration of each trial, which lasted 30 seconds. If their support foot shifted position or their non-support foot touched the floor then the trial ceased and was recorded as unsuccessful.

Three successful trials were attempted for each stance and testing ceased if there were three unsuccessful attempts at any particular stance. There was a 15-second rest between trials and 60 seconds between stances and devices.

Validity

To work out what the Stillness test was actually measuring, the ideal testing scenario would have been to have one WBB transmitting identical CoP data simultaneously to a laptop and a Wii console. This would have given a percentage score from the Stillness score to correlate with the raw CoP data.

As each individual WBB only allows its output signal to be received by one Bluetooth linked device, the data of each separate trial had to be collected simultaneously by two different devices. To achieve this, two WBBs were stacked on top of each other, transmitting a separate but simultaneous Bluetooth signal for each trial. One was linked to a Wii console with its data analysed by the Wii-Fit software. The other was linked to a laptop running the custom software, producing CoP coordinate data. This configuration was tested for accuracy and then the simultaneous balance data were analysed for correlation.

125

Procedures - Validity

Testing Configuration of WBBS

The accuracy of stacking the WBB on a force platform has been shown to be accurate,[291] but the stacked WBB configuration was still tested. Two WBB were placed on top of each other and positioned on the centre of the Biodex SD force platform (BIODEX, USA). First, the uppermost WBB was paired via Bluetooth with the laptop running custom software. A single subject then performed multiple standing trials. The data output was collected simultaneously on the laptop and also by the Biodex SD force platform. The lower of the two WBBs was then paired via Bluetooth with the laptop and the subject repeated multiple standing trials with simultaneous data collection by the Biodex SD and the lower WBB. The output of the WBBs in each position was correlated with the Biodex SD data. The Biodex SD was not used for the remainder of the reliability or validity testing.

Deriving Stillness formula

The two WBBs were stacked one on the other and placed on a solid surface. One was connected to a Wii console running Wii-Fit software and the other to a laptop running the custom software. A single subject performed 50 trials (DSEO) in 5 sets of 10 with the subject varying the variability and direction of their CoP position in a methodical manner to replicate as many potential standing balance situations as possible. The simultaneous data sets from the custom software and the Wii-Fit Stillness test were collected from these 50 trials (see Table 3).

Table 3 - 50 trials of Stillness correlated with simultaneous CoP data from custom software

Trials	Variation
0 – 10	Low sway amplitude and velocity in any direction
11 – 20	High ML sway amplitude – low velocity
21 – 30	High AP sway amplitude – low velocity
31 – 40	High sway amplitude - all directions – low velocity
41 - 50	High sway amplitude - all directions – high velocity

AP – Antero-posterior direction

ML – Medio-lateral direction

Each trial produced two sets of data. The WBB that was paired to the Wii console generated a percentage Stillness score. The WBB paired to the laptop generated a set of x, y coordinates representing the CoP was generated.

The laptop data sets were analysed with formulae of known valid measures of standing balance.[175] For each trial, the results of each formula were then correlated with the Wii Stillness scores for the corresponding trial.

Data Analysis

The data from the WBBs were transmitted via Bluetooth connections. The Wii console analysed the WBB data using the 30-second Stillness test in its Wii-Fit software (Nintendo, Japan). This produced a percentage score, a higher percentage suggesting less movement of the centre of pressure.

The laptop computer used custom-written software (Labview 8.5 National Instruments, Austin, TX, USA). The Labview software was calibrated as detailed in Clark and colleagues' paper.[180]

The Labview software produces a series of x, y coordinates, at a specified frequency, representing the CoP measured. Our data were sampled at 40 Hz and filtered with an eighth order Butterworth filter with a low-pass cut-off frequency of 12 Hz, as per Clark and colleagues. The Stillness test in the Wii-Fit software generates a percentage score. The data filtering and cut-off levels used by the Wii-Fit software were not available to the author and are not readily supplied by Nintendo for commercial reasons.

The Biodex force platform also produces x. y coordinates representing the CoP measured. These data are stored in the integrated console and exported to a .csv file.

The median value of each set of three trials was used. The median was chosen to reduce the opportunity for outlying data to skew the results.

The maximum CoP (X-Y) amplitude for each trial of the laptop data was converted into a percentage score with the equation of the line of best fit from Figure 12.

Maximum X-Y amplitude (cms) = ABS (Max X – Min X) + ABS (Max Y – Min Y)

This resulted in a derived Stillness % score from the laptop data, enabling direct comparison with the Wii Stillness % score from the Wii console.

Agreement between these two datasets was examined by means of Bland-Altman plots.[292] The Bland-Altman plots (also known as a Tukey mean difference plot) plot the difference in percentage score between the Wii console data and the custom laptop data against the mean results When a sample has a large standard deviation, there may be a good correlation between two methods of measuring but this correlation does not necessarily mean that there is agreement between the two measuring devices. One solution is to use a Bland-Altman difference plot, which is used to analyse the agreement between two different methods for measuring the same parameter. The plots have two functions.

1) There should be no relationship between the two methods across the range of measurements. If a proportional bias exists, the methods will not agree across the range of measurements. The data points will show a linear relationship.

2) The 95% limits of agreement lines tell us how far apart the measurements taken by the two methods are. If this range is not clinically relevant, then the two methods of measurement can be used interchangeably.

A two-way, random-effect, single measure (median of the three trials) intraclass correlation coefficients ($ICC_{(2,1)}$) model was used to assess reliability. In conjunction with the ICC values, standard error of measurement (SEM) and minimum detectable change (MDC) values were calculated to assess

the concurrent validity between the laptop-derived Stillness score and the Wii Stillness score, as well as the within-device test-retest reliability and measurement error over the two testing sessions. Point estimates of the ICCs were interpreted as follows: excellent (0.75-1), modest (0.4-0.74), poor (0-0.39).[275]

All statistical analyses were conducted with the Statistical Package for the Social Sciences. (SPSS, IBM Version 21.0) The MDC, which is based on the reliable change index score, was calculated using the equations reported previously by Jacobson and Truax.[276] It is expressed as the percentage test-retest change in score, required to find a significant difference at an alpha level of 0.05 based on the Day 1 mean value.

4.1.3   RESULTS

Validity

The comparison of the CoP coordinate data from the Biodex with each of the stacked WBBs paired with the laptop showed an excellent correlation (R^2= >0.99). The position of the WBB tested had no effect (see Figure 9).

_ _ _ _ _   = 95% confidence intervals

----------   = 95% prediction lines
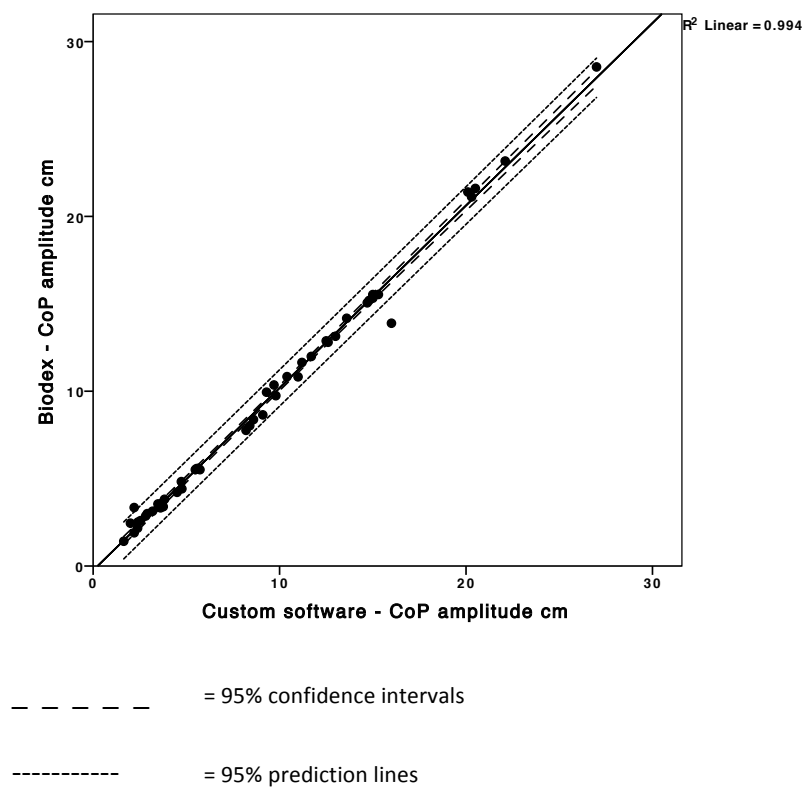
Figure 9 - Centre of pressure data from stacked WBBs linked to the laptop, placed on a Biodex force platform.

Calculating the Stillness measure

The Stillness score was correlated against a number of known balance measures. Each formula was applied to the raw CoP data from the WBB connected to a laptop running the custom software, giving

a result for each trial. The results for each formula were then correlated against the Stillness

percentage scores from a stacked WBB connected to a Wii console.
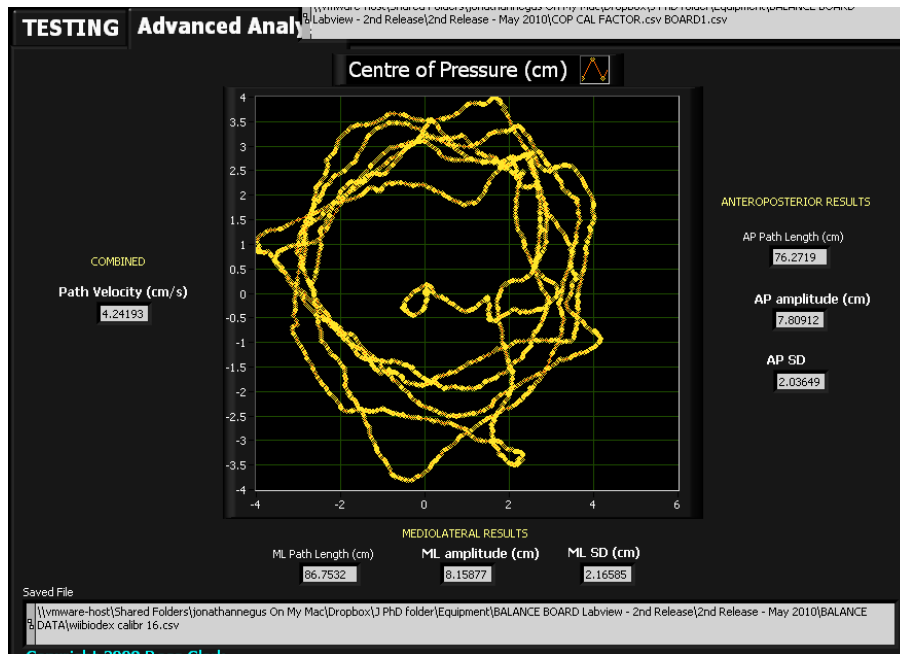
Figure 10 – Centre of pressure graph from a balance trial

Figure 10 demonstrates a line which represents the path taken by the CoP coordinate throughout a

single balance trial. The length of the line connecting all these dots is the path length.

Table 4 illustrates some of the measures of standing balance investigated in the relationship between

the Wii-Fit Stillness score and the raw CoP data from a WBB connected to a laptop: path length, sum

of distance from origin, standard deviation, and variance of total and directional coordinates and 95%

ellipse area, amongst others. The graphs demonstrate that Stillness did not correlate well with path

length or sum of distance from the origin.

Table 4 - Validated measures of standing balance with Pearson correlation coefficients – WBB connected to Laptop correlated to Biodex platform data

Measure	Formulae	Key	R
Path length	$= Sum(SQRT(((ABS(X^{n+1}-X^n)))^2+((ABS(Y^{n+1}-Y^n)))^2$	ABS – magnitude without a sign	-0.691
Distance from start	$= Sum(SQRT((X^n-X^1)^2+ (Y^n-Y^1)^2)$	$X^n = n^{th}$ x-axis value $X^1 = 1^{st}$ x-axis value	-0.505
95% ellipse	$5.99 = \{[x(t)-ux]/sx\}^2 + \{[y(t)-uy]/sy\}^2$	u = mean s = st dev t = timepoint	-0.873
Sum of x-axis & y-axis amplitudes	$= ABS(MAX(X)-MIN(X))+ ABS(MAX(Y)-MIN(Y))$	ABS – magnitude without a sign	-0.958

A selection of the results of correlating Wii-Fit stillness with more commonly used standing balance measures is shown in Figure 11.
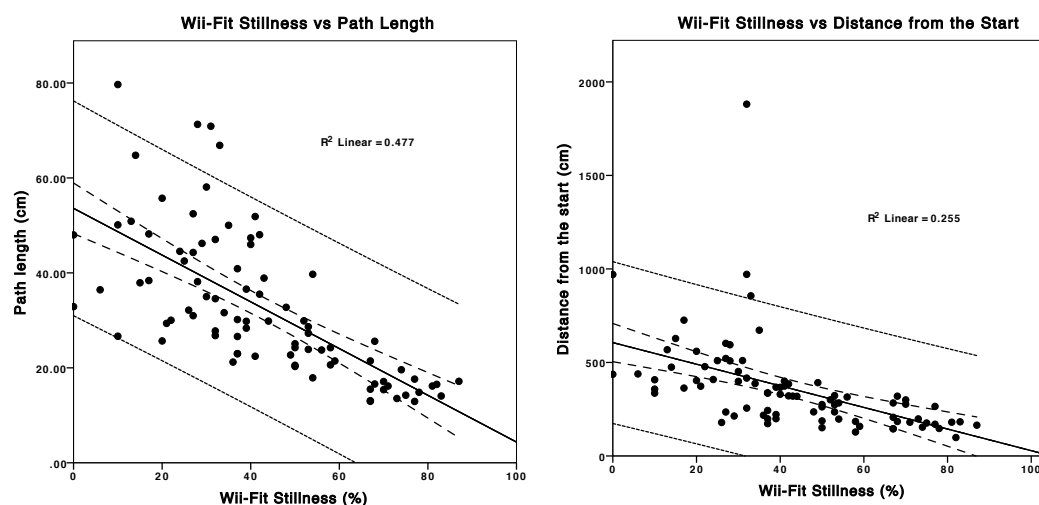
Figure 11 – Wii Stillness correlated with a) path length and b) sum of distance from origin

Ultimately, the measure of balance which showed the greatest correlation was the sum of the mediolateral (ML) and anteroposterior (AP)) ranges. The range can be calculated from the value representing the maximum excursion of CoP at one end of the x-axis added to the value of the maximum excursion at the other end of the x-axis. This value is the range of the x-axis. This was

summed with the range in the y-axis. Using Figure 10 as an example, it was calculated as the most extreme CoP point on the x-axis, which is approximately 8cm, and in the y-axis it is approximately 8cm. This gives a maximum amplitude value for this trial of approximately 16cm. This figure was applied to the formula derived from the line of best fit in Figure 12 to give a Stillness percentage.

When the following formula was applied to the raw CoP data, the results demonstrated an excellent correlation: $R^2 = 0.919$.

Maximum X-Y amplitude (cms) = ABS (Max X – Min X) + ABS (Max Y – Min Y)
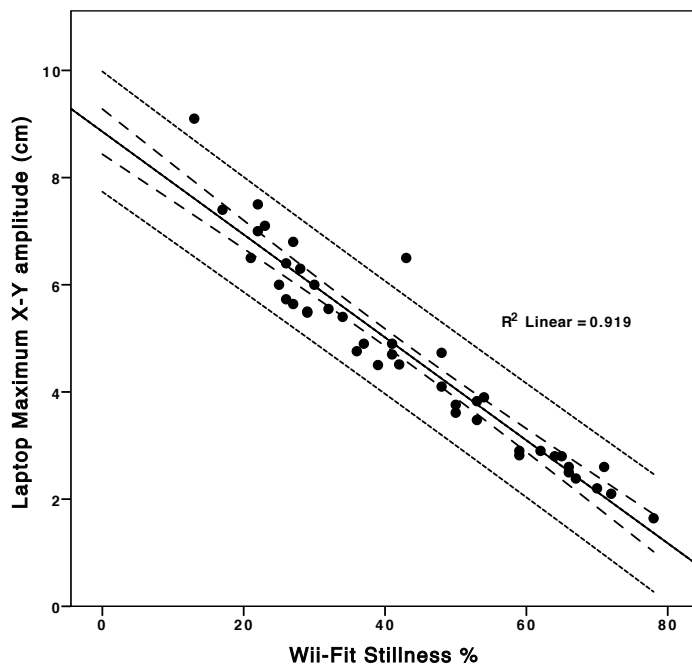
Figure 12 - Correlation of Wii Stillness % score against laptop maximum X-Y amplitude data

The equation for the line of best fit was X = (Y - 8.86) / - 0.1). Three outliers at very low Wii Stillness % scores were excluded to maximize the accuracy of the line of best fit. The outliers were in an extremely low range of measurement which was deemed less clinically relevant to the scores that would be investigated. The lack of predictive accuracy in this low range (less than 20%) was taken into account.

133

The CoP data in centimetres were transformed into a percentage score with the equation above so that the data from both measuring devices (Wii console and Custom software on the laptop) were in the same units, as required when Bland-Altman plots are used for comparison.

validity

Table 5 - Reliability and validity analysis of Stillness % and Labview sum of range converted to % scores during each of the three analysed standing balance trials.

	Laptop sum of range converted to %		Wii Stillness %		Mean Diff	(95 % CI)	ICC	(95 % CI)
USEO	%	SD	%	SD				
Day 1	30.3	(13.6)	30.5	(8.8)	0.2	(-16.4 to 17.2)	0.72	(0.45 to 0.87)
Day 2	29.5	(16.3)	33.0	(11)	3.5	(-16.0 to 23.1)	0.70	(0.42 to 0.86)
Mean diff (95% CI)	0.80	(-24.53 to 21.85)	2.52	(-1.1 to 0.67)				
ICC (95% CI)	0.65	(0.34 to 0.83)	0.85	(0.68 to 0.93)				
SEM	7.74		3.44					
MDC %	71%		31%					
DSEO								
Day 1	57.0	(14.3)	65.3	(10.8)	8.3	(-9.6 to 26.3)	0.74	(0.49 to 0.89)
Day 2	59.1	(8.35)	64.3	(10.9)	5.1	(-8.1 to 18.4)	0.77	(0.54 to 0.89)
Mean diff (95% CI)	2.18	(-17.53 to 21.89)	1.04	(-14.61-18.24)				
ICC (95% CI)	0.64	(0.31 to 0.82)	0.88	(0.74 to 0.95)				
SEM	8.68		3.80					
MDC %	43%		16%					
DSEC								
Day 1	35.7	(18.9)	42.6	(20.2)	7.3	(-10.4 to 25.6)	0.91	(0.80 to 0.96)
Day 2	38.9	(18.2)	43.3	(17.7)	4.8	(-12.0 to 21.7)	0.89	(0.77 to 0.95)
Mean diff (95% CI)	3.18	(-17.50 to 23.86)	0.69	(-14.57 - 19.88)				
ICC (95% CI)	0.85	(0.68 to 0.93)	0.89	(0.75 to 0.95)				
SEM	7.42		6.86					
MDC %	58%		45%					

USEO: Unilateral-leg stance, eyes open; DSEO: Double-leg stance, eyes open; DSEC: Double-leg stance, eyes closed; SD: Standard Deviation; diff: difference; ICC: Intraclass correlation coefficient; CI: Confidence Interval; SEM: Standard error of the measurement; MDC: minimum detectable change, expressed as a percentage of the Day 1 mean value

The complete results of the inter-device ICCs between the laptop custom software and the Wii console Stillness score for each balance stance with confidence intervals are shown in Table 5. In summary -

DSEC validity was excellent in both testing sessions   - 0.91 & 0.89

DSEO validity was modest to excellent   - 0.74 & 0.77

USEO validity was modest   - 0.72 & 0.70

USEC - for the single limb stance with eyes closed, 14 participants were unable to complete at least one trial on both testing days for both machines. Because of this paucity of test-retest data, meaningful analyses were not possible for this stance.

The Minimum Detectable Change values showed considerable variation ranging from 16 to 96% for both the Wii console Stillness score and the laptop-measured maximum X-Y amplitude. This was expected, given the results of previous studies.[180]

The Bland-Altman plots for the Stillness % score and the % score derived from applying the line of best-fit equation to the laptop amplitude data are shown in Figure 13. It is usual practice to interpret Bland-Altman plots informally with no further analyses. The discrepancy between the two devices was approximately +/- 20% in all stances in 95% of cases, as evidenced by the limits of agreement lines set at +/- 2 s.d..

No relationship was evident in the Bland-Altman plots for the DSEO and DSEC stances.
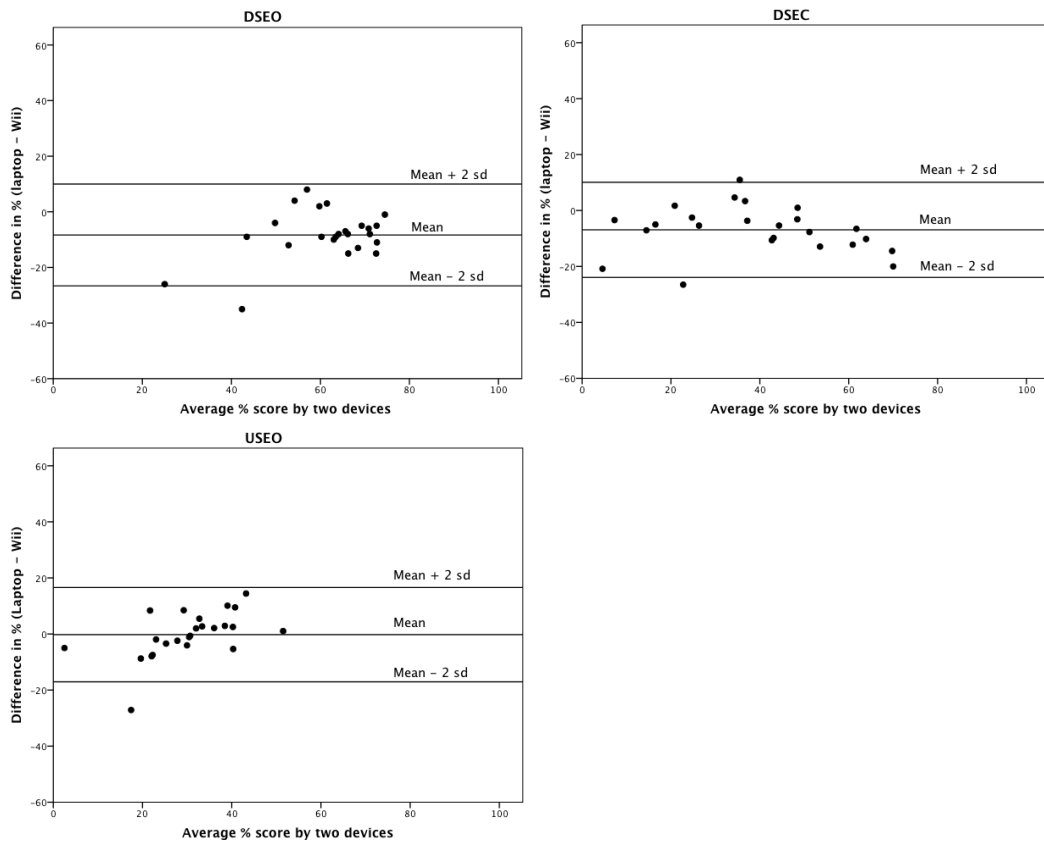
Figure 13 - Bland-Altman plots representing comparisons between a WBB linked to the Wii console and a WBB linked to a laptop (amplitude converted to %) under the three testing conditions: (DSEO) double-limb, eyes open; (DSEC) double-limb, eyes closed; (USEO) single-limb, eyes open. The mean line represents the mean difference between the devices, with the upper and lower lines representing the limits of agreement (2 s.d.)

4.1.4   DISCUSSION

This study demonstrated the Wii-Fit Stillness score to be a reliable measure of standing balance in three of the analysed stances. There were excellent correlations between the two testing sessions in these stances.

The reliability in the USEC stance was not analysed, as a majority of the subjects were unable to maintain their balance in this stance for the full 30 seconds. A patient population which is older or has balance pathology would be likely to struggle even more, suggesting that a USEC stance over 30 seconds is not a good test for these populations. Springer and colleagues showed that the ability to stand on one leg diminishes rapidly from the age of 50 with eyes open or closed.[293] Whereas the mean time to failure for the eyes-open condition is more than 40 seconds for subjects under 50 years of age, it drops to 15 seconds for those in their seventies. When the eyes are closed in unilateral stance, the same age groups can only manage seven and two seconds respectively.
Bohannon and colleagues found that the mean time in seconds whereby healthy subjects could maintain a one-legged stance with their eyes closed was 28.8 (s.d. 2.3) for those aged 20 to 29, but only 4.3 seconds (s.d. 3.0) for those aged 70 to 79. [294]

This suggests that a more relevant test for the USEC stance would be either a 10- second trial length or a measure of the total length of time a subject can maintain the stance and analysis of the centre of pressure data over that time period. In this study, the Wii-Fit software has a rigid time parameter of 30 seconds. If the subject does not reach 30 seconds, no meaningful centre of pressure data are recorded by the Wii. For the safety of the elderly participants and for accuracy and completion of the data, the USEC stance was not utilized in the later trial.

The measured reliability of the Wii-Fit Stillness software score is not surprising. The results which the software produces from an identical input on two different occasions will differ only in terms of the variability of the measuring device and the software. Both the WBB and the Wii-Fit software will be

affected by 'noise' artifacts and by sampling rate, partly because there is no cable shield between

each load cell and the main board. Pagnacco and colleagues reported that the WBB channels are not

sampled simultaneously and that there was a large amount of noise in the system compared with

their posturographic machine.[257] This led to an over-estimation of mean velocity by the WBB of 46%

compared with the force platform.

The WBB uses a first order filter and has a data frequency output of approximately 70 Hz. The data

output frequency cannot be altered by the user, differently from other force platforms which have a

range of output data frequencies, e.g. CAPS Lite 10-1000Hz.[257]

The CoP resolution of the WBB is about 0.5mm, which is low precision compared with other gait

analysis or standing balance platforms (e.g. CAPS**™** Lite force platform = 0.01mm).[257] The noise leads

to an overestimation of velocity by 46% with the WBB compared with the CAPS**™** Lite device, even

when the data are filtered.

The validity of Nintendo's Stillness measure is more complicated. Albeit this score can provide valid

information about the maximum range of the subject's centre of pressure movements, it is prone to

underscoring a subject's balance ability through the inclusion of outlying data points. For example, a

subject with very good balance could complete a trial with the vast majority of their CoP points in a

tight circle around the origin, save for one excursion. The score would only reflect the maximum limits

of that one excursion, not the bulk of the data points close to the centre. Other validated measures,

such as path length, are less affected by single, aberrant CoP excursions. This is a limitation of the

Stillness measure. In order to minimize this effect, the median or best score should be used.

While working out the algorithm used by the Stillness test (Figure 12), there were three outlying data

points in the low percentage range on the X-axis. The low percentage range equates to large CoP

amplitudes or 'poor balance'. These needed to be excluded from the line of best-fit calculation in

order to maximize the correlation with the majority of the data to maximize the predictive power of

138

the algorithm. The exclusion of these data points means that in balance trials with a large amplitude likely to lead to a 'Stillness' score below approximately 20%, the conversion algorithm is unlikely to be correct. However, in all but the USEC stance, a score of less than 20% was very uncommon. Therefore, it made sense to calculate the line of best fit in the zone between 20 and 100% where the majority of data points were plotted.

However, it was important to examine why these outliers existed. They all represented a larger amplitude measurement in the custom software than in the Wii Stillness score. A possible explanation could be a difference in the measurement threshold of the two software programs. The Wii Stillness appears to have a maximum threshold range, which can be exceeded by the CoP amplitudes of some trials with large amplitudes (low Stillness percentage scores). The CoP trace, visible on the television screen during the Wii-Fit Stillness tests, appeared to reach this threshold in a number of the USEC trials. The Labview custom software is able to measure beyond this threshold. This means the custom software measures a larger amplitude than the Wii, represented on the graph by the outliers sitting above the line of best fit.

As a laboratory measure of balance the Stillness test has a number of limitations. It produces a single percentage result giving no access to raw data from the Wii console. The AP and ML components of the CoP range cannot be discriminated, not allowing for further investigation of a low score. The CoP path is traced on the screen at the conclusion of each trial but there is no accurate way to extract these data. The resolution of the line drawn on the screen means that even digital photography of the screen at the conclusion of the trial would not give enough information to produce CoP data.

The length of the test is also set at 30 seconds, reducing the usefulness of the score for more difficult tests such as USEC, when participants are less likely to be able to complete the whole test. It is worth noting that Clark and colleagues, in their paper validating the WBB as a force platform, could customize their testing protocols and their unilateral stances were only held for 10 seconds, explaining why they completed a higher number of successful trials. [180]

Finally, we know from Clark and colleagues[180] that the WBB cannot measure force in the horizontal plane, limiting its use as a gait analysis force platform.

The recommendation from this study is that the median or maximum score of three trials is analysed. This study used median score to maintain consistency with the original validation by Clark and colleagues, but a later analysis of the data showed that the maximum or best score was equally valid and is more clinically relevant. It was the maximum score which was used in the clinical trial with patients. The subjects often had an outlying trial with a much lower score than the other two. If the mean score were to be used, the scores analysed would be lower than the expected compared with other validated balance measures.

Minimum Detectable Change (MDC) is the minimum amount of change in a patient's score which ensures the change is not a result of measurement error. For the DSEO, stance was calculated as 16% (see Table 5). Applying this to the median Stillness score on day 1 of 65.3%, the MDC in a test-retest study would mean a difference in Stillness score of 11% (MDC * day 1 mean score = 0.16*0.653). For DSEC it was 19% and for USEO it was 9.5%. This suggests that the USEO stance is the more clinically relevant because the minimum value of the difference between two separate tests needed in the USEO stance is half that of the DSEC stance.

These MDC values suggest that the maximum amplitude method of assessing standing balance may be clinically relevant in assessing a younger and healthier population using the USEO stance but less so with the DSEO and DSEC stances. However, in the elderly patient population or those at risk of falls, the Stillness score may prove more useful as the subjects are likely to have a lower starting score with potential to make greater improvements, greater than the MDC of all three stances. It is important to be aware that the results of validity tests can only be applied with confidence to a similar population in terms of age or co-morbidity. Therefore, the MDC needs to be calculated for the population in question.

The Bland-Altman plots showed little difference between the two software programs (Wii-fit on Wii console and Custom software on laptop) for the three analysed stances across the range of scores, demonstrating good concurrent validity (Figure 13). This suggests no proportional bias throughout the measurement range for the two devices. In the USEO stance, the points demonstrate a positive slope with an increasing average % score by the two devices. As the average percentage score of the two machines increases, indicating a better balance trial result, the WBB connected to the laptop custom software is outputting a higher percentage score than the WBB connected to the Wii console running the Wii-fit software. Conversely, at lower percentage scores, especially below 20%, the Wii console is measuring a larger % value than the laptop custom software. This is in agreement with the earlier findings shown by the outliers in the line of best fit of Figure 12, showing a measurement discrepancy between the two devices at percentage values below 20%. The fact that this does not happen in the DSEO and DSEC stances may be because of a greater velocity of CoP movement in a single-leg stance compared with a double-leg stance.

The Wii Stillness ICCs on the same device when measured on different days were excellent (0.85 - 0.89) and the ICCs comparing the different devices on the same day were modest to excellent (0.70 - 0.91) for the three analysed stances (see Table 5).

It is worth noting that the subjects never received visual or auditory feedback from the Wii for this validity testing. The Stillness test is designed with three levels of visual perturbation as well as auditory feedback. As the relative contribution of each of these could not be measured from raw data, all of this feedback was removed. In practice, clinicians may leave this perturbation in the test. For the TKR-POWER study, the subjects were always tested in the same manner as in this validation, with no visual or auditory perturbation.

The Wii Stillness score is one of a number of balance tests offered by the Nintendo Wii-Fit software. Although it is sold as a video game for recreation and is not classed as a medical device, it has been used in clinical research and, anecdotally, in clinical practice.[17, 18, 264]

Although the WBB has been shown to be an alternative to laboratory-grade force platforms, albeit with significant caveats, the results of this study cautions against assessing standing balance with the Stillness component of the Wii-Fit software.[180, 257]

Software parameters, such as the inability to modify the length of the test and the inclusion and summation of maximum amplitude of all CoP movements, limit its clinical utility. This is particularly relevant in an older population where the ability to hold stances for longer than 30 seconds or the occasional wide excursion of movement before gaining control or when starting to fatigue is likely to be more prevalent. The inability to further examine the contribution of different planes of movement or the spread of the measurements in those planes and therefore to target particular muscle groups for more focused exercise strategies is another potential limitation.

In conclusion, this study validated the measurement of the Stillness score by the Wii-Fit software and it proved reliable for three of the four tested stances. For future studies the recommendation would be that, although subjects can be tested standing on both legs with eyes open or closed, if standing on one leg, their eyes need to be open. The best score from three trials should be recorded to reduce the influence of outliers. This will maximize the relevance of the data collected by this machine.

4.2 Investigating the associations between balance measured on the Nintendo Wii and functional outcomes in patients with end-stage knee osteoarthritis

4.2.1 Introduction

Osteoarthritis is a common, painful and disabling condition, which affects the knee more frequently than any other joint of the lower limb. The prevalence of knee OA increases with age.[295, 296] Patients with end-stage osteoarthritis of the knee have functional limitations which commonly include prolonged sitting, squatting, kneeling, rising from a chair, and getting in and out of a car.[297, 298] These limitations are largely because of a decline in the neuromuscular control of the knee. This coordinated activity provides stability to the knee joint helping to absorb much of the load placed on it during weight-bearing activities.[299]

Functional control requires adequate muscle strength, knee joint proprioception, and standing balance. There is a steady decline in joint position sense (JPS) and knee proprioception with age in subjects with normal knees. [201, 300, 301] This is worse in those with osteoarthritis and in these patients, is less age-dependent.[201, 302, 303] Knee OA is also associated with a 50-60% reduction in maximum quadriceps torque.[304]

Patients who have had a total knee replacement (TKR) have a better JPS in the operated knee than before the operation,[300] and better JPS than patients with osteoarthritic knees which have not been replaced,[201] but they show a significant impairment compared with age-matched controls.[300] Interestingly, in the patients who had a TKR, there was no difference between the operated and non-operated knee.[300, 305] This has led to the conclusion that the non-operated knee appears to be affected by impaired JPS, independently of whether it is itself arthritic. [201, 283, 300, 302] Therefore it has been postulated that in studies on JPS and proprioception, it is better to use the patient's own contralateral knee as a control than knees of age-matched normal or arthritic patients. This may be

appropriate for studies of a surgical technique or implant design, but in a rehabilitation-based study, the intervention will affect both legs and the aim is to return the patient to the level of function of an age-matched control.

Balance is an integral component of many of the activities of daily living. The task of maintaining balance becomes increasingly difficult with age, as evidenced by the high frequency of falls in older adults.[156, 306, 307] Patients with OA have worse balance than age-matched controls.[164, 278-281] Understanding the mechanisms and interactions of knee OA on balance will lead to a better understanding of the possible mechanisms of disability in this population and may direct treatment for patients.

Standing balance has a direct relationship with threshold JPS in subjects over 70 years old. The ability to recover from a perturbation does not have this direct relationship with threshold JPS in these older subjects. In the context of standing balance, a perturbation is an outside influence such as an uneven standing surface or a physical stimulus which causes a deviation in the balanced state. There was no negative effect on standing balance in subjects with a TKR. [308]

There is good evidence for the association between functional deficits and knee extensor strength[149, 199, 209] but the evidence for associations between functional deficits with standing balance is limited. [277, 309]

In patients with knee OA, reduced knee extensor strength, with body weight accounted for, was moderately associated with impairment of standing balance (Spearman's r=-0.40, p=0.002).[310]

Some studies have reported that not all patients with knee OA have a greater standing CoP displacement than healthy controls.[311-313] This could be explained by the interaction of this association with muscle strength, investigated by Pua and colleagues. [309] They looked at the interactive effects of knee strength and standing balance on function in patients with knee OA. They found a complex

disordinal relationship with more than one strategy employed to maintain balance but used SF-36, not a knee-specific measure (see Figure 14).[309]

In subjects with low knee extensor (quads) strength (<1.10 Nm/kg), balance was negatively related to physical function. The balance variable was the standard deviation of the CoP excursion in the antero-posterior or forward and backwards direction. (The association described in the paper was positive because poor balance gave a higher value for the balance variable.)

In patients with greater knee extensor strength, better balance (s.d. of AP CoP) was positively related to function. The optimum knee extensor strength for threshold for separating patients with and without a poor gait speed was 1.22 Nm/kg. The association was only observed in the AP and not the mediolateral or side-to-side direction.
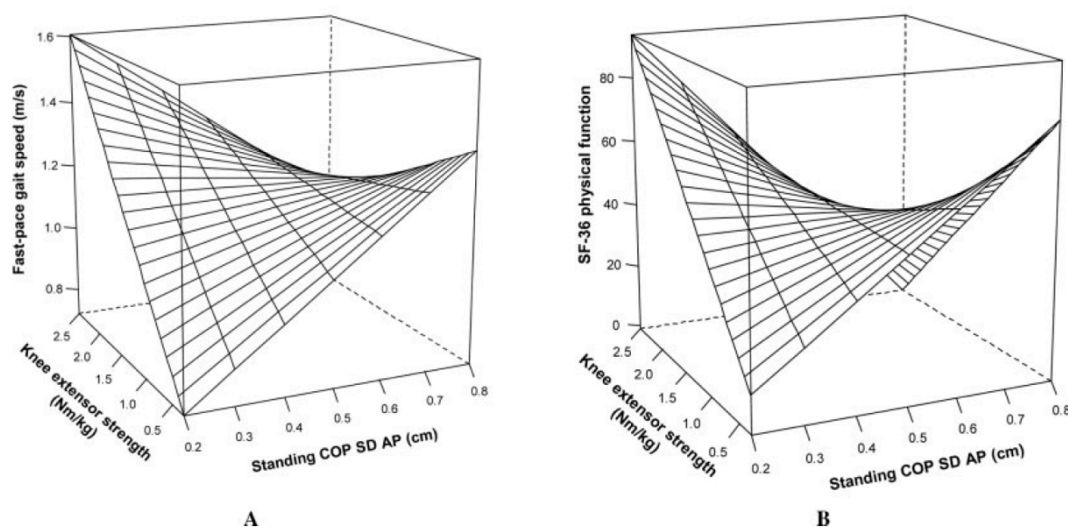
Figure 2. Three-dimensional plots showing the influence of standing center of pressure (COP) SD anteroposterior (AP), knee extensor strength, and their interaction on fast-pace gait speed (A) and Medical Outcomes Study 36-item Short Form health survey (SF-36) physical function (B), adjusted for sex, age, body weight, and SF-36 bodily pain.

Figure 14 – The disordinal relationship between balance, strength and function – Pua et al 2011 [309]

There is growing evidence that CoP displacement may represent the exploratory activity of the CNS to ensure an adequate level of sensory input. In the presence of substantial muscle weakness, the CNS may actually increase CoP excursion in an attempt to increase sensory input or it may employ a global stiffening strategy with global muscle contraction.[314, 315] The patients with knee extensor weakness were using a co-contraction strategy to maintain balance which led to a small CoP excursion but is

associated with a low level of physical function. The patients with knee extensor weakness who had a greater CoP excursion were actually associated with better function, suggesting the exploratory strategy. This departs from the conventional wisdom that lesser CoP displacement is associated with better postural control and physical function.

Schwartz and colleagues found that there was a strong correlation between improved quality of life according to SF-36 and several measures of static and dynamic balance. [195] They performed a linear regression which showed that balance may have a more significant influence than pain on the outcome of TKR as measured by a functional test (four square step test) and the Oxford Knee Score. They showed that in patients who were one year on from a TKR, those whose balance had improved more had better functional outcomes.

Van der Esch and colleagues examined the relationship between balance, muscle strength, and function in patients with OA of the knee. They found the interactions between muscle strength and balance to be more complicated than a purely linear relationship.[316] Bivariate analyses of proprioception, muscle strength and function showed that muscle weakness was related to functional limitation (r= -0.55 to -0.66, p<0.001) and poor proprioception was related to greater limitation in function (r=0.26 to 0.30, p<0.05) and muscle weakness (r= -0.42, p<0.001). However, these relationships were weak. Their multivariate analysis showed that 54% of the variance of the walking time measure of function was explained by muscle strength, proprioception, and their interaction. In the presence of poor proprioception, muscle weakness was associated with even more severe deterioration of functional ability than when proprioception was accurate. The results were similar for the Get-Up-and-Go measure of function. Therefore 'poor proprioception aggravates the impact of muscle weakness on limitations in functional ability. In the absence of adequate motor control through a lack of accurate proprioceptive input, muscle weakness affects a patient's functional ability to a greater degree'.[316]

Studies have been cited in the literature as having found no significant association between standing balance and physical function in bivariate or multivariate analyses.[279, 310, 311] Of these, Hassan and colleagues used the WOMAC score to assess function in 140 patients.[310] Objective functional assessments such as gait speed have been associated with balance and strength whereas the functional subscale of WOMAC has not.[316] Hurley and colleagues compared 103 patients with osteoarthritis of the knee with 25 controls and, although there was no association between standing balance measured by postural stability and functional outcome overall as measured by a battery of objective functional tests, there was a subset of 23 patients who could not maintain the bipedal eyes-open stance.[311] When this group was compared with the rest of the patients who could, there was a significant difference between the two for the functional outcome. Hinman and colleagues examined 33 patients with knee osteoarthritis and 33 controls using a sway meter and a functional step test.[279] They did find significant inverse but weak associations between the step test and 7 of the 12 postural sway variables. They felt these relationships were not strong enough to be used for predicting balance from the step test.

The CoP displacement measured in a standing balance test represents the net neuromuscular response in terms of controlling the passive centre of gravity. Standing balance is therefore a complex and multivariate phenomenon. As both extensor strength and standing balance have been linked to these neurophysiological processes, they should be examined together.

Physical function can be measured subjectively by general measures such as SF-36[309] or disease-specific questionnaires. The WOMAC index is a validated measure used in knee OA and has subscales to examine pain, stiffness, and function.

The WBB can be used as a basic force platform[19, 180] but it has limitations.[257] The software available from retail outlets which comes with the balance board is Wii-Fit. It contains a number of balance 'tests'. One is the Stillness % score, which is a measure of maximum CoP range. The WBB and Wii-Fit

are used by therapists to measure balance performance as well as therapeutically as a rehabilitation tool.

Therefore, this study aims to measure the associations between knee extensor strength measured in Kg by handheld dynamometer and standing balance measured using the Wii Stillness measure on the Wii Balance board with physical function using the knee-specific questionnaire – WOMAC or the UG (up and go test) in patients with severe osteoarthritis of the knee.

4.2.2 METHODS

Sixty patients with end-stage knee osteoarthritis on the waiting list for a total knee replacement were recruited from one orthopaedic centre in Sydney, Australia between March 2011 and September 2012. They were recruited within three months of surgery as part of a randomized controlled trial investigating the use of the Nintendo Wii as a post-surgery rehabilitation tool (TKR-POWER study). Patients were excluded if they 1) had secondary knee OA because of inflammatory disease or sepsis, 2) had a pre-existing neurological, inner ear or other pathology affecting balance, or 3) did not speak English. The study was approved by the local ethics committee and all patients provided informed consent.

Patients attended the research institute after being booked for their TKR. They underwent a physical assessment of height, weight, and knee range of motion. They then completed the following surveys.

Patient-reported outcome measures

Western Ontario McMaster University Osteoarthritis index (WOMAC).[55, 60, 62] An osteoarthritis-specific subjective questionnaire with a total score comprised of pain, stiffness and function subscales. It is scored on a five-point Likert scale from zero to four with a maximum score of 96, a higher score denoting worse outcome. The subscales are scored on: pain (20), stiffness (8), and function (68).

Functional measures

1) Up-and-Go (UG) test over 2.44m[234] (a shortened version of the Timed-Up-and-Go TUG test).

2) Standing balance – Double stance eyes closed (DSEC) - on both legs with eyes closed

3) Knee flexion and extension strength test.

1) Up-and-go

Patients sat in a chair with arms on the armrests. They were instructed to get up after a countdown to go, and walk up to and then around a cone placed 2.44m away, return to the chair, and sit down. The time was recorded from the 'go' signal to the point at which they sat down, recorded to one-tenth of a second. They had a practice trial to ensure they understood the instructions and then the average of two trials was recorded. This is as per the protocol of Rikli and Jones.[234] The Up-and-Go test over 2.44m (8ft) is a shortened version of the Timed-Up-and-Go (TUG) test which was over three metres.[233] TUG is measured in seconds – lower score indicates better function

2) Standing balance was assessed with both the Wii balance board (Nintendo) and the Biodex balance platform. The patients stood with feet a comfortable distance apart, arms by their side, and eyes closed for three tests of 30 seconds each. If they touched the side rail or a foot shifted position, the trial ceased and time in seconds to that point was recorded.

The WBB was placed on top of the Biodex and data were simultaneously recorded from the Wii console connected via Bluetooth to the WBB in the form of a Wii Stillness % score and from the Biodex via its integrated console, recording anteroposterior and mediolateral CoP coordinates. The range of the CoP represents the distance between the most positive and negative CoP trajectory positions in the respective planes and the standard deviation of the CoP represented the variability of the CoP around the mean value. Three successful trials were attempted and testing ceased if there were three unsuccessful attempts. There was a 15-second rest between trials.. The median of three trials was used to reduce the effect of outliers. Although the previous chapter (4.1) describes the

maximum score as the most valid, the median score was used as per the protocol used by Clark and colleagues as this study was designed prior to the results of Chapter 4.1 being available.[180] Wii DSEC is measured in percentage – higher score indicates better function whereas the DSEC AP and ML values are amplitudes measured in cms and a higher score indicates worse function.

3) Maximal volitional isometric contractions of the knee extensors at 90 degrees of knee flexion were performed with a hand-held dynamometer. The patient was in a seated position on the side of the bed, holding onto the edge of the bed with the non-tested leg hanging free. The dynamometer was placed 5cm proximal to the medial malleolus. The maximum of three trials was recorded in kilograms, converted to Newtons and divided by the patient's height to normalize the data. Standardized verbal encouragement was given for each trial to encourage a maximal effort.

Data Analysis

- Demographic comparison – Student's t-test

- Pearson's correlation coefficients were used to examine bivariate relationships between –

    - Balance and muscle strength

    - Balance and function

    - Muscle strength and function

- Linear regression models were constructed to examine the relationship between balance, muscle strength and function

    - The dependent variables were the functional measures - either the WOMAC function subscale score or the UG score

    - The first analyses used independent variables measuring balance (Wii DSEC) and quadriceps strength

    - The second analyses also included age, BMI and gender as independent variables.

Categorical data were presented as counts (%) and continuous data as mean (s.d.)

p-values of <0.01 were considered significant

The data from the WBB were transmitted via Bluetooth connections. The Wii console analysed the WBB data using the 30-second Stillness test in its Wii-Fit software (Nintendo). The Stillness test produces a percentage score with a higher percentage suggesting less movement of the centre of pressure. The percentage score is a measure of the sum of the AP and ML range of CoP. The authors do not know any data filtering and cut-off levels used by the Wii-Fit software. The median value of each set of three trials was used. The median was chosen to reduce the potential for outlying data to skew the results.

Table 6 shows the demographics of the group and the outcomes are shown in Table 6. There is a statistically significant improvement in time taken for the UG test (p<0.0001) as well as the functional subscale of the WOMAC score (p<0.0001).

Table 6 – Demographics

AGE (years)	70.42 (6.9)
Gender (m:f)	30:30
BMI	30.2 (4.48)
TUG (s)	7.4 (2.1)
Knee Extensor strength (N/m)	2.80 (0.86)
Knee Flexion strength (N/m)	2.80 (0.86)
DSEC Wii (%)	45.1 (13.7)
WOMAC Function	23.6 (9.40)

mean (s.d.)

Table 7 – Pearson correlation coefficients (p-value)

	Age	UG	Gender	Quadriceps strength	Hamstrings strength	Wii DSEC	DSEC AP	DSEC ML	WOMAC function
Age	1	**0.400** (**0.002**)	-0.173 (0.187)	-0.770 (0.558)	-0.118 (0.370)	-0.083 (0.526)	0.205 (0.116)	0.028 (0.832)	-0.375 (0.005)
UG	**0.400** (**0.002**)	1	**-0.372** (**0.003**)	**-0.378** (**0.003**)	**-0.492** (**<0.0001**)	-0.029 (0.826)	-0.118 (0.367)	-0.032 (0.777)	-0.020 (0.884)
Gender	-0.173 (0.187)	-0.372 (0.003)	1	**0.450** (**0.0003**)	**0.452** (**0.0003**)	**-0.427** (**0.001**)	**0.332** (**0.009**)	0.214 (0.100)	0.200 (0.148)
Quadriceps strength	-0.770 (0.558)	**-0.378** (**0.003**)	**0.450** (**0.0003**)	1	**0.669** (**<0.0001**)	**-0.294** (**0.023**)	0.209 (0.109)	0.023 (0.841)	-0.269 (0.037)
Hamstrings strength	-0.118 (0.370)	**-0.492** (**<0.0001**)	**0.452** (**0.0003**)	**0.669** (**<0.0001**)	1	**-0.293** (**0.023**)	0.235 (0.071)	-0.027 (0.837)	-0.115 (0.263)
Wii DSEC	-0.083 (0.526)	-0.029 (0.826)	**-0.427** (**0.001**)	**-0.294** (**0.023**)	**-0.293** (**0.023**)	1	**-0.402** (**0.001**)	**-0.616** (**<0.0001**)	0.108 (0.438)
DSEC AP	0.205 (0.116)	-0.118 (0.367)	**0.332** (**0.009**)	0.209 (0.109)	0.235 (0.071)	**-0.402** (**0.001**)	1	**0.383** (**0.002**)	-0.049 (0.685)
DSEC ML	0.028 (0.832)	-0.032 (0.777)	0.214 (0.100)	0.023 (0.841)	-0.027 (0.837)	**-0.616** (**<0.0001**)	**0.383** (**0.002**)	1	-0.162 (0.177)
WOMAC function	**-0.375** (**0.005**)	-0.020 (0.884)	0.200 (0.148)	-0.269 (0.037)	-0.115 (0.263)	0.108 (0.438)	-0.049 (0.685)	-0.162 (0.177)	1

UG – Up and Go test, DSEC – Double stance eyes closed, AP – Anteroposterior, ML - Mediolateral

Looking at age, the functional subscale of the WOMAC score and TUG time have poor to modest correlations that demonstrate a decreasing function with increasing age. TUG time is also negatively correlated with male gender, knee flexion and extension strength meaning that a lower and therefore better TUG time is correlated with male gender and stronger muscles around the knees.

Males have a modest negative correlation to the Wii balance score (Wii DSEC). The Biodex balance data enables analysis of the two components that make up the Stillness score, maximum amplitude in both AP and ML directions. The only significant correlation was in the AP direction not ML which was

at a similar level of correlation to the overall DSEC correlation. This positive correlation means male gender is associated with a poorer score.

Knee extension strength is strongly correlated with knee flexion strength. Both extension and flexion strength are correlated with balance as measured by Wii DSEC.

Balance as measured by Wii DSEC is negatively correlated with gender (modest), knee extensor (poor) and flexor strength (poor) and both the AP and ML measurements of balance on the Biodex.

The Pearson correlations demonstrate no significant association between standing balance as measured by the Wii-Fit Stillness score and function as measured by the function subscale of WOMAC or TUG (Table 7).

Linear regression models found no significant effects of balance on function as measured by TUG or WOMAC function.

Table 8 - Linear regression models testing the association of Functional measures against balance and strength

	Dependent	Independent	R^2	p-value	Coefficient of significant predictors	p-value
1	WOMAC F	Q Strength, DSEC	0.079	0.095	N/A	
2	UG	Q Strength, DSEC				
3	WOMAC F	Q Strength, DSEC, age, BMI, gender	0.249	0.007	Quad S = -4.209	0.009
4	UG	Q Strength, DSEC, age, BMI, gender	0.340	0.0001	Age = 0.096	0.008

WOMAC F – WOMAC functional subscale, UG – Up and Go test, DSEC – double stance eyes closed, Q strength – Quadriceps strength, BMI – Body mass index

4.2.4   DISCUSSION

.

The aim was to investigate the association between balance and function in a sample population of patients with severe end-stage knee osteoarthritis. The study used a new gaming device that is easily available to most clinicians and is gaining popularity in the research and clinical community to measure balance. There was no correlation between standing balance as measured by DSEC Wii Fit Stillness and function as measured by WOMAC function or the UG test. This could be because of the following reasons:

- No association between standing balance and function in this patient population

- The WOMAC function score does not correlate with balance deficits

- The UG measure of function does not correlate with balance deficits

- The Wii Stillness DSEC measure of standing balance is not valid for this patient population

No association between standing balance and function in this patient population

There were no significant correlations between the balance and function outcomes in this study. The regression models showed that when looking at just quadriceps strength and Wii DSEC measured balance as predictors, the models weren't predictive. When age, BMI and gender were added in, the quadriceps strength was predictive for WOMAC function and age was predictive for the UG result. However, the R^2 values were low suggesting that these variables were not accounting for much of the variability.

No previous study has looked at standing balance measured with the Nintendo Wii and its association with functional outcomes in these patients. This study concentrated on balance ability in an attempt to investigate further, why TKR patients are less satisfied with the results of their operation compared with those who have had a total hip replacement. Joint proprioception and the ability of the patient to control the stability and movements of that joint are important to promote the feeling of the joint

being more natural. Combining the damage from chronic arthritis and surgical trauma to the ligaments and the joint capsule leads to significant proprioceptive disruption.[201, 300, 317]

Replacement surgery should produce a mechanically stable knee through the balancing of ligament tensions, but the dynamic stability is provided through the muscles and reflex arcs firing those muscles.

Albeit there may be no association, standing balance is a debatable measure of functional balance. Schwartz and colleagues found that there was only a significant improvement in one of five static standing balance measures in osteoarthritic patients one year after a knee replacement whereas three out of five dynamic tests showed a significant improvement.[195] However, their modeling with linear regression indicates that one year after TKR there is a strong correlation between improved quality of life and improvement in a number of measures of static and dynamic balance. This suggests that balance may have a significant influence on the outcome of TKR. However, all the data from our analysis demonstrate no significant association of our chosen static balance measures with function or quality of life in patients with osteoarthritis.

Males had significantly more quadriceps and hamstring strength on the side affected by arthritis compared with females. However, they had significantly worse Wii Stillness balance scores. The analysis of the force platform data showed this to be because of movement in the AP rather than the ML direction. This may be due to poor balance or a specific balance strategy in that ML direction. This strategy could be due to increased muscle strength tolerating greater excursion in the AP plane without losing balance. Gender specific effects have been shown in obese subjects where obese males show an increased sway in the ML direction compared to obese females. [318, 319] Menegomi et al showed that the increased body weight produced a significant increase in postural activity in men. The increase in postural activity in the ML direction has been reported to be from a more hip centred balance strategy than an ankle based strategy. [320]

The WOMAC function score does not correlate with balance deficits.

The function subscale of the WOMAC score showed a negative correlation with the age of the patients. The older the patient, the poorer their functional WOMAC score. This association is well known and is the same trend as age-matched non-arthritic controls.[321]

The UG measure of function does not correlate with balance deficits

An increased time to complete the UG test was significantly correlated with increasing age, female gender, and decreased quadriceps and hamstring strength., Age was not correlated with decreased muscle strength. This may have been due to the limited age range of the cohort. There may have been other reasons why more elderly patients take longer for the TUG than just having weaker legs. These could be less upper limb strength (to help them rise from the chair), co-morbidities, a feeling of unsteadiness during the test, or simple reluctance to move too quickly for fear of falling.

The Wii Stillness DSEC measure of standing balance is not valid for this patient population

Measuring standing balance with a force platform can cost tens of thousands of dollars and often requires additional equipment such as a laptop and specialized data collection software. Although they can collect detailed information on the balance of the subject, the level of detail may not be clinically relevant, making them unnecessary and uneconomical.

The previous chapter (4.1) described the 'Stillness' measure used by the Wii-Fit software. It is a 30-second balance test, which scores a subject's balance ability with a percentage value. It measures the maximum range of the centre of pressure in the AP and ML directions and sums the two results before converting them into a percentage. This is a gross measure of standing balance which is only likely to be useful as a screening tool. In order to quantify its ability to screen the balance of patients

157

with severe OA, the baseline values needed to be established and then checked for associations with validated patient-reported and more objective functional measures.

The Wii balance board has been shown to be a valid and reliable tool for measuring standing balance in healthy volunteers as well as in patients with certain pathologies such as Parkinson's disease.[19, 180] It has the advantage of being widely available to clinicians and the public alike, and is relatively cheap and extremely portable because of its size and weight. There are concerns about the accuracy of the WBB when used as a device measuring balance and sway (posturographic). Of note, these concerns were raised by authors with a commercially available posturographic machine. [257]

A number of studies have used the Wii-Fit software as part of a balance and rehabilitation protocol. [17, 18, 21, 265, 267, 268] Some have used the in-built balance test results such as the Stillness % score or the Wii-Fit age as a balance outcome. These tests have not been validated and in some cases it is not clear how the software generates the result from the raw CoP data it collects from the WBB. Negus and colleagues tested the validity and reliability of the Stillness test as a measure of standing balance.[322] The single percentage score generated by the software is a summation of the range in both the AP and the ML directions. The range or amplitude of CoP is a gross measure of standing balance which is used in many studies but it is mostly reported as separate AP and ML values. The Wii does not allow any further interrogation of the combined result, limiting further investigation. This measure is also prone to underestimating balance because of one-off excursions of CoP being captured as the result of an individual transient wobble, even if for the majority of the test the CoP is maintained in a tight cluster around the centre. This is why in Chapter 4.1 there was a recommendation for multiple tests and recording the median or best score.

Some studies have used the Wii-Fit age, which is very difficult to assess. The Wii-Fit age is based on two balance tests, which are randomly selected each time the assessment is selected. Therefore, the only way to reliably test an improvement in Wii-Fit age is through multiple tests. Other than Stillness, however, none of these tests have been validated or the background calculations explained.

The results of this study reinforced those of previous studies which describe associations between muscle strength and balance as well as muscle strength and functional limitation. An association between the quadriceps strength of the affected leg and the time taken to perform an UG test is understandable. The quadriceps are the main muscle group used to drive the action of rising from a chair and then in an eccentric mode sitting back down again. Lower body strength declines with age at the rate of approximately 15% per decade.[200]

Female gender was associated with worse TUG scores. This may be because of the lower quadriceps strength to weight ratio of females. As female patients start with less quadriceps strength, and as their arthritis causes a further decrease in strength, they cross the line of functional minimum before males do.

Females also showed significantly decreased hamstring strength. The loss of strength for co-contraction could affect balance, explaining the effect of the association between gender and Wii DSEC max scores.

Balance is a difficult clinical assessment to quantify. Measures such as path length and spread can provide information on standing balance. Path velocity can be used for postural sway. More active tests such as limit of stability or perturbation tests can assess dynamic balance. Each test looks at a different aspect of balance, which will have varying levels of relevance for individual patient pathologies (see Chapter 4.1).

Balance control has been shown to be complicated and not just a function of motor output but motor behavior as an integrative outcome of perceptual, cognitive motor and sensory processes. [183, 184]

This study used the DSEC Wii-Fit Stillness score as a balance measure, even though it has been shown to have a number of limitations. This measure was chosen because the Wii balance board and Wii-Fit are increasingly being used in both research and clinical practice to investigate standing balance in

159

patients with varying clinical pathologies. Therefore, it is important to further investigate the clinical relevance of this software's output as a measure of standing balance.

This study did not find a link between balance and function in patients with end-stage osteoarthritis of the knee. The subjects went on to have a total knee replacement and it has been shown in other studies which not only do functional outcomes improve for those patients who improve their balance outcomes, [207] but the patients with a greater improvement in balance outcomes show more improved functional outcomes.[195]

Therefore, the preoperative data from these patients will be used to investigate the improvement in balance measures from before TKR surgery to one year after surgery and compare the results with functional outcomes.

4.3   THE USE OF TECHNOLOGY IN THE TKR AGE GROUP (CIGAR)

4.3.1   INTRODUCTION

Traditionally, the younger generation embraces technology and is more comfortable and competent with its use than the older generation. The computer and the internet have become integral in the lives of the majority of those in the developed world. Those who have not mastered the basics of this technology are at a disadvantage in a number of areas.

In 1989, Davis and colleagues said that 'Understanding why people accept or reject computers has proven to be one of the most challenging issues in information systems research'.[130] Studies have shown that attitudes play a key role in predicting the acceptance of the user as well as their satisfaction with the use of computers.[131, 132] This is still likely to be true today. Laver and colleagues looked at the acceptability of using the Nintendo Wii as a rehabilitation device in a geriatric rehabilitation hospital. They found that the subjects perceived conventional therapy to be more effective after using the Wii for five days a week during their stay.[261] This was relatively early in the evolution of using the Wii-Fit for rehabilitation and the authors acknowledged that with more evidence of its benefit the subjects might have been more receptive. The lack of perceived benefits may also have been exacerbated by the generic nature of the available exercises on the Wii-Fit. They are designed to be very game-like and may not seem like 'serious' rehabilitation to the more elderly population.

Further studies tried to determine the effects of attitudes and beliefs on individuals' use of computers.[133-135] Many suffer from a lack of stability over time.[136] The rapidly changing face of technology, from the hardware which is available to the average abilities of the general public, means that a score which is specific to one type of technology is likely to date very quickly. It is noteworthy that the majority of these papers were conducted on students and undergraduates. Such validation is unlikely to be appropriate for a population in their seventies. A study of demographic data from 1988 indicated that only 1% of adults over 65 years old used a computer.[137] More recent data from the UK

show that although less than 10% of over-65-year-olds used a computer daily in 2006, by 2013 the number was 37%.[138]

In the field of research, many patient-reported outcome measures (PROMS) which were filled out on paper forms have now been developed as an online questionnaire (www.myclinicaloutcomes.co.uk). This leads to easier collection, database maintenance, and data analysis. However, in order to collect data online, the subject must have access to the internet, preferably at home, and the ability to use it.

The TKR-POWER study plans to ask patients in their seventies and eighties to use a Nintendo Wii at home for three months. It is likely that a number of patients will never have seen or heard of the Wii, let alone know how to use it. There can be a reluctance to use technology which is unfamiliar, especially among the older generation. Therefore, it is important to establish an understanding of the technology with which this age group are comfortable using and to what extent they use it already in their everyday lives.

The Attitudes Towards Computers Questionnaire (1992) was designed for the elderly and tested on 101 adults in a community home between the ages of 57 and 87 (mean 75.1 s.d. 5.93). [144] It examined whether the attitudes of older adults became more positive as they gained computer experience. The results demonstrated that older adults' computer attitudes are modifiable and that direct experience with computers is an effective means of change. The authors also found that it was important for the elderly subjects to experience success when operating the computer.

It has been known for some time that attitudes drive behaviour and that these attitudes can be changed through experience. [145-148] In the case of attitudes towards computers, experience includes direct use of a computer or indirect use through observation of other users. This indirect observation is now pervading the ubiquitous nature of technology in our daily life from smart phones to smart televisions. The media make everyone very aware of available technology such as gaming devices and

tablet devices. In the case of the Nintendo Wii, an extensive advertising campaign targeted the elderly through the use of older actors or depiction of games played with grandchildren.

All of the scales and instruments contained either too much detail or did not ask about aspects of computing on which the studies were focused. The subjects were already going to be subjected to a large quantity of questionnaires and the new one needed to be concise (see chapters 2 and 3). This consideration was the driver of the decision to create a simple questionnaire focused on specifics relevant to the TKR-POWER study. The questions assessed the ability to access the Internet at home, access to an email account, shopping on the Internet as a proxy for general internet ability, and the awareness and use of gaming consoles. The final questionnaire was named the Computer, Internet & Gaming Ability for Rehabilitation score (CIGAR).

This study aims to determine the ability of patients, who are of average age for a TKR, to use computers, gaming devices and related computing functions such as internet and email. It will use a non-validated questionnaire with questions targeted towards the equipment relevant to the planned rehabilitation study using the Nintendo Wii.

4.3.2 METHODS

One hundred and thirty-six patients enrolled in the TKR-POWER RCT. The questionnaire was administered on enrolment into the study, prior to TKR surgery. It was offered in both a paper and an online format (see appendix vii).

The questions were chosen based on the applicability to the future rehabilitation study. This study will require the use of a Nintendo Wii, an internet based questionnaire and email. There was no previous validated study in the literature that was relevant to constructing this questionnaire. The questionnaire was not validated.

The questionnaire consisted of the following questions.

1. How comfortable do you feel using computers in general?

2. Do you have family or friends who help you with technical and computing issues?

3. How often do you check your emails?

4. Do you check your emails on a computer, a smart phone or both?

5. If you have internet at home, what type do you have?

6. How often do you use the internet?

7. How comfortable are you using the following –

    a. Desktop computer

    b. Internet clothes shopping

    c. Internet music / book shopping

    d. Internet grocery shopping

    e. Internet banking

    f. Nintendo Wii / Wii fit

    g. Playstation 3 / X-Box 360

    h. Nintendo DS i.e. braintrainer

    i. Facebook

    j. Twitter

4.3.3　Results

For the first 155 patients who were screened for inclusion to the TKR-POWER study, there were 128 responses with 27 (17%) not completed.

All of the subjects randomized in the TKR POWER study were included in this study (n=136). The remaining 19 were screened but not randomized. Comparing these two groups, 18% of those subjects randomized in the TKR-POWER study did not complete the questionnaire compared with 11% who were screened but not randomized.

The mean age of the subjects was 71.7 years (s.d. 7.61). There were 70 male subjects (46%).

Question 4 yielded a poor response rate as there was confusion over the response options. The responses to this question are excluded from the results.

The results are shown in the following graphs (see Figure 15).

How comfortable are you using computers?
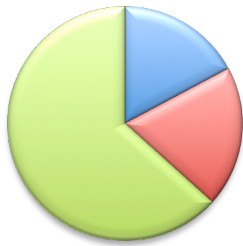
- Uncomfortable n=22 (17%)
- Will use if have to n=25 (20%)
- Comfortable n=81 (63%)

Comfortable - desktop
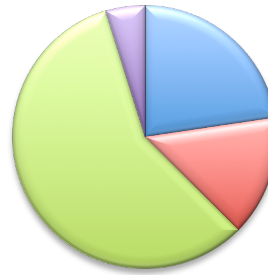
- Uncomfortable n=29 (22%)
- Will use if have to n=19 (15%)
- Comfortable n=74 (58%)
- No response n=6 (5%)

How often do you check your emails?
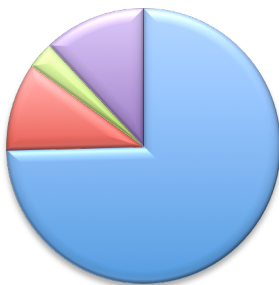
- Daily n=94 (75%)
- Weekly n=13 (10%)
- Monthly n=4 (3%)
- Don't have email n=15 (12%)

What type of internet connection do you have at home?
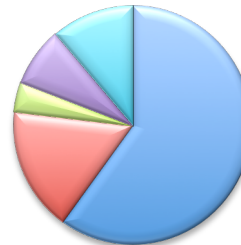
- Broadband n=75 (60%)
- Wireless n=21 (17%)
- Dial-up n=5 (4%)
- Not sure n=10 (8%)

How often do you use the internet?
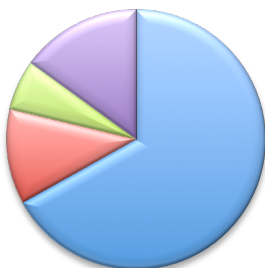
- Daily n=85 (67%)
- Weekly n=15 (12%)
- Monthly n=8 (6%)
- Don't use at all n=19 (15%)

How comfortable are you using the Internet?
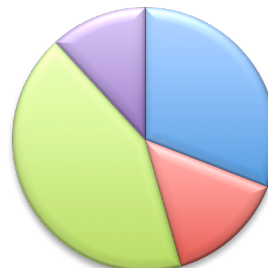
- Uncomfortable n=21 (31%)
- Will use if have to n=10 (15%)
- Comfortable n=29 (42%)
- No response n=8 12(%)

Uses of the internet

58%
41%
30% 27%
18%
13% 12%

Banking Music Books None Facebook Clothes Grocery

How comfortable are you shopping on the internet?

- Uncomfortable n=70 (55%)
- Will use if have to n=22 (17%)
- Comfortable n=19 (15%)
- No response n=17 (13%)

How comfortable are you using: Nintendo Wii, DS, X-Box, Playstation?

- Happy to use n=10 (8%)
- Not comfortable using n=111 (92%)

How comfortable are you using social media: Facebook. Twitter, YouTube?

- Happy to use n=23 (18%)
- Not comfortable using n=105 (82%

Figure 15 – Graphs of the responses to the CIGAR questionnaire

4.3.4 DISCUSSION

The completion rate for this survey was 83% of the 155 subjects surveyed. This was an acceptable sample for this population. There was a similar rate of non-completion of 18% as regards those who were randomized in the TKR-POWER study and 11% as regards those who did not go on to be involved in that study (p=0.465 (Kruskal-Wallis)).

The results demonstrate that the majority of this population are using a computer and feel comfortable doing so. Sixty-three per cent are very comfortable or happy using computers and 73% are happy to use their desktop computer. Seventy-seven per cent have a high-speed internet connection at home and 75% check their emails daily, so computer use is a significant part of their daily lives already. The use of emails is becoming a necessary part of modern life and so this question was used to act as a baseline between the technology which the subjects need to use and the tools which they choose to use.

As regards use of the internet, 67% used the internet daily and only 15% did not use it at all. This value did not surprise me because the more elderly in our society are exactly those who can benefit from using the internet for information, banking or shopping. Communicating via social media is less necessary but still provides a valuable service for keeping in touch with families and friends. There was a disparity between the 67% who used the internet daily and the 42% who felt comfortable using it. Some 15% were happy only occasionally and 31% were not happy at all. This would suggest that a proportion of the respondents used the internet regularly even though they did not feel comfortable doing so.

The majority used the internet for banking (58%) and buying music (41%) and books (30%). Nearly a third of respondents (27%) did not report using the internet for any of the given tasks but only 15% reported not using the internet at all. It would be interesting to know what the former do use it for. This was not part of the questionnaire.

It was also interesting to note that although 42% were happy to use the internet, only 15% were comfortable shopping on the internet. The 17% happy to shop on the internet occasionally is similar to the 15% who were happy to use the internet occasionally. However, a majority (55%) were not comfortable using the internet to shop.

The survey investigated the comfort levels associated with using social media to get an idea of the non-essential uses of the internet and it was low (18%). For social media to be relevant to an individual, they need to be able to connect with friends and family. Therefore, if the respondent peer group is not connected on social media, which these data would suggest is the case (Facebook demographic data), then family must drive the usage. It is likely to evolve slower than that among the younger generations who manage more of their social life through social media.

The use of gaming systems such as the Nintendo Wii (12%) was low as well. This was expected and it was of interest in the light of the proposed intervention arm in the randomized TKR-POWER trial. The proportion of respondents comfortable using the Nintendo Wii, X-box, or Playstation, was expected as it was hypothesised that they would have experience playing with their children and grandchildren. The Nintendo DS's popularity was hypothesised to be slightly higher as a proportion through the use of the brain training systems which have been aggressively marketed as a tool for maintaining memory and mental agility, especially in the elderly. However, the proportions for each category of gaming system were within 2% of each other.

This was a non-validated questionnaire developed specifically for this study because there was no validated questionnaire which could provide the necessary data, as those available were out of date or too technical.

It would have been useful to debrief the participants cognitively to determine if they understood or whether all the questions being asked were relevant– for example it was not specifically asked whether they searched for information on Wikipedia.

The initial hope was to categorize information on smart phone use but this was too complicated in the event. In an age of smart phones and tablets, however, it would be sensible to incorporate these devices into a new questionnaire.

An obvious use of technology which was not mentioned was Skype or Facetime. These programs allow many people to keep in contact with family and friends, especially if they are overseas or live in outback Australia. They fall into the category of essential services, much like banking, but could explain the gaps in the usage categories.

The respondents are all demographically similar. They are all patients living in the North Shore area of Sydney, have private medical insurance, and plan to have their knee replaced by a specialist in a private hospital. There were no data on education level or household income, which may have shed more light on the explanations. These results should be generalized to other populations in different socioeconomic situations with great caution.

This questionnaire was designed to answer some important questions about technology use in the elderly in relation to the proposed TKR-POWER study. The results suggested that the majority of people in this age group, who live in an affluent area of a major city, are comfortable about using computer technology in their homes. The majority have access to high-speed internet at home which they use on a daily basis for emails and internet use. This internet use is mostly for banking, followed by book and music shopping. Finally, the exposure and comfort level as regards various gaming systems is low at around 10%, which means that any study using these gaming systems could potentially encounter resistance from subjects unfamiliar with the technology. However, given the ability to use computers and the internet for necessary purposes, it is possible that they would learn to use the gaming systems and become comfortable with them.

This study aimed to determine the ability of patients, who are of average age for a TKR, to use computers, gaming devices and related computing functions such as internet and email. This study population with an average age of nearly 72 years old, found the majority to be comfortable using computers, emails and the internet, with all three incorporated into their daily life in one way or another. This was important in the planning of the TKR-POWER study (Chapter 5.1) due to the emphasis on technology throughout the planning, data recording and intervention stages. There was concern that this population may not be able to manage this technology. It was apparent that the majority did not feel comfortable using gaming devices. The TKR-POWER study relies on the intervention group using the Nintendo Wii-Fit. The results of this study suggested that in the design of the TKR-POWER study, the set-up and use of the Nintendo Wii needed to be explained clearly with plenty of support, but that this was a population that could manage technology when needed. The results also suggested that the planned options of online questionnaire completion was a possibility for the study subjects.

CHAPTER 5 – Clinical Studies

5.1   TKR-POWER STUDY. PATIENT OUTCOMES USING WII-ENHANCED REHABILITATION AFTER TOTAL KNEE REPLACEMENT

5.1.1   INTRODUCTION

The number of total knee replacement (TKR) surgeries performed annually is increasing across the developed world. The Australian Joint Registry notes that over 40,000 TKRs are performed each year in Australia, and the number is increasing by approximately 6,000 each year.[2] Patients with severe, painful, osteoarthritis of the knee expect that a TKR will reduce or eliminate their pain and improve their physical function and overall quality of life. [110, 323-325] There is good evidence that a TKR leads to the relief of pain but studies consistently report that 20% of patients remain dissatisfied after surgery. [113, 115, 326, 327]

Patient satisfaction scores average approximately 80%, leaving 20% of patients less than satisfied. [113, 115, 326, 327] Health-related quality of life scores remain lower in patients who have had TKR surgery than the scores of an age-matched population. This is despite significant improvements in their scores from their preoperative levels. [110, 113, 324, 328-330] Functional deficits remain in most patients one year after TKR surgery and beyond. These deficits include reductions in quadriceps strength, gait speed, and stair-climbing ability. [152, 331-333] It is not known if these deficits are the inevitable result of disease and surgery or if they can be improved with targeted rehabilitation.

The provision of postoperative TKR rehabilitation differs around the world. Rehabilitation in some form is recommended in the first six weeks.[47, 241, 242] This can be provided to both inpatients and outpatients. The outpatient provision can be one-to-one, group-based, or a home visit.

Post-acute rehabilitation in the United States is mostly provided as an inpatient service or via home visits. In contrast, in Australia and the UK it tends to be conducted via outpatient services. A nationwide survey in Australia reported that 73% of hospitals performed routine referrals to community or outpatient physiotherapy.[47] In England and Wales, 72% of 22 high-volume public orthopaedic centres questioned offered routine outpatient physiotherapy, 11 as a group service and 5 on a one-to-one basis.[241] The reasons behind the specific rehabilitation prescribed to a patient can be as much to do with anecdote, financial incentives, and sociopolitical influences as scientific evidence supporting a particular choice. [42]

A systematic review in 2007 analysed the results of six randomized trials which looked at rehabilitation after a TKR. [37] The studies comparing the location of the rehabilitation found no difference between outpatient therapy and home-based or usual care. [34, 251] A more recent randomized superiority trial with an intervention period of six weeks found that one-to-one therapy does not provide superior self-reported or performance-based outcomes compared with group-based therapy or a monitored home programme.[247] These studies are in agreement with others in finding that unsupervised exercise at home can be as effective as physiotherapist-led rehabilitation on a one-to-one or group basis. [43, 249, 250]

Moffet and colleagues randomized TKR patients to an intensive functional rehabilitation programme or usual care but started the intervention two months after surgery. They measured functional ability and health-related quality of life in 75 patients. [39] The patients attended 12 supervised sessions lasting 60 to 90 minutes each, over an average of six weeks. They found a significant increase in the walking distance and improvement in WOMAC score in the intervention group with less pain and stiffness at four and six months after the operation but not at the final follow-up at one year.

The Maximum Recovery After Knee Replacement (MARKER) study was set up to examine the long-term effectiveness of starting intensive rehabilitation six weeks after TKR surgery rather than on discharge from hospital. The hypothesis is that patients are unlikely to be able to exercise at an

intensity which would contribute significantly to muscle strength in the first few weeks after TKR. This could be because of swelling, pain, or even anaemia. [40]

Historically, in trials investigating adherence with a rehabilitation study up to 50% of patients did not follow the recommended regimens showing that the majority of the non-compliant patients had low motivation levels in the week after surgery during inpatient rehabilitation. [32] Campbell and colleagues [109] found that patients with osteoarthritis who were prescribed physiotherapy tended to continue with the exercises which they found easiest or felt were giving the most benefit. Patients who could incorporate them into their regular routine tended to show approval. Patients who stopped the exercises talked of the various difficulties of doing them alone and cited a perceived lack of progress or improvement. Patients who saw the causes of arthritis as immutable were less likely to exercise than those who saw the intervention as a possibility to minimize the impact of the arthritis. There was a close relation between high levels of adherence and those who saw improvement and hence felt the intervention to be effective.

The Nintendo Wii-Fit is a popular piece of fitness software used by members of the general public ranging from children to the elderly. It is easy to use and can deliver a varied exercise routine; it endows users with the ability to focus on balance when using the Wii balance board. It has been used in a number of trials ranging from stroke rehabilitation to balance training in the geriatric population. [18, 19, 180, 266, 267] The inherent ability to produce scores and demonstrate progress over time could improve adherence with a given programme. Various levels of acceptability have been reported by the older patient population using the Nintendo Wii-Fit. [18, 261]

The Nintendo Wii with Wii-Fit software and balance board cost approximately $300 as of 2012. The cost per unit device used for each patient needs to be considered in the context of rehabilitation in the home with no outpatient visits or supervision. The benefit paid by health insurance companies in Australia per TKR surgery was between $17,888 and $20,423 in 2012-13. At most the hardware is a few percent of the overall cost, before the cost of physiotherapist supervision is taken into account.

It was assumed for this study, that for an intervention to be feasible, clinically effective, and cost-effective it should be home-based, intensive, and start six weeks after TKR surgery. The six-week start time means that the more intensive exercise only begins once the pain and swelling have largely settled.

The aim of the proposed clinical trial is to investigate the effectiveness of using the Nintendo Wii-Fit as a rehabilitation device at home, following primary TKR surgery using the change in the WOMAC total score as a primary outcome. The intervention group will perform the Wii-Fit exercises in place of the routine home exercises. The hypothesis is that compared with usual care, the Wii-fit will improve functional outcomes and increase adherence to the recommended rehabilitation programme.

5.1.2   Methods

The TKR-POWER study (Total Knee Replacement – Patient Outcomes Using Wii-Enhanced Rehabilitation) was a prospective, parallel-group, double-blinded, randomized clinical trial registered with the Australian New Zealand Clinical Trials Registry (ACTRN12611000291987). It was conducted according to the extension of the CONSORT (Consolidated Standards of Reporting Trials) guideline for clinical trials. [334]

Ethical approval was obtained from the Northern Sydney Central Coast Area Health Service Human Research Ethics Committee (Harbour) (EC00333) (protocol no: HREC/10/HAWKE/123), after submission of the National Ethics Application form (http://www.neaf.gov.au/), and the University of Sydney Ethics Committee (Protocol No: 13996). All participants were required to provide informed consent prior to their involvement in the study.

Study Participants

Potential participants were screened and recruited at the rooms of the three operating surgeons when the decision was taken to undergo elective TKR.

Inclusion criteria were: 1) over 18 years old; 2) primary unilateral or bilateral TKR; 3) English speakers. Exclusion criteria were: 1) previous septic joint; 2) revision surgery; 3) inability to return for all extra follow-up visits; 4) dementia, Parkinson's, or other medical condition severely affecting balance.

The trial ran from March 2011 to September 2012. There was no change to the usual medication or rehabilitative care during the perioperative inpatient or postoperative period up to six weeks after the operation.

Research staff performed baseline preoperative assessments. Once the subjects had been randomized by the clinical researcher (JN), JN did not perform any further assessments. All the other research staff were blinded to the group allocation and instructed not to ask the patients about their specific protocols. Participants were reminded at each assessment visit not to reveal their group. The patients were blinded inasmuch as they did not know the details of the two rehabilitation strategies employed in the study. Therefore, neither group was aware of what the other group's rehabilitation consisted of or whether they were in a control or intervention group.

Randomization

Randomization was performed with a computerized random number generator to generate a sequence for patients with a 1:1 ratio for each group. An independent administrative officer sealed randomization codes in sequentially numbered, opaque envelopes. Randomization was performed after recording all the outcome data at the baseline appointment. The groups were assigned to the

participants by one clinical researcher, who opened the next numbered envelope as each patient

reached the six-week mark.

Interventions

Control group

Participants were provided with written instructions and diagrams detailing the exercises to be

continued at home between 6 and 18 weeks post-op, which is standard practice at the hospitals in

the study (see appendix ix). They were given a paper diary in which to record their daily exercise. The

diary served as a source of instruction on when to perform which exercises. All diaries and exercises

were explained by the lead author (JN) in an interview at randomization lasting approximately 20

minutes.

The exercises were taken from the standard rehabilitative exercise sheet given to patients for home

use at one of the major rehabilitation hospitals where the majority of the participants received

treatment. These included advice on performing squats, step-ups, and calf raises for a total of 30

minutes each day. Participants in the control group were free to perform any other exercises

prescribed or otherwise, including physiotherapy, hydrotherapy, walking, and sports.

Wii Group

Participants received a set of paper diaries, which detailed three stages of Wii exercises of increasing

difficulty every four weeks. (see appendix viii) They were given a Wii console and balance board plus

the Wii-Fit software to take home. The exercises included yoga, strength training, aerobics, and

balance work. The Wii was installed by the lead author (JN).

Participants also received the same standard rehabilitation exercise program instruction sheet and

paper diary as the control group and were asked to record any of the standard rehabilitative exercises

which they performed. (see appendix ix) Participants in the Intervention group were free to perform any other exercises prescribed or otherwise, including the routine exercises given to the control group, additional physiotherapy, hydrotherapy, walking, and sports.

The Wii Fit and Balance Board

The Wii is the computer console running the Wii-Fit software. It is connected to a standard television. The Wii balance board connects to the Wii console wirelessly via Bluetooth technology. The participant controls the choice of exercise by using an infra-red camera-based handheld remote control. They then perform their exercises while standing on the balance board. The participants are aiming to control the position of their centre of pressure on the board, which is relayed to them via a positioning dot on the screen or the movement of a character or device on screen. The Wii-Fit exercises are divided into yoga, strength, balance, and aerobic tasks.

Outcome Assessments

Assessments were conducted preoperatively when the participant was first recruited (and repeated if the recruitment date was more than three months before the operation). Assessments continued postoperatively at 6 weeks (study baseline), 18 weeks, 6 months and 12 months. All assessments after six weeks were performed by investigators blinded to the treatment allocation. Demographic information collected at baseline included socio-economic status, co-morbidities, height, weight, history of lower limb arthritis, trauma, and previous orthopaedic surgery.

Primary outcomes

The primary outcome was the self-reported Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index total score, which used a five-point Likert scale (WOMAC LK3.1). The WOMAC consists of 24 items covering three subscales: pain (5 items), stiffness (2 items), and physical function

(17 items). Each item could score from zero (no pain, stiffness, or difficulty) to four (maximum pain,

stiffness, or difficulty). This led to a minimum score of zero and a maximum score of 96, a high score

representing a worse outcome. The WOMAC is a widely used questionnaire specifically designed to

evaluate knee and hip OA. Its clinimetric properties have been validated in patients undergoing TKR.
[93, 335, 336]

Secondary Outcomes – see Table 9

Table 9 - Secondary outcomes and timeline of assessments

TEST	DESCRIPTION	STATISTICAL TEST
Measured at all Assessment Timepoints		
Oxford Knee Score[85]	12-item self-reported functional score for TKR. Range 0-48.	Unpaired t-test of change score
SF-12[91]	12-item Short-form Health Survey. Data from the 12 items is used to construct the physical and mental component summary scores.	Unpaired t-test
VAS – pain score[102]	The patients bisect a 10cm line with 0 being no pain and 10 being extreme pain. Standard question - 'How much pain are you feeling in the operated knee on an average day within the last week?'	Unpaired t-test
AKCRS[39, 40]	The objective score – includes knee ROM & alignment. The ROM is measured with a goniometer. Knee flexion and extension are measured as an active movement while the participant is supine.	Unpaired t-test of ROM and score
Balance[19, 180, 204]	Static standing balance tests are performed with the Wii-Fit giving a Stillness score and the Biodex balance platform giving raw coordinate data of double leg with eyes open (DSEO) or closed (DSEC) stances.	Unpaired t-test
Strength[229, 337]	Maximal isometric knee flexion and extension strength for both legs – isometric muscle strength assessed with a handheld dynamometer seated for quadriceps and prone for hamstrings. The best of three attempts is recorded in Kg.	Unpaired t-test
Up-and-Go[234]	This is the shortened version of the Timed Up-and-Go test. It is the time taken for the participant to get up from a chair, walk 2.44m, turn around a cone and return to the chair to sit down.	Unpaired t-test
End of Intervention Period – 18 Weeks		
Exercise diary	Participants record their exercise frequency in a paper diary. The Wii group also have their exercises recorded by the Wii console.	Unpaired t-test of mean number of exercises per week
At 18 Weeks and 1 Year Only		

	A 3-question self-reported assessment.	
Patient satisfaction	1) How satisfied are you with your operated knee? 2) How satisfied are you with your rehabilitation? 3) Does the operated knee feel normal?	Chi Squared test

At each follow-up assessment, participants were asked about the occurrence of any adverse events. Adverse events were defined as musculoskeletal or cardiovascular events resulting in hospitalization and any falls for whatever reason.

Sample size

The sample size was calculated to detect a minimal clinically important difference in the WOMAC total score at 18 weeks after total knee replacement between the control and Wii groups. The minimal clinically important change was defined according to criteria suggested by Bellamy and colleagues: a standard deviation of 14 for calculating sample size with use of this scale, and a change of at least 7.4 points following a weak intervention, such as a low dose of Voltaren**™** (diclofenac sodium). [60]

The sample size calculation was constructed to detect a difference of 8 with a standard deviation of 14, a ratio of treatment to control of 1:1, a type-I error of 5% and power of 80%. The sample size was calculated to be 49 in each group. After allowing for a loss to follow-up of at least 20%, a sample size of 128 (64 in each group) was chosen.

Statistical analysis

Data were analysed according to intention to treat principles. Unpaired T-tests assessed the differences between the groups at six weeks, eighteen weeks and one year. The change in the WOMAC total score from six weeks to eighteen weeks and six weeks to one year between the two groups was compared by ANOVA analysis.

There were no statistical methods applied to any missing data.

To account for any baseline differences, a linear regression model including the WOMAC total score at week 6 as a covariate (ANCOVA) was used to compare difference in the WOMAC total score at week 18 between the two study groups. The same analysis was applied to the WOMAC total score at one year as a secondary analysis.

A general estimate equation model was applied with WOMAC total at 18 weeks and one year as dependent variables and independent variables were group, time of assessment, preoperative, and six-week WOMAC total score. Models were built for further analysis of potential confounders and to assess the effect of the intervention throughout the study taking all timepoints into account in one analysis. The confounders analysed were age, BMI, gender, balance, SF-12 (mental) and flexion. Age, BMI, gender and flexion have previously been demonstrated as variables that affect recovery from TKR surgery. [329, 338, 339] As age was correlated with WOMAC function score and gender was correlated with balance in the analyses shown in chapter 4.2.3 (page 88), these two variables were included in the regression models. The models also investigated the potential confounders of balance and mental state (SF-12 mental) as they were of particular interest to the study question.

The 20,50,70 method of Bellamy was used to assess discriminant validity.[81] The proportion of subjects who reached the threshold level of at least a 20% improvement from baseline was compared between treatment groups. The same comparison was then performed for a 50% and a 70% improvement.

Appropriate statistical methods were used for the secondary outcomes.
- T-test for parametric continuous data
- Mann-Whitney-U and Chi-squared for non-parametric data

5.1.3 Results
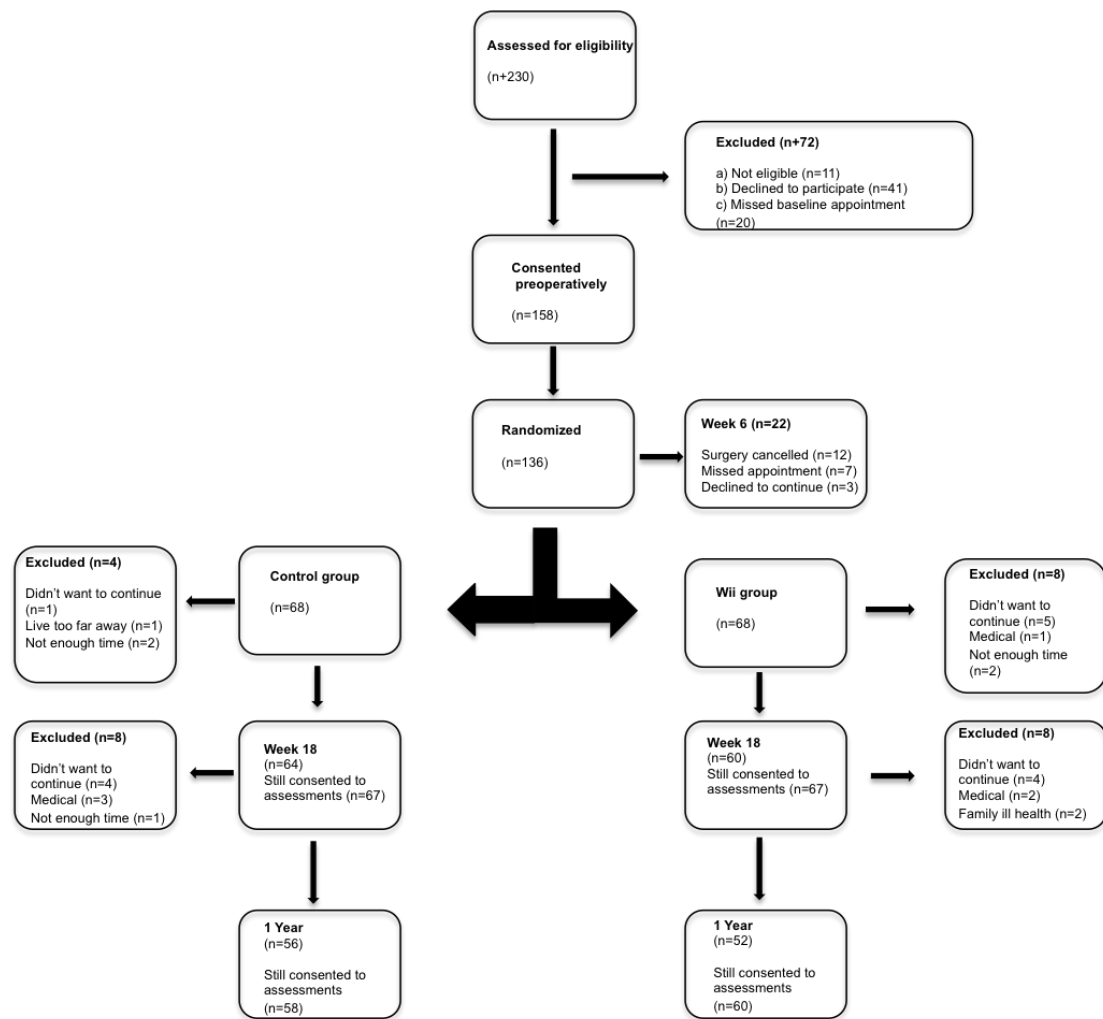
Figure 16 – Flow chart of recruitment numbers for the TKR-POWER study

From March 2011 to September 2012, 230 subjects were screened for eligibility after they had been booked for their TKR surgery (see Figure 16). Some 158 consented to the study and 22 withdrew before randomization at six weeks following surgery. This left 136 patients to be randomized (95 unilateral and 41 bilateral) in the trial; 68 were allocated to each group.

Between the randomization at six weeks and the end of the intervention period at eighteen weeks, four participants withdrew from the control group (5.9%) and eight from the Wii group (11.8%). Of

these that declined the intervention, two in the control group (2.9%) and one in the Wii group (1.5%) declined to attend the assessments and were deemed lost to follow up.

Between the end of the intervention period at eighteen weeks and the one-year follow-up assessment, eight participants withdrew from the control group and eight from the Wii group. Only one of the eight in the Wii group agreed to continued assessments giving a loss to follow up of 14.7% for the control group and 11.8% of the Wii group.

This left 66 in the control group and 67 in the Wii group who were continuing assessments at 18 weeks (133 total (97.8%)). At one year there were 58 in the control group and 60 in the Wii group who were continuing assessments (118 total (86.8%)) and therefore included in the intention to treat analysis. All subjects were analysed according to their randomised group and all data that was collected was included in the analysis regardless of whether the patient continued with their allocated intervention.

There was no difference between those lost to follow-up and those who were followed up in terms of age, sex, BMI, and preoperative WOMAC.

Participant characteristics – Preoperative – (see Tables 10 & 11)

All patients had primary TKRs performed by three surgeons. Each surgeon used a consistent prosthesis and protocol for each of their patients. The first used posterior cruciate sacrificing prostheses, resurfaced the patellae and cemented all components. The second used cruciate retaining prostheses, resurfacing the patellae and routinely cementing the femur and patella with some tibial components cemented and some uncemented. The third used all cemented cruciate retaining prostheses and resurfaced the patellae.

Conventional surgical techniques were used for all patients including a medial parapatellar approach and computer navigation.

All patients were given the option of up to two weeks of inpatient rehabilitation if requested by the patient or advised by the rehabilitation specialists. They all had access to outpatient rehabilitation in the first six weeks postoperatively. There was no change in treatment during these first six weeks of any of the study participants.

There were 95 subjects who had unilateral TKR surgery (50 control group and 46 Wii group).

There were 41 subjects who had bilateral TKR surgery (18 control group and 22 Wii group).

They had a mean age of 71.5 years (7.6). There was no significant difference between the control and the Wii group with regard to age, sex, or BMI.

There were no significant differences (p>0.05) in the other preoperative measures.

Participant characteristics – 6 Week BASELINE – (see Table 11)

There were no significant differences (p>0.05) between the groups at the Week 6 baseline assessment.

Table 10 - Demographics

	Control group	Wii group	P value
Age	72.3 (7.8)	70.6 (7.2)	0.201
Sex Male : Female	31:37	32:36	0.864
BMI	30.4 (5.5)	30.1 (4.1)	0.717

BMI – Body Mass Index

Table 11 - Baseline characteristics PREOP and 6 WEEKS

Outcome	Mean (sd)				P value	
	Control group		Wii group			
	PREOP	WEEK 6	PREOP	WEEK 6	PREOP	WEEK 6
WOMAC Pain	7.46 (3.06)	5.42 (3.7)	7.31 (3.15)	6.83 (10.4)	0.790	0.314
WOMAC Stiffness	3.39 (1.36)	2.73 (1.57)	3.64 (1.76)	2.76 (1.41)	0.382	0.916
WOMAC Function	24.51 (9.02)	15.5 (10.4)	24.54 (10.65)	17.9 (9.12)	0.985	0.166
WOMAC Total	35.36 (12.36)	22.5 (12.6)	35.49 (14.36)	26.3 (12.6)	0.956	0.089
OKS	26.95 (7.11)	32.90 (7.66)	27.25 (7.28)	31.75 (7.53)	0.828	0.370
SF-12 PCS	36.28 (8.50)	40.05 (8.10)	36.84 (7.97)	38.27 (8.84)	0.699	0.452
SF-12 MCS	54.08 (9.67)	54.78 (5.98)	55.82 (9.51)	52.87 (8.71)	0.300	0.357
VAS – pain (mm)	50.3 (27.8)	27.3 (21.8)	49.8 (29.1)	30.6 (22.0)	0.879	0.075
Wii DSEO max (%)	57.68 (12.29)	58.6 (11.3)	60.68 (9.71)	60.3 (10.4)	0.134	0.364
Wii DSEC max (%)	41.95 (13.76)	38.9 (12.9)	40.70 (12.91)	40.7 (13.8)	0.653	0.448
Quadriceps strength (kg)	20.83 (8.84)	17.8 (6.30)	21.64 (7.55)	17.9 (5.18)	0.586	0.950
Hamstrings strength (kg)	10.76 (4.66)	9.67 (3.74)	11.37 (4.41)	10.01 (3.29)	0.459	0.575
Flexion (degrees)	109.8 (15.34)	111.74 (12.87)	110.98 (11.34)	103.36 (15.70)	0.369	0.051
UG (secs)	8.21 (2.59)	7.92 (1.96)	7.73 (2.17)	8.08 (3.32)	0.173	0.677

WOMAC – Western Ontario & McMaster Osteoarthritis Index, OKS – Oxford Knee Score, SF-12 – Short Form -12, VAS – Visual analogue scale, DSEO – Double stance eyes open, DSEC – Double stance eyes closed, UG – Up and Go test

Table 12 - Results at 18 weeks

Outcome	Control group Mean (s.d.)	Wii group Mean (s.d.)	Diff between groups Mean	P value
WOMAC Pain	2.93 (3.26)	1.94 (2.06)	0.99	0.044
WOMAC Stiffness	2.00 (1.29)	1.97 (2.38)	0.03	0.927
WOMAC Function	10.22 (9.58)	8.13 (6.75)	2.09	0.163
WOMAC Total	15.15 (12.88)	11.33 (8.57)	3.82	0.054
OKS	40.23 (5.90)	40.92 (5.10)	-0.69	0.502
SF-12 PCS	45.62 (9.32)	49.05 (7.17)	-3.43	0.129
SF-12 MCS	57.27 (5.97)	60.60 (3.70)	-3.33	0.016
VAS pain (mm)	12.9 (14.1)	11.4 (11.2)	1.5	0.543
Wii DSEO max (%)	66.9 (9.79)	68.6 (7.84)	1.7	0.347
Wii DSEC max (%)	45.0 (10.6)	51.4 (13.7)	6.4	0.009
Quadriceps strength (kg)	22.75 (7.33)	22.40 (6.28)	0.35	0.789
Hamstring strength (kg)	11.27 (4.18)	11.61 (4.79)	-0.34	0.690
Flexion (degrees)	111.7 (12.9)	110.1 (14.0)	-1.6	0.533
TUG (secs)	6.69 (1.35)	6.54 (1.72)	0.15	0.590

WOMAC – Western Ontario & McMaster Osteoarthritis Index, OKS – Oxford Knee Score, SF-12 – Short Form -12, VAS – Visual

analogue scale, DSEO – Double stance eyes open, DSEC – Double stance eyes closed, UG – Up and Go test

Table 13 – Results at 1 year

	Control group	Wii group	Diff between groups	P value
Outcome	**Mean (s.d.)**	**Mean (s.d.)**	**Mean**	
WOMAC Pain	1.35 (2.57)	1.16 (1.60)	0.96	0.662
WOMAC Stiffness	1.17 (1.13)	1.25 (1.15)	0.02	0.716
WOMAC Function	6.81 (7.58)	4.96 (5.79)	1.82	0.168
WOMAC Total	9.33 (10.4)	7.42 (7.95)	0.89	0.303
OKS	42.66 (4.17)	43.40 (5.15)	-0.74	0.474
SF-12 Phy	46.59 (12.41)	46.66 (9.67)	-0.079	0.988
SF-12 Men	55.37 (8.51)	54.37 (12.37)	1.00	0.847
VAS (mm)	8.62 (11.26)	6.92 (10.77)	1.70	0.593
Wii DSEO max (%)	64.9 (10.6)	67.3 (9.5)	-2.4	0.280
Wii DSEC max (%)	46.6 (14.1)	47.3 (13.3)	-0.7	0.823
Quadriceps strength (kg)	26.12 (7.06)	25.53 (8.00)	0.59	0.734
Hamstrings strength (kg)	14.16 (4.67)	13.05 (4.72)	1.11	0.287
Flexion (degrees)	116.9 (15.5)	117.9 (12.9)	-1.0	0.734
TUG (secs)	6.61 (1.80)	6.44 (1.72)	0.17	0.667

WOMAC – Western Ontario & McMaster Osteoarthritis Index, OKS – Oxford Knee Score, SF-12 – Short Form -12, VAS – Visual analogue scale, DSEO – Double stance eyes open, DSEC – Double stance eyes closed, UG – Up and Go test

There was no significant difference between the groups at any point for group average WOMAC total score. This was not the primary outcome. The primary outcome was absolute mean change of WOMAC total score from the six week post-op assessment and the end of the 12-week intervention period (18 weeks post-op assessment) (Table 14). A one-way ANOVA of the data yielded the following results.

Table 14 - ANOVA of WOMAC total score mean difference from 6 weeks to 18 weeks

	Control mean difference (95% CI)	Wii mean difference (95% CI)	Difference (95% CI)	P-value
Week 6 to 18	-8.03 (-10.5 to -5.6)	-15.3 (-17.9 to -12.7)	7.2 (3.7 to 10.8)	<0.0001
Week 6 to 1 year	-12.2 (-15.7 to -8.8)	-18.0 (-21.5 to-14.6)	5.8 (1.0 to 10.6)	0.019

(mean (95%CI))

Both groups improved their WOMAC total scores significantly to the 18-week and the one-year point. However, the Wii group improved nearly twice as much as the control group (15.3 vs. 8.0) at 18 weeks and 50% more at the one-year mark (18.0 vs. 12.2). The difference between the improvements in the two groups was significant at both these time points.

Table 15 - ANCOVA of WOMAC total score mean difference from 6 weeks to 1 year

	Control mean difference (95% CI)	Wii mean difference (95% CI)	Difference (95% CI)	P-value
Week 6 to week 18	-9.1 (-11.0 to -7.21)	-14.2 (-16.1 to -12.4)	5.1 (2.5 to 7.8)	<0.001
Week 6 to 1 year	-14.0 (-16.4 to -11.4)	-16.4 (-18.8 to -14.0)	2.4 (-0.57 to 7.12)	0.176

(mean (95%CI))

Even though there is no significant difference between the groups at six weeks, the p-value was 0.089, suggesting that the difference of 3.8 points should be accounted for in the analysis. To account

for this difference, the ANCOVA analysis was used. This compared the difference between the two

groups at the defined timepoints, adjusting for the six-week baseline measurements (Table 15).

Table 16 – Generalised Estimating Equation (GEE) Models adjusting for potential confounders and baseline WOMAC.

Model No.	Independent variables	Effect size	95% CI	p-value
1	Group (Control vs Wii)	3.99	1.60, 6.38	0.001
	Time of assessment (1 year vs 18 weeks)	-3.39	-5.09, -1.69	<0.001
	Preoperative WOMAC Total	0.10	0.01, 0.20	0.04
	WOMAC total at week 6	0.36	0.26, 0.47	<0.001
2	Group (Control vs Wii)	3.58	1.17, 5.99	0.004
	Time of assessment (1 year vs 18 weeks)	-3.36	-5.06, -1.67	<0.001
	Preoperative WOMAC Total	0.11	0.01, 0.21	0.03
	WOMAC total at week 6	0.38	0.27, 0.48	<0.001
	Age	0.16	0.00, 0.33	0.05
	BMI	0.15	-0.10, 0.41	0.24
	Male gender	1.06	-1.32, 3.44	0.39
3	Group (Control vs Wii)	4.42	1.32, 7.51	0.007
	Time of assessment (1 year vs 18 weeks)	3.93	1.73, 6.12	0.001
	Preoperative WOMAC Total	0.15	0.02, 0.27	0.027
	WOMAC total at week 6	0.22	0.08, 0.36	0.003
	Age	0.18	-0.04, 0.40	0.116
	BMI	0.13	-0.21, 0.46	0.461
	Male gender	1.81	-1.62, 5.24	0.304
	Flexion	-0.04	-0.19, 0.11	0.623
	SF-12(MCS)	-0.00	0.00, 0.00	0.436
	Wii DSEC	-0.10	-0.21, 0.02	0.111

(Wii DSEC = double leg stance, eyes closed; SF-12 (MCS) – Short Form 12 Mental component subscale; Effect size = coefficient for effect of Wii intervention vs control, CI = confidence interval)

The regression was performed using General Equation Estimate models (Table 16). Models 2 & 3

include additional independent variables. These variables were chosen for the model as they

demonstrated statistical significance in the ANCOVA analysis or have been shown in the literature to

affect outcome following TKR. The models take all the independent variables into account with

WOMAC Total score throughout the study as the dependent variable and calculates the co-efficient as

an indicator of the size of the effect being in the intervention group has on the Total WOMAC score at

any time point.

Using the models to take multiple assessments into account, the analyses showed that the Wii group

had a better average difference in WOMAC total score from the six-week to the one-year timepoint

without confounding adjustments. The addition of confounding adjustments for variables such as age, BMI, gender and flexion improved the coefficient of the model but all models shown in table 16 were statistically significant.

Table 17 - Analysing the 18-week results using the 20 & 50 of the 20,50,70 method of Bellamy and colleagues.[81]

	Control	Wii	Difference	P value
W 20 18 weeks	41 (71%)	56 (93%)	22%	**0.001***
W 50 18 weeks	23 (40%)	36 (60%)	20%	0.028
W 20 1 year	47 (81%)	55 (92%)	8%	0.093
W 50 1 year	37 (64%)	50 (83%)	13%	**0.016***

((W=WOMAC Total score; 20 or 50 = Improvement from baseline of 20 or 50% (* a Bonferroni adjustment gives a significant result as p<0.0125 (0.05/4))

There was a significant difference between the groups with a greater number of the Wii group improving by 20% of their 6 week baseline WOMAC total score at 18 weeks (p=0.001) but not at 1 year (p=0.093). When looking at those who had improved their score from 6 week baseline by 50%, there was no significant difference between the groups at 18 weeks (p=0.028) but it was significant at 1 year favouring the Wii group (p=0.016). (Table 17). After a Bonferroni adjustment for the two analyses of 20% and 50%, the level of significance was set at 0.025.

It was deemed highly unlikely that there would be enough patients showing a 70% reduction in scores because of the high level of function already shown at 6 weeks, so it was decided a priori not to assess the 70% improvement. Therefore, only the 20 & 50 part of the 20,50,70 method of Bellamy was used to assess discriminant validity.[81] A post hoc analysis comparing the groups by those showing a 70% reduction in scores showed no significant difference for the total score at 18 weeks (Wii n=20 vs control n=12 (p=0.124)) or at 1 year (Wii n=38 vs control n=28 (p=0.101)).

Secondary outcomes

The mean scores for the secondary outcome measure are presented in Table 12 for 18 weeks and Table 13 for one year.

Satisfaction

Analysis of the satisfaction questions revealed no statistically significant difference between the groups in terms of those who responded (see Table 18). The patients expressed a very high level of satisfaction with their operated knee and their rehabilitation.

Table 18 – Chi-square test of satisfaction answers - Q1 – How satisfied are you with your operated knee? Q2 – How satisfied are you with your rehabilitation? Q3 – Does the operated knee feel normal?

	Control		Wii			p-value
	Dissatisfied	Satisfied	Dissatisfied	Satisfied		
Q1	1 (3%)	30 (97%)	2 (5%)	35 (95%)		0.665
Q2	1 (3%)	30 (97%)	1 (3%)	36 (97%)		0.900
	Not normal	Normal	Not normal	Normal		
Q3	10 (32%)	21 (68%)	7 (19%)	30 (81%)		0.209

There was a low number of satisfaction assessments returned and as such the lack of statistical significance with question 3 may be a type II error. Of the total cohort of 136 patients, only 68 (50%) returned a satisfaction assessment. However, this was partly due to a clerical error in distributing them and once the assessments were administered, there were 68 returned out of a possible 110 (69%).

Table 19 - Table of falls in the two groups throughout the study period

	% Questionnaire returned	Control group	Wii group	P-value
18 weeks	Control 74% / Wii 81%	12 / 52 = 23%	8 / 56 = 14%	0.240
1 Year	Control 68% / Wii 63%	8 / 39 = 21%	7 / 38 = 18%	0.849

Table 19 illustrates that there was no significant difference between the number of falls experienced by the subjects in either group. None of the falls occurred while subjects were on the Wii balance board.

5.1.4 DISCUSSION

The aim of this randomized controlled trial was to compare the effectiveness of the Nintendo Wii-Fit and balance board in place of the usual rehabilitative care in terms of improving function in people who have had a primary Total Knee Replacement.

The primary outcome was the change in total WOMAC score from six weeks to eighteen weeks and then secondarily from six weeks to one year post-operatively. The ANCOVA analysis adjusted for baseline showed a statistically significant improvement in primary outcome for the Wii group compared to the control group at 18 weeks. The regression analysis demonstrated that the Wii group was significantly improved compared to the control group throughout the study period post randomization, which lasted out to one year post-operatively. Importantly, a variation of the Bellamy 20,50,70 method demonstrated clinical significance by showing that there was a significant difference in the numbers of patients in the Wii group compared to the control group who improved by more then 20% of baseline at one year in total WOMAC (p=0.093). This was also true for a 50% improvement in WOMAC score at 1 year (p=0.016). It is extremely uncommon for a rehabilitation study to show a significant difference in primary outcome to one year, many months after the intervention has ceased. This intervention did result in an enduring improvement although this has been shown using one of many available methods that try to demonstrate clinical significance. There is no consensus yet as to the best method to do this.

The review of the literature in chapter 2 (section 2.2) and the survey of Australian arthroplasty surgeons on p.21 (section 2.23) showed us that although the location of rehabilitation can vary in the six weeks following TKR surgery, from that point onwards, the majority of patients are usually discharged from rehabilitation therapy with advice to continue their exercises at home. The rehabilitation therapists usually give the patients written instructions and diagrams demonstrating which exercises they should continue. For the purposes of this randomized trial, this was replicated for the control group to represent 'usual care'.

The patients in the control group received usual care and were blinded to what was involved with the intervention for each group. Therefore, they only knew that they were being studied and had been given exercises to do. There was a concern that this could lead to increased adherence on the part of the subjects in trying to please the investigators by doing their exercises. The patients all knew that this study was focused on rehabilitation and was part of a PhD thesis. There were comments made from some patients in the control group that there was a desire to 'work hard' to help the study. This may have led to increased adherence to the standard exercise protocol, above the usual level of adherence to a standard protocol, observed in a group of patients not enrolled in a study. If this were the case, it could have biased the results in favour of the null hypothesis. However, it was not possible to control for this effect within this study design or to differentiate between the groups in terms of increased adherence due to trial participation. This could be the focus of future studies looking at results of patients following a standard rehabilitation protocol, knowing that they are in a control group.

The evidence suggests that this effect only lasts for the early stages of rehabilitation and, regardless of the artificial study setting, the adherence to a protocol tails off as the intervention period continues. This is described in Chapter 5.2 where adherence to exercise protocols by two groups was investigated.

The Nintendo Wii has aroused interest in various medical specialties since its introduction. The introduction of the Wii balance board and Wii fit software has led to a number of studies on its use in a musculoskeletal population.[289] This study is the first appropriately powered RCT to look at the effect of rehabilitating postoperative TKR patients with the Wii and is one of the few studies that has had an emphasis on balance exercises in the intervention and measured balance as an outcome measure.

Current evidence supports the practice of post-acute rehabilitation at home following TKR surgery.[34, 37, 43] Exercises which focus on balance as well as strength and flexibility can improve self-reported

outcomes[205] and intensive therapy may well provide more benefit if initiated six to eight weeks after the surgery.[39, 40] However, few trials concentrate on an intervention with a particular focus on balance as shown by the review of Minns-Lowe et al. [37] The TKR-POWER study had at its centre, an intervention with a large component of balance training. One of the original hypotheses of the authors was that by improving the proprioception of the knee and the overall balance of the patient, we may get closer to a replaced knee feeling normal. The data showed that more of the Wii group felt they had a normal knee but it was not to a statistically significant level. The 'Wii DSEC max %' outcome did show a statistically significant difference of 6.4% (p=0.009) at 18 weeks in favour of the Wii group.

The WOMAC score was similar in both groups at the preoperative assessment. At the week 6 assessment there had been a significant improvement in total WOMAC score in both groups. Both had improved by between 10 – 13 points. There was a difference between the two groups of 3.8 (p=0.089) points. This was neither statistically or clinically significant but was factored into the analysis because the patient group was a high functioning group. This was demonstrated by the post intervention period total WOMAC score of the controls which was 15.2. This is very close to the total WOMAC score of 14.2 of age-matched controls without significant arthritis.[66]

Secondary outcomes

This study included many secondary outcomes which were analysed for statistical significance. It is noted that the study was designed and powered to detect a significant difference in the mean difference of the WOMAC total score at 18 weeks and 1 year from the 6 week baseline. While some of the secondary outcomes did demonstrate a statistical significance to the p<0.05 level, this is without adjusting for multiple analyses making the results an interesting base from which to investigate further but on which no firm hypotheses can be confirmed or rejected.

Oxford Knee Score

Both groups also showed a significant improvement in Oxford score at 6 weeks from their baseline preoperative scores without any significant differences between the two groups. The disparity between the results of the WOMAC score and the Oxford score raises questions. It may be that the positive result from the WOMAC analysis is an example of type-I error. Similar studies that repeated similar positive results or found no difference would help discriminate this possibility. It is also apparent that the Oxford score measures different parameters to the WOMAC index. WOMAC goes into more detail in all three areas of its pain, function and stiffness subscales. WOMAC focuses 17 of its 24 questions on function whereas half of the Oxford score's 10 questions relate specifically to pain. The Wii rehabilitation was targeted at balance and functional improvement which may explain the larger effect as measured by WOMAC over Oxford score.

When designing the study, the Oxford score was considered as a primary outcome, not least because it is widely used and quick to complete. However, calculations which were backed up by the paper of Minns-Lowe et al suggested that in order to detect a difference of 2.5 on the Oxford knee score, the sample size for a study of this type would need to be 521 per arm. This score of 2.5 is not clinically significant based upon the mean clinically important difference when calculated using one half of a standard deviation with most studies putting the standard deviation around 10 (MCID 5).[91] This study was never powered to detect a difference in the Oxford score between the two groups, but the data is useful for comparison with other patient groups.

SF-12

The two groups both started with a SF-12 Physical Component Summary (PCS) score of 36, which is lower than the population mean of 50 (95% CI +/- 6.97). At 18 weeks, both groups had returned to within the 95% CI for PCS (Wii = 49.1, control = 45.6) and there was no significant difference between the groups at 18 weeks or one year. This improvement in the PCS score following a TKR is expected.

Preoperatively, their SF-12 Mental Component Summary (MCS) score was above average at 54.8 (control) and 55.8 (Wii) but within the 95% CI of +/- 6.24. There was a significant difference between the two groups at 18 weeks for the MCS with both groups having scores higher than the 95% CI but the absolute difference between their scores of 3.53 is unlikely to be clinically significant. This difference did not last to one year and the scores of both groups returned to within the 95% CI of the population mean. These results raise the question of whether the higher MCS score in the Wii group was from carrying out the Wii intervention itself or as a result of the improvement in their subjective result, as measured by WOMAC score, causing them to feel better.

VAS - Pain

There was no significant difference in the Visual analogue scale for pain score at 18 weeks. Both groups were reporting significantly reduced pain levels from the preoperative baseline and this continued out to one year where the control group scored 8.62 (11.26) and the Wii group 6.92 (10.77). While there was no statistically significant difference in VAS pain score between the two groups, the Wii group did have significantly lower pain scores as measured by the WOMAC pain subscale score than the control group. The WOMAC pain subscale focuses on functional pain in the knee whereas the VAS is a more generic score. Therefore, as the WOMAC pain subscale demonstrated, the Wii group were reporting less functional knee pain and the lack of significant difference in the VAS scores may have been due to pain from other sources.

Falls

There was a reduced falls frequency in the first 18 weeks which was not statistically significant. This study was not appropriately powered to detect such a difference. Any intervention that could be shown to reduce the frequency of falls would prove very clinically and economically important.

Satisfaction

There was no significant difference between the two groups when looking at satisfaction. Only 50% of the patients returned a satisfaction survey due to a clerical error during the early part of the study. This left the results of this questionnaire underpowered and open to type II error. Both groups expressed very high satisfaction rates with both their knee and their rehabilitation scoring over 95% satisfied or very satisfied. This is much higher than the commonly quoted literature values of 80% satisfied. This reduced the likelihood that the study would find a significant difference between the groups with such high satisfaction levels in the control group. The third question in the satisfaction survey asked if their operated knee felt normal or not. This question of a 'normal knee' was added because we had concentrated so much of the rehabilitation training on balance exercises. One hypothesis behind the decreased satisfaction levels in the literature is the loss of proprioception in the operated knee. Therefore, even if pain free, it may not feel normal. This may be a dissatisfying result for some patients. The study postulated that the emphasis on balance retraining could have improved proprioception in the operated knee, making it feel more normal, regardless of overall satisfaction with the knee. There was no significant difference between the groups (Wii 81% normal, Control 68% p=0.209). Future studies should concentrate on a higher return rate of satisfaction questionnaires and be powered appropriately for the 'normal knee' question.

Range of motion

The degree of knee flexion achieved by the patients was very similar to their preoperative average. This result was expected because patients tend to converge towards the mean after a TKR. Those with very poor flexion tend to achieve good gains in range whereas those with a very high degree of flexion tend to lose flexion postoperatively. At 6 weeks, both groups had lost flexion from the preoperative average of 108-111 down to 98-99 degrees. At 18 weeks they had both returned to levels very similar

to their preoperative baseline. Both groups did continue to make gains out to one year, which is expected with the control group finishing at 116.9 degrees and the Wii group at 117.9 degrees.

Knee strength

There was no significant difference in strength of knee flexion of extension between the two groups at any point. This was not surprising as neither of the rehabilitation protocols involved any exercises using significant resistance. The lack of resistance training is an area that could be addressed in future studies, incorporating it into a Wii protocol. This does often require further equipment and training which could reduce adherence and incur expense. It could also be argued that in this patient population, the resistance involved in these exercises seems to be enough. Most exercises are using a variable amount of the patient's own body weight to generate resistance. By virtue of the total WOMAC score returning to the levels of age-matched controls without arthritis, either protocol does seem to return the patient's strength to the required level for their age.

Up and Go test

There was no significant difference between the two groups for the time taken for the Up-and-go test (UG) at 18 weeks or one year. The UG test combines strength and balance in a functional assessment and it was hypothesised prior to the study that the UG test may have demonstrated a difference. However, as discussed previously, the population being studied was functioning at a high level and the UG time of the control group (6.7 seconds) at 18 weeks compares very favourably with the literature. The group of community dwelling non-fallers in the study by Rose et al had an average age of 76.7 years and averaged 7.1 seconds (s.d.1.4) for the UG test.[238] While this group is 4 years older than the control group (mean = 72.3 years) in the TKR-POWER study, they are community dwellers who could walk without assistive devices. Therefore, it would seem that the lack of a significant difference is due to the patient population already being at such a high level that improvement would be very difficult to achieve using this test.

Balance testing

The assessment of balance was described in Chapters 2 and 4. It is a complicated process with many different testing options using a variety of instruments from a simple pencil attached to a belt through to a very expensive laboratory-grade force platform. The work in Chapter 4 detailed the use of the Nintendo Wii balance board as a force platform using the Wii-fit software to analyse the data. For the TKR-POWER study, the patients were assessed using the Wii-fit as a novel tool for assessing balance. There was no significant difference between the groups at preoperative baseline or at 6 weeks for the eyes open or the eyes closed tests. At 18 weeks, the Wii group showed a statistically significant difference in the eyes closed stance scoring 51.4% against the control group who scored 45.0% (p=0.009). The difference between the groups was 6.4% but this difference was not present at one year. There is a possible effect from the familiarity that the intervention group would have had with the Wii balance tests because they were using the equipment regularly at home whereas the controls didn't.

The validity study for the Wii-fit Stillness score in chapter 4.1.3, showed that the Stillness score was a measure of total amplitude in the 'x' and 'y' axes. While amplitude has been used as a validated measure of standing balance, it is usually reported in the separate axes with standard deviations. The Wii-fit software does not offer this option. Therefore, the Stillness score offers limited scope for analysis of any balance deficiencies found during testing. In order to maximize the utility of the score, as per the findings of chapter 4.1.3, the TKR-POWER analysis attempted to eliminate outliers by taking the best score of three.

In chapter 4.1.3, the standard error of the measurement was found to be 6.96%. Therefore, this could suggest that the difference found in the TKR-POWER study, while statistically significant, are right on the edge of the measurement error of the Wii-fit and Wii balance board combination. It is also important to note that when validating standing balance measurements, the results are often specific

to that patient population. The validation in chapter 4.1.3 was in healthy young volunteers and so the results may not be applicable to the patients with arthritis in their seventies.

While the results show a statistically significant difference in Wii DSEC standing balance between the groups at 18 weeks as measured by the Wii Stillness score (p=0.009), this analysis was not adjusted for multiple analyses and the clinical significance of this outcome measure in this population is still not well understood. There needs to be more research before clinical meaning is ascribed to this finding.

Adherence

The difference in adherence to the rehabilitation protocols is covered in Chapter 5.2. It would be of interest to attempt a correlation of the intensity of exercise in this trial compared to other studies in the literature. Unfortunately, the interventions are so diverse that the analysis would be difficult to run in any meaningful manner.

STUDY STRENGTHS AND WEAKNESSES

The trial methodology was designed in accordance with CONSORT criteria to minimize bias where possible. The participants were assigned to their groups by random, concealed allocation with an audit trail. Given the nature of the intervention however, it was not possible to provide a true placebo control. To minimize bias, the outcome assessors were blinded to the group allocation and the participants were blinded to the nature of each intervention group. Consequently, neither group knew if they were control or intervention and only the Wii group knew about the Wii once they were randomized to it. The assessors and the subjects were reminded throughout the study not to discus the type of exercises they were doing. All questions relating to the exercises or the Wii were directed to the lead author. The statistical analysis was blinded to treatment allocation and in accordance with 'intention to treat' methods.

The randomization process was not stratified. The sample size was deemed large enough to account for all other variables of concern. Analyses were adjusted for confounders and covariates deemed to be potentially important such as age, BMI and flexion. The baseline characteristics of each group were consistent with those reported elsewhere and there were no significant differences between the groups. The baseline variables were found to improve significantly at all postoperative assessment points, in line with what is expected after TKR surgery. Despite this, the analyses of the primary outcome WOMAC total score were adjusted for baseline values according to recommendations in the peer-reviewed literature (see chapter 2, section 2.3.2.1a).

The intervention group received a home visit from the study author to install the equipment. It would be interesting to assess the willingness of this population to use the equipment were it not set up for them. The results of the CIGAR survey demonstrated that a high proportion did not feel comfortable with gaming technology and as the aim of the study was to see if performing the exercises themselves improved function, adding the extra step of setting up may have decreased adherence and reduced the chances of answering the study question. The study of the equipment's application in the real

world, including set-up, is a potential future study. The effect on adherence of a home visit could have been mitigated by a home visit to the control group to show them the exercises. The exercises were demonstrated by the lead author, in an extended appointment at the research unit instead. The Wii group did not receive this extended appointment.

The follow-up stopped at one year as it has been shown that there are few functional improvements in TKR patients beyond that time.[229, 340, 341] However, some of the patients were still exercising up to and presumably beyond one year. Data were collected at the one year assessment on the levels of exercise in each group beyond the end of the intervention period. This will prove the basis of further analysis of the effect of the Wii intervention on continued adherence to an exercise protocol which is investigated further in chapter 5.2.

The WOMAC osteoarthritis index was used as the primary outcome as it has been validated and shown to be clinimetrically relevant as a functional outcome assessment in the total knee replacement population. It can be subdivided into pain, stiffness, and functional subscales, allowing for further sub-analysis of each patient. It has been used extensively in the literature, allowing for comparison with other studies.

The data collected looking at contralateral leg osteoarthritis was not analysed. This is known to have a potential effect on the overall score and although the randomization process should have accounted for any differences, that data will be worth investigating.

It has been shown by some authors that there is discord between objective observed functional disability and subjective patient-reported outcome measures.[342, 343] However, Wittink and colleagues only used the SF-36, which is a general health-specific and not a disease-specific measure.[343] The conclusion that both objective and subjective measures should be assessed is valid. It must also be borne in mind that the evidence shows the best predictor of the postoperative subjective score is the

preoperative subjective score. [13, 67, 228, 269, 270] This can be mitigated somewhat by the use of relative change scores and the analysis of levels of response such as the WOMAC 20,50,70 method.[81]

The latter is a method for calculating discriminant validity, an area which has proven to be controversial over the years. There is no agreed, valid method for calculating a minimum clinically important difference with many variants and formulas.[59, 64, 71, 73-76, 78-80, 344, 345] The 20,50,70 method allows for 'a categorization of individual patients according to whether they achieve levels of improvement at or above prescribed response thresholds.'[81]

The data from these patients needs to be further examined to assess the potential for a patient selection tool. As mentioned earlier, the ability to identify the 20% who will be unsatisfied before their operation would be a valuable tool for directing further preoperative management. Such data may help towards that goal. They may also help to identify those patients who would benefit most from extra supervision from a physiotherapist. This could limit the need to provide supervised physiotherapy to all patients, enabling it to be targeted just at those who need it. This would save resources and increase the benefits to those most in need.

There is a discrepancy between the positive results of the WOMAC Total score and the lack of significant difference in the Oxford knee score, knee strength or balance amongst other secondary outcomes. It has been noted earlier that the TKR-POWER study was only powered to detect a difference in the WOMAC score and the lack of difference in these other measures may be a type II error. It may also be that the difference in WOMAC total score is a type I error. It would be sensible to perform more studies looking at this intervention in a TKR population before firm clinical recommendations can be made. These studies could be powered using a post-hoc analysis of the data form this study to examine more of the secondary outcomes as well as confirming or disputing the WOMAC total score findings.

The overall goal is to use this Wii technology to allow the majority to maximize their outcomes with the most cost-effective use of resources. It could reduce patients' time and transport costs and the costs to the healthcare system. However, it is important to have a system which can identify before or during the rehabilitation process the patients who need a more careful eye to be kept on their progress. That analysis has not yet been possible.

Wii equipment is widely available in retail outlets for a modest price and can be reused or shared once patients have finished with it. If effective, it could optimize outcomes for patients and further reduce the need for costly supervised outpatient rehabilitation in this group. It is also possible that because of the interactive and personalized nature of the intervention, it can improve compliance with and adherence to prescribed exercise treatment.

The results of this study are widely applicable in the developed world. The technology is available in most countries and age has been shown to be no barrier with the oldest person in the study being 88 years old when recruited. The cost is within the realm of many patients and should a cost-analysis prove beneficial to the healthcare system, it may be available through hospitals or insurers. The cost analysis could include investigate a range of durations of the intervention period through varying the exercise intensity and types of exercise recommended.

The development of software tailored to this population and specific to their rehabilitation needs is now required in order to maximize the potential patient benefits and applicability.

As part of the study, more data were collected on anxiety, depression, and locus of control. The mental state of the patient is an important but poorly understood factor in recovery. Self-efficacy in the early post-op period has been shown to be a better predictor of long-term outcome than preoperative self-efficacy.[346] It will be the task of further work to see if pre- or postoperative mental state and locus of control predicts outcome in a similar manner. This may also be of benefit to a predictive model of those patients who are not satisfied with the results of their surgery.

The TKR-POWER study was sufficiently powered to detect a clinically meaningful benefit as demonstrated by the change in the WOMAC total score. It has shown that the Wii-Fit can improve outcomes for patients after their knee replacement and that according to the regression model these improvements continue to be present up to one year compared with the control group. The statistical analysis showed some variability in the results depending on how they were analysed. The one-year results were significant when the mean difference scores were compared using ANOVA but not when adjusting for baseline using ANCOVA. Importantly, the regression models took all preoperative and baseline WOMAC scores into account and compared the groups throughout the study period and found the Wii group to be significantly improved as measured by total WOMAC score throughout the intervention period to one year post-operatively. This indicates the importance of the model's analysis and taking confounders into account.

There are many ways to assess a minimum level of clinical difference or importance (see Chapter 2). However, the subgroup analysis of the WOMAC score using the 20,50,70 method of Bellamy and colleagues demonstrates that a greater proportion of patients in the Wii group achieved a 20% and 50% improvement in their function over baseline compared with the control group.

The intervention has been shown to be safe with no major clinical incidents reported by the patients. There were no falls off the balance board. There was no statistically significant difference in the number of falls experienced by subjects in either group during the study period. There was a difference at the 18 week follow up in favour of the Wii with only 14% experiencing a fall compared to 23% of the control group. This is an area to investigate further by powering a similar study specifically for falling and also assessing the other outcome measures of the fallers compared to the non-fallers.

The results support the study hypothesis that the Wii is safe and effective for use in the home for this patient population and using it for three months improves functional outcome as measured by the WOMAC score.

The study was designed and carried out using an intention to treat analysis. Of the patients who withdrew from the study, the majority stopped performing the exercises specified by the study for health or personal reasons but still continued to attend their assessment appointments when requested. All subjects with data were included in the final analysis although there were no statistical methods employed to account for missing data. Future work is planned to analyse the patients that withdrew to check for any significant differences from those who continued. It is also planned to run analyses imputing any missing data where appropriate. This was not felt to be critical to the initial data analysis as the groups were evenly matched for withdrawals not attending for assessments at eighteen weeks (Control 3% vs Wii 1.5%) and at one year (Control 14.7% vs Wii 11.8%).

There was a high level of patient satisfaction in this study and there was no significant difference between the groups. Some 89% were very satisfied with their operated knee in the Wii group and 77% were very satisfied in the control group. However, if the patients who were either satisfied or very satisfied were combined, the satisfaction levels were above 95% in both groups. This is significantly higher than the satisfaction rates reported in the literature and so high that demonstrating a difference between groups would be very difficult. Therefore it is not surprising that none of the satisfaction questions relating to the operated knee or the rehabilitation process were discriminatory. The third question asked if patients' operated knees felt normal. The control group felt their knee was normal in 68% of cases compared with 80% of the Wii group. This was not a significant difference. Only 68 satisfaction surveys were received because of an administration delay in getting them to the first 50 patients in the study. A future study should be powered appropriately to ask this 'normal knee' question .

5.2 ADHERENCE TO A TKR REHABILITATION PROTOCOL

5.2.1 INTRODUCTION

Historically, up to 50% of patients do not follow the recommended regimens in trials investigating adherence. This is similar to physiotherapy and drug trial rates. [103-106] There have been many attempts to increase adherence, with mixed results. A study of two differing rehabilitation protocols following TKR surgery had a loss to follow-up of nearly 50% in the first year and the majority of these patients had low motivation levels when assessed during their immediate postoperative inpatient rehabilitation. The patients were lost to follow-up as they either refused to continue or were not contactable.[32]

Campbell and colleagues explored non-compliance (non-adherence) with physiotherapy in patients with osteoarthritis of the knee.[109] Only 8 of 20 subjects were fully compliant at the 2-5-month stage post-surgery and the rest only partially. The patients tended to continue with the exercises which they found easiest or felt were of the greatest benefit. They also demonstrated a bimodal pattern of adherence with an initial phase when still attending physiotherapy sessions and then a later stage. The initial stage was easier to explain, patients describing an obligation to the therapist, altruism towards the researchers, and a direct benefit to themselves in a period when they were still recovering from pain and stiffness. The reasons for continuing adherence were more complicated. Campbell raised the following themes.

There was a tendency towards a positive attitude towards the exercises on the part of those patients who could incorporate them into their regular routine. Patients who stopped the exercises talked of a number of difficulties in doing them alone and a perceived lack of progress or improvement. Those with more severe knee symptoms tended to continue their exercises. Patients who saw the causes of arthritis as immutable were less likely to exercise than those who saw the intervention as a possibility to minimize its impact.

There was a close relation between high levels of adherence and those who saw improvement and hence felt the intervention to be effective. This study aimed to investigate levels of adherence in patients rehabilitating from TKR surgery. The difference in adherence was assessed between the group using the Nintendo Wii and a control group.

5.2.2 METHODS

The 136 patients who were participating in the TKR-POWER study completed exercise diaries to record their exercises. TKR-POWER was a prospective, parallel-group, double-blinded, randomized clinical trial registered with the Australian New Zealand Clinical Trials Registry (ACTRN12611000291987). It was conducted according to the extension of the CONSORT (Consolidated Standards of Reporting Trials) guideline for clinical trials. [334]

Ethical approval was obtained from the Northern Sydney Central Coast Area Health Service Human Research Ethics Committee (Harbour) (EC00333) (protocol no: HREC/10/HAWKE/123), after submission of the National Ethics Application form (http://www.neaf.gov.au/), and the University of Sydney Ethics Committee (Protocol No: 13996). All participants were required to provide informed consent prior to their involvement in the study.

Study Participants

Potential participants were screened and recruited at the rooms of the three operating surgeons at the time a decision was taken to undergo elective TKR.

Inclusion criteria were: 1) over 18 years old; 2) primary unilateral or bilateral TKR; 3) English speakers. Exclusion criteria were: 1) Previous septic joint; 2) revision surgery; 3) inability to return for all extra follow-up visits; 4) dementia, Parkinson's, or other medical conditions severely affecting balance.

Timeline

The trial ran from March 2011 to September 2012. There was no change to the usual medication or rehabilitative care during the peri-operative inpatient or postoperative period up to six weeks after the operation.

Randomization and blinding

For details of randomization and assessments see Chapter 4.

The patients were blinded by not knowing the details of the two rehabilitation strategies being employed in the study. Therefore, neither group was aware of what the other group's rehabilitation consisted of or whether they were in a control or intervention group. The use of the Wii was only known to the Wii group and they did not know that the other group were not receiving the Wii. The outcome assessors were blinded to the treatment allocation. Both patients and assessors were instructed not to discuss the exercises or allocation.

Diary monitoring

Both groups maintained exercise diaries. All patients were given a diary to record the frequency with which they performed the routine exercises prescribed by the rehabilitation facility on discharge from outpatient physiotherapy. There were nine possible exercises in the standard diary, each of which could be completed a maximum of seven times a week. There was no division into different phases or change of intensity. In the Wii diary, there were 13 exercises in phase 1, 16 in phase 2 and 19 in phase 3, each for a maximum of seven times a week. Each 4 week phase had the option of adding progressively more challenging Wii exercises for the Wii group. (See Appendix viii)

The diary data was analysed from the ninth week of the trial, which was the third week of the intervention. This was because some patients were delayed getting to the clinic to collect exercises or receiving their Wii-Fit hardware in those first two weeks.

The intervention group also had diaries to record the frequency of the exercises performed on the Wii. They had three diaries which split the intervention period into four-week phases. Each phase added extra exercise choices to those already prescribed. The difficulty of the exercises increased as the patient progressed through each phase.

Each diary consisted of a column allocated to each prescribed exercise. Each row corresponded to one week of the 12-week intervention period.

For all diaries, the subjects were asked to record each exercise performed that day with a tick in the corresponding box. There was a maximum of one tick per day and seven ticks per box to indicate whether a particular exercise had been performed every day that week.
At each assessment, the research staff asked about any extra ongoing physiotherapy and any falls or incidents since the last assessment.

Data analysis

The data were checked for normality with Shapiro-Wilk statistics.
The mean number of exercises per week was compared between groups specifically in terms of :

- overall adherence between groups

- adherence over time

- particular favourite exercises

Adherence was assessed as recording a minimum of four exercises per day four times a week.

Statistical analysis

The data were tested for normality and then analysed using Student t-test.

5.2.3 RESULTS

Table 20 – Number of patients recruited and rate of diary completion and return

	Control group	Wii group	
	Standard diary	Standard diary	Wii diary
Number of patients	68	68	68
Diaries returned	37	42	44
Blank diaries returned	1	17	4
Rate of return	54%	62%	65%
Rate of completion	53%	37%	59%

Of the 136 patients in the study, 81 (60%) returned a diary. The rate of return was slightly higher in the Wii group (62% vs. 54%) for the standard diary and even higher for the Wii exercise diary (65%). The rate of completion of a diary was higher for the Wii group (59% vs. 53%) compared with the standard diary of the controls.

There were more blank standard diaries returned by the Wii group than by the control group. This was expected as the Wii group had been told that they did not need to complete any of the standard exercises (Table 20).

The mean number of exercises completed per week for each group was normally distributed (Shapiro-Wilk p=0.574 and 0.717).

The average number of times each exercise was completed per week over the course of 10 weeks was 5.6 (s.d. 0.38) for the Wii group and 3.6 (s.d. 0.59) for the control group (p<0.001) (see Figure 17).

The average number of times each exercise was performed per week was markedly different throughout the intervention period. The Wii group started with a higher average number of each exercise performed per week at 5.6 (s.d. 0.59) compared with 4.2 (s.d. 0.4). The difference between the groups became larger as the trial continued. The number of each exercise performed per week dropped off in the control group in the latter half of the trial where the Wii group maintained their rate to the end.
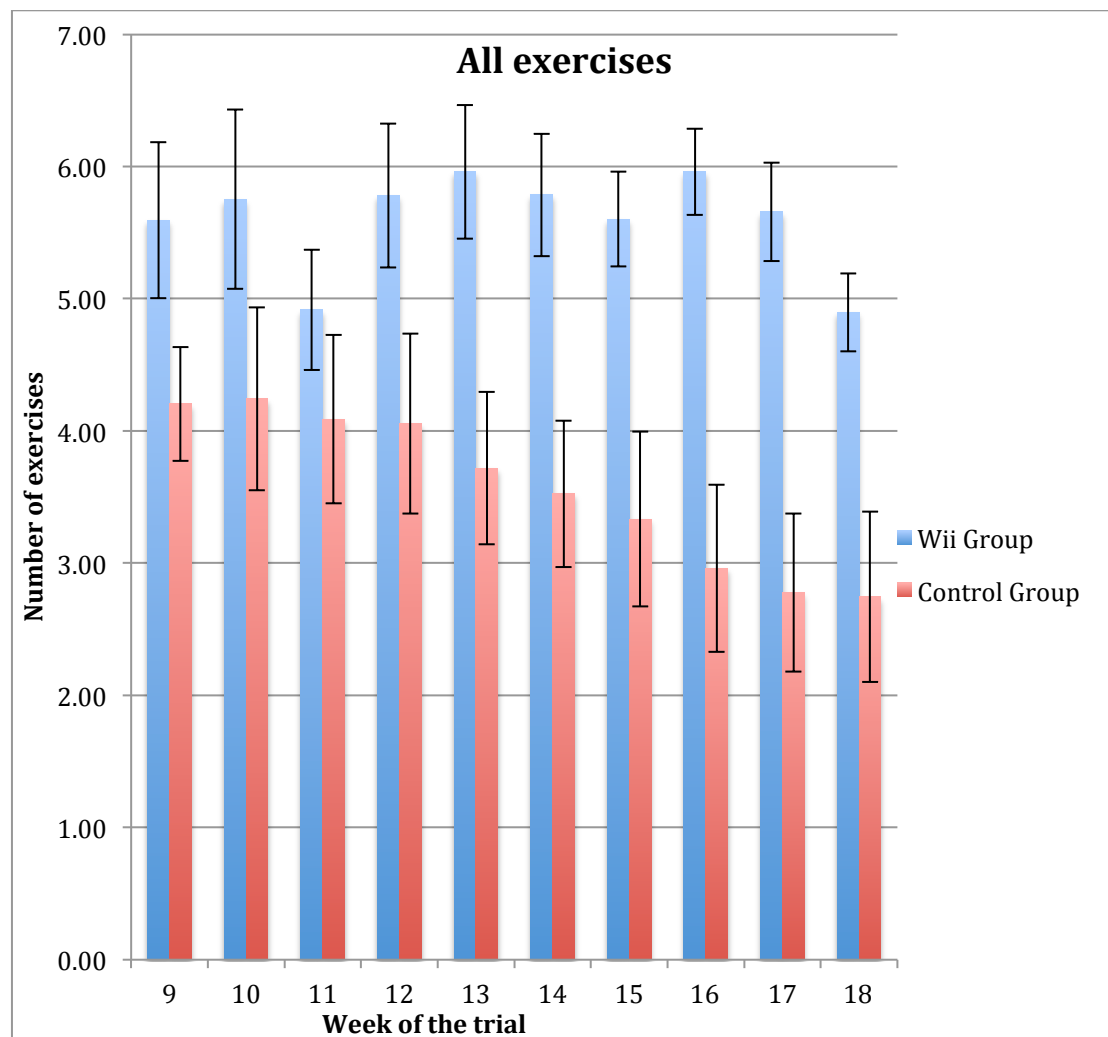
Figure 17 - Graph comparing mean number of exercises completed per week. Error bars +/− 2 standard deviations

Table 21 - The most and least popular Wii exercises in the trail during each phase

Phase 1		Phase 2		Phase 3	
Most popular	Least popular	Most popular	Least popular	Most popular	Least popular
Deep breathing (Y)	Hula hoop (B)	Ski slalom (B)	Bridge (Y)	Tilt city (B)	Balance bridge (M)
Half moon (Y)	Stillness (B)	Balance bubble (B)	Jogging (A)	Single leg extension (M)	King of the dance (Y)
Warrior (Y)	Cycling (A)	Penguin (B)	Supa Hula (B)	Balance bubble (B)	
Torso twists (M)	Parallel stretch (M)	Sun salutation (Y)	Driving range (B)	Palm (Y)	
Table tilt (B)		Chair (Y)		Spine extension (M)	
Rowing squat (M)		Side lunge (M)		Lunge (M)	

Table 21 shows the 6 most popular and the least popular exercises performed on the Wii according to the Wii diaries. (Y = Yoga, M = Muscle, B = Balance, A = Aerobic) As the 12 week intervention period progressed, the Wii group moved from more Yoga and balance exercises to strength-based muscle exercises.

Figure 18 shows the number of times that the six most popular exercises were performed by each group. The difference between the two groups is significantly different with the Wii group averaging 5.68 (s.d. 2.64) exercises per week and the control group averaging 3.73 (s.d. 0.72).
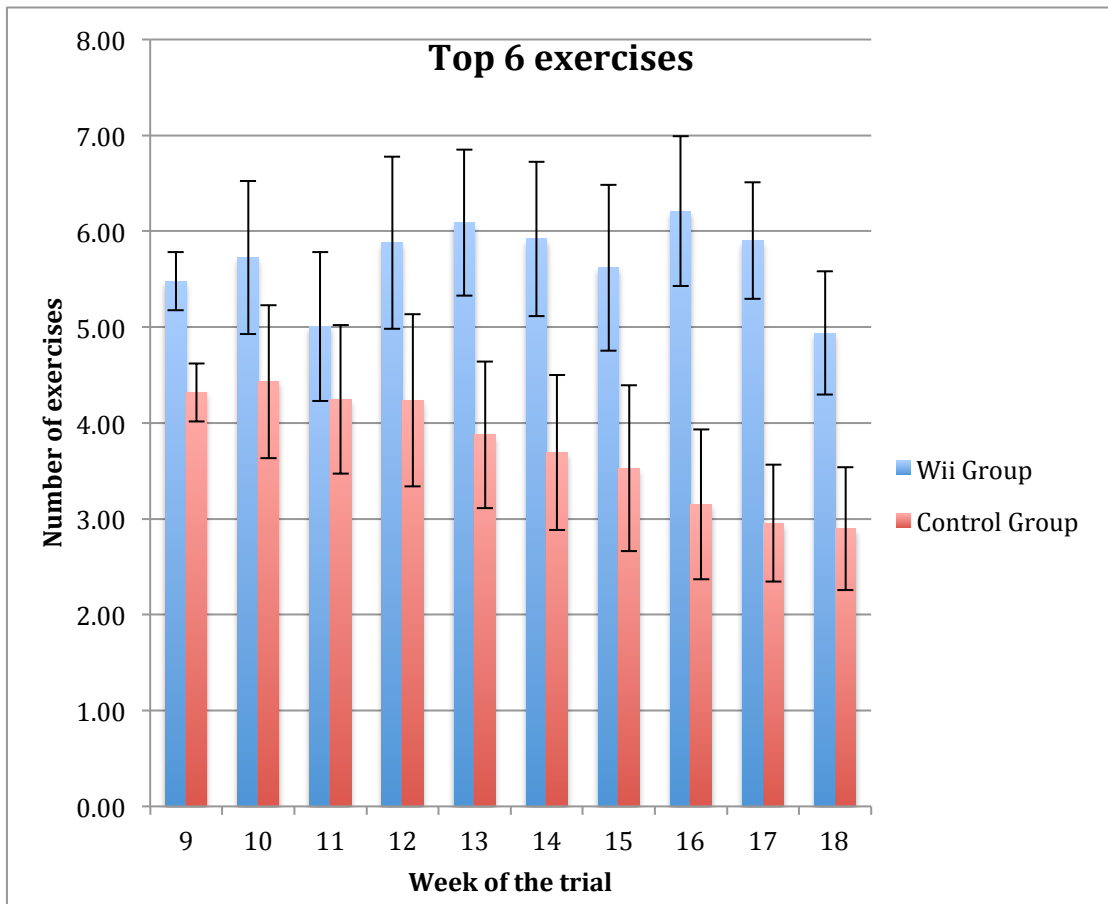
Figure 18 – Graph comparing average number of top 6 most popular exercises completed per week between groups Error bars +/- 2 standard deviations
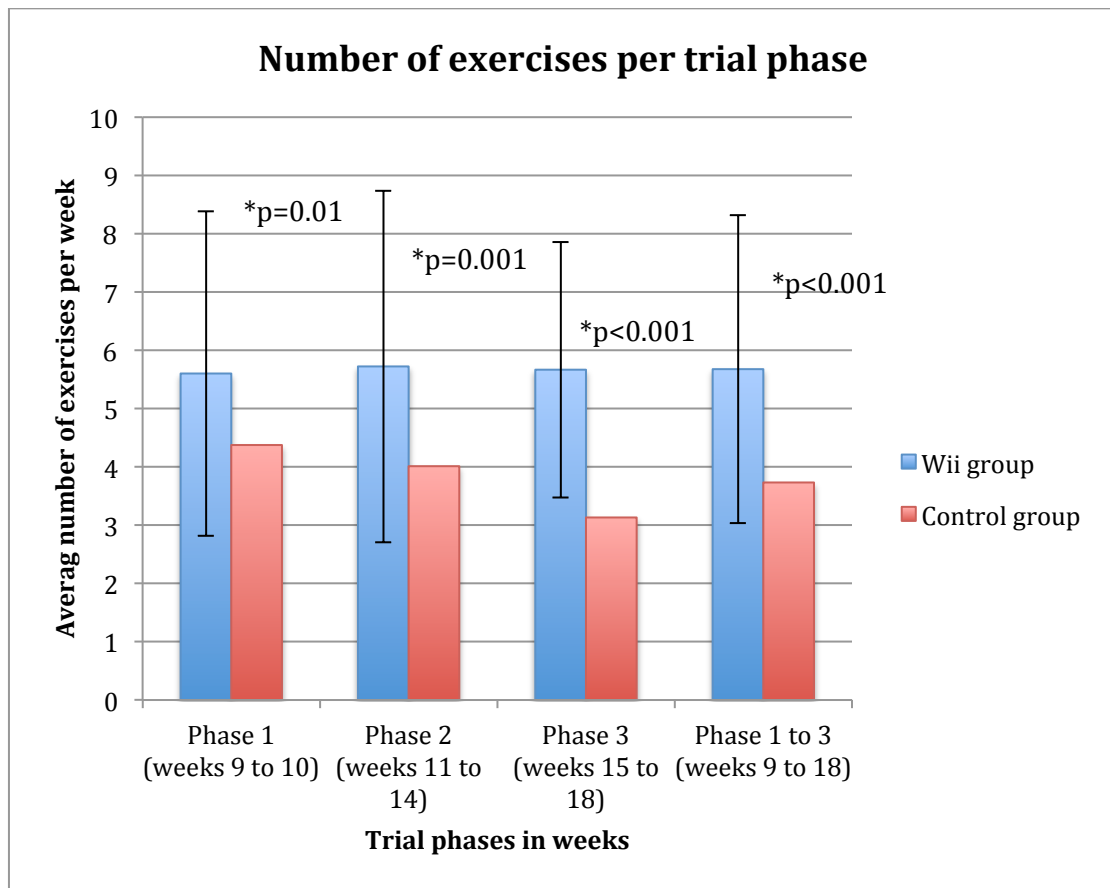
Figure 19 – Graph comparing the number of exercises completed per week during each of the trial's phases

There was a significant difference between the groups throughout all phases of the trial for the average number of exercises performed per week. This is shown in Figure 19 where the groups are compared in each of the three phases of the intervention period as well as overall. The differences between the groups are significantly different in each analysis.
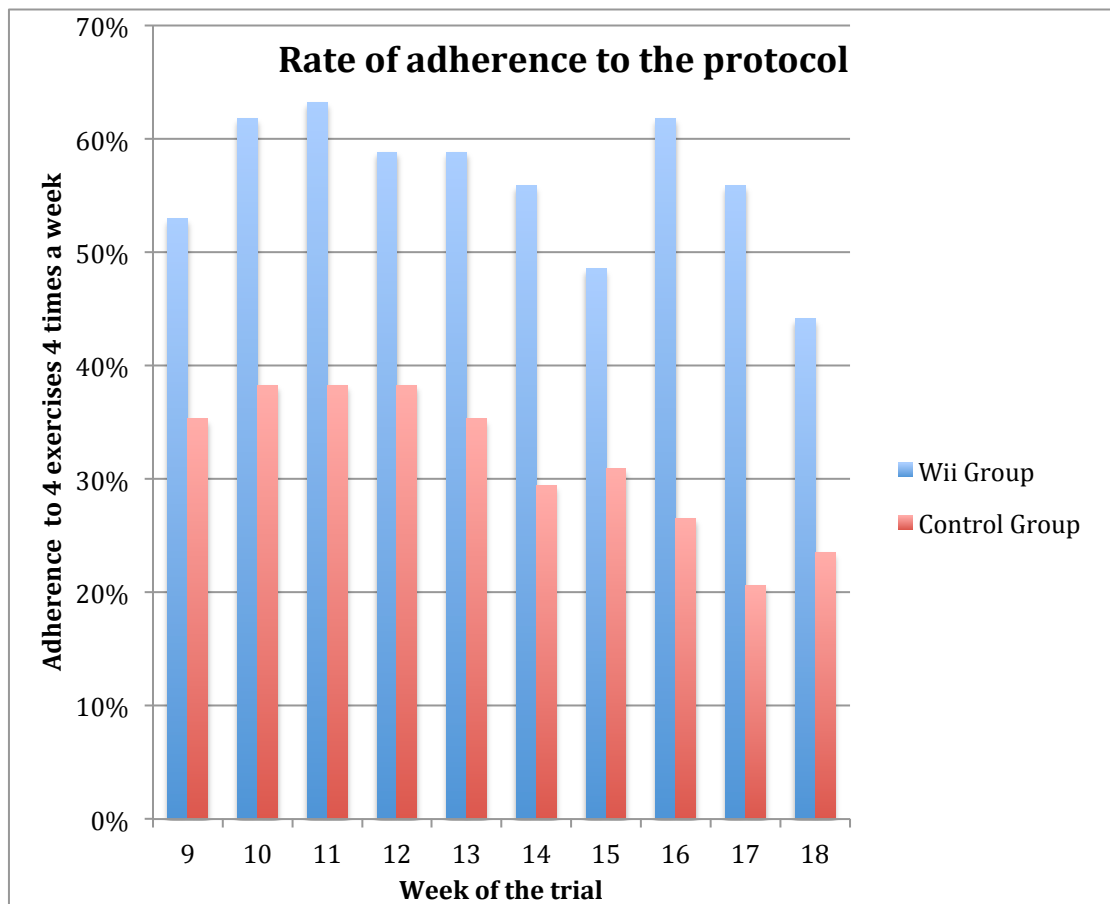
Figure 20 – Graph comparing the rate of adherence to the protocol minimum of 4 exercises per day 4 times a week

The percentage of patients adherent to the target of four exercises a day four times a week was different between the two groups. The Wii group maintained an adherence rate of between 20 and 35% higher than the control group, especially in the later phase of the trial (see Figure 20).

5.2.4   DISCUSSION

This study looked at the adherence to two different rehabilitation protocols within the TKR-POWER study. One group had written instructions on 9 standard exercises that remained unchanged for the 12-week study period. The other had a 3-phase written protocol with increasing intensity for exercises to be carried out on the Wii. They were both instructed to carry out 30 minutes of exercise a day minimum and had no other restrictions.

The results show that not only did the Wii group perform significantly more exercises over the 12 weeks but they continued to increase their rate of exercise beyond the peak of the control group and maintained the increased rate for longer.

The number of exercises is a surrogate for time spent exercising. Each of the Wii-Fit exercises takes several minutes to carry out. This equates well to the time needed for each of the exercises on the control group's list. It is sensible to compare the number of exercises from each as well in terms of time.

In terms of total time spent exercising, both groups were given the same instructions, i.e. to exercise for 30 minutes per day. The exercise sheets for each group were designed to take 30 minutes to complete if each exercise was completed once with the directed number of repetitions. The patients could only tick the exercise once, so if they repeated the exercise that day the time spent was not reflected in the diaries. This was to simplify the data collection and maximize diary completion. Some patients did actually tick the boxes more than the maximum seven.

Adherence was established as completing a minimum of four exercises a day four times a week. This allowed for patients taking some days off while still maintaining a regular pattern of exercise. It represented an exercise rate of more than once every second day. The Wii group were significantly more adherent to this level than the control group by between 20 and 35%. One reason for this could

be that the Wii group had a greater variety of exercises to attempt. The diaries recorded a higher number for those performing many different exercises just once compared with those repeating a single exercise over and over again.

However, the instructions were to perform each exercise only once every day as this would take up the suggested 30 minutes. The rate of performance averaged between three and six times per week, suggesting that the patients were not doing every exercise once a day and were often not completing some exercises at all in a week. Therefore, the adherence rates are likely to reflect a real increase in the volume of exercise performed.

Compared to the literature, the adherence rates for the Wii group were similar to the home exercise group of Piva et al which were 64% and 67% for their two groups.[205] Ko et al received 73% of their exercise diaries from all participants although two of the three groups were supervised. They reported moderate adherence to the home-based exercises with 65% of the monitored home program performing >80% of the required sessions.[347]

There was a relatively low return rate of diaries. It will be a focus of any future studies to ensure a higher rate or diary return.

The popularity of the Wii exercises are demonstrated in Table 21. The popular exercises in each phase either fulfilled the category of 'seeming to give some understandable benefit' or being fun. The warrior pose and torso twists for example, make the patient work the legs and the trunk in such a way as to feel the benefit. The hula hoop and the supa-hula are complicated movements which feel much more like games and are likely to be frustrating as they are difficult to complete well. The cycling and jogging games involve walking on the spot and anecdotally many patients expressed a preference to walking outside if that was the exercise of choice. Finally, some exercises were just too hard. The king of the dance exercise is very difficult, as it involves extreme levels of balance, flexibility, and strength. This is difficult for the healthy younger age group, let alone a postoperative group.

The Wii-Fit provides variety and a stepwise increase in difficulty of available exercises. Although it could be argued that the control group did not have a stepwise increase in the difficulty of the exercises provided, the study was interested in replicating real-world rehabilitation. When patients are discharged from outpatient rehabilitation, often at six weeks, it is common that they have nothing beyond the exercise sheet given to them on discharge. The study looked to test the hypothesis that by offering them more variety and the option of increasing the difficulty of the exercises, it would encourage them to maintain adherence to their exercise protocol for longer.

CHAPTER 6 - Discussion

6.1 SUMMARY OF THE AIMS OF THIS THESIS

The objectives of this thesis were to evaluate the Wii-Fit as a balance testing tool and a rehabilitation adjunct for patients to use at home following total knee replacement (TKR) surgery with the overall objective being to improve outcomes and patient satisfaction levels. The pilot studies aimed to quantify and validate what the Wii-Fit Stillness balance test measured, test for associations between balance and function, and finally assess the use of technology by the older age group which would be involved in the clinical trial. The TKR-POWER clinical trial compared rehabilitation using the Wii with usual care and the adherence trial quantified the volume and specifics of the exercises performed by the subjects.

Although primary TKR surgery has been shown to be very cost-effective, up to 20% of patients reported being dissatisfied at the one-year follow-up. One avenue of investigation for improving outcomes is rehabilitation.

The literature review raised many interesting points about the current practice of rehabilitation following knee surgery. There is little evidence-based consensus on the ideal type of exercise, and the timing and quantity of rehabilitation. There is a gathering weight of evidence to suggest that the location where the rehabilitation takes place and even the level of supervision is not critical as long as the patient is doing the exercises. However, there is emerging evidence that adherence with exercise regimens is poor.

Balance is a relatively neglected function in the assessment and rehabilitation of TKR patients which becomes highly relevant when considering the increased falls risk of this group (see chapter 2.3.2.2b). Although balance training is easy to incorporate into a protocol, especially with a Wii, it is unclear how best to measure and interpret balance in this population.

This study aimed to validate the Wii-Fit as a balance-measuring device which could be used quickly and cost-effectively by clinicians when assessing patients.

The validity of the Nintendo Wii-Fit software as a measure of standing balance was first assessed in healthy volunteers. The Stillness score was measured in patients with severe osteoarthritis of the knee who were awaiting a total knee replacement. The results were examined for any associations with preoperative function with the aim of identifying people who should be targeted given that preoperative function is a good predictor of better postoperative outcomes.

The TKR-POWER study aimed to assess the effectiveness of the Wii-Fit software as a rehabilitation tool when used in conjunction with the Wii balance board. The patients had full use of the Wii at home for three months and a raft of outcome measures were recorded throughout their recovery until one-year post-surgery.

The level of adherence to the given rehabilitation protocols was recorded to compare the two groups and assess the preferred exercises of the patients as the intervention period progressed.

6.2   SUMMARY OF THE RESULTS AND INTERPRETATION

The assessment of standing balance is complicated by a lack of standard measurements and unclear implications for the findings. The evidence is mixed as to the level of effect which standing balance, as measured on a force platform, has on the functional abilities of these patients. The Wii balance board has been shown to provide a measure of centre of pressure which is similar to a force platform in the horizontal plane (x and y) but not in the vertical (z) plane, albeit with reduced accuracy because of limitations of hardware and software. This limits its use as a force platform and makes it more of a posturographic tool for measuring stillness or sway. The Stillness measure used by the Wii-Fit software was shown to be a reliable measure of standing balance in healthy individuals in three of four tested stances. The eyes-closed stance standing on one leg could not be tested with Stillness as the time parameters of the test are set to 30 seconds. Most of the subjects could not maintain the

stance for this long. Therefore, even if it could be shown to be a reliable measure, it cannot be a valid test for standing balance if tested for over 30 seconds.

The validity of the Stillness score in the three tested stances is limited. The algorithm used to calculate the score sums the maximum range of CoP coordinates in the x and y planes. Although this measure has been used in the literature, the mean value and standard deviation for each of the separate planes are generally described individually.

Standing balance measurements can show the majority of movement occurring in either the 'x' or the 'y' axis. This results in larger CoP ranges and standard deviations in one direction over the other. Each type of balance movement can be due to different pathologies, which can help to target intervention for specific pathologies.

The literature also often demonstrates that the results are only significant in one plane for certain patient populations. The Stillness score cannot differentiate the two planes, summing them together to create a score. The measurement of maximum amplitude is described as a poor measure of balance and without the ability to discriminate the components of the two axes, then at most, it could only be recommended as a screening tool. In a situation where a deficit was picked up, the clinician would need to investigate the patient in more detail using a force platform and software combination with the ability to measure specific validated tests allowing for more detailed analysis of data from both axes.

The Stillness balance measure for a person standing on both legs with eyes closed was not associated with function at the preoperative stage or predictive of function at one year postoperatively in a population with severe end-stage osteoarthritis of the knee necessitating TKR. Therefore, in this population, the use of the Wii-Fit Stillness cannot be recommended as a tool for screening or directing management. It may be that there is an association with different stances and that the data need further investigation.

When the plans for asking patients in their mid-70s to rehabilitate using the Wii-Fit at home for three months were initially considered, concerns were expressed about their ability and willingness to use the technology. The Computers, Internet and Gaming in Rehabilitation (CIGAR) questionnaire allayed some of those concerns. It illustrated that of this group of 155 patients the 128 respondents showed a good baseline level of technical ability. This was demonstrated by the 60% who were comfortable using a computer, the 67% who used the internet daily, and the 75% who checked their emails daily. However, exposure to gaming systems such as the Nintendo Wii and the Microsoft X-Box was limited to approximately 10% of the respondents. There was ongoing concern that this population might only use the technologies which they felt they had to or which offered a perceived direct benefit. Laver and colleagues used a discrete choice experiment to show how elderly subjects given a Wii rehabilitation protocol described a desire to use the standard rehabilitation over the Wii.[261] The subjects wanted a programme to have proven benefits before they put any effort into it. As the TKR-POWER study was the first of its kind, such assurances could not be given. Consequently, some resistance to the Wii was anticipated as a benefit could not be proven compared with routine care. As such a small proportion of this population had been exposed to gaming systems such as the Nintendo Wii, a more positive effect of the novelty factor was anticipated.

The main study in the thesis was the randomized controlled trial 'TKR-POWER'. The 136 patients who were randomized were followed for a year from their surgery. There was a statistically and clinically significant difference between the intervention and the control group in the mean difference in WOMAC total score. This improvement in function was evident at the end of the 12-week intervention period, which is not uncommon in physical therapy trials. However, the Wii group had a better WOMAC total score on average throughout the year of the trial, as shown by a linear regression model. This is the only trial in the literature which demonstrates a lasting difference in change from the 6 week baseline between groups at long-term follow-up, more than six months after the intervention has finished. This may have been specifically because of the effect of the Wii exercises. The extra balance component could have made the difference. Other potential reasons

include: the competitive element with the ability to chase higher scores, the real-time feedback on performance with coaching and encouragement to do better, and the variability leading to increased frequency and volume of exercises.

The Wii group had a poorer WOMAC score at the six-week baseline. This difference was not statistically significant and the groups had very similar scores at preoperative baseline. However, it did warrant an analysis of the data which accounted for baseline differences. An ANCOVA analysis was used which accounted for the phenomenon of regression to the mean. This showed that difference between the groups in ANCOVA analysis was less than that shown by the T-test analysis. The difference was still significant at 18 weeks but not at one year. Many studies in the orthopaedic literature only use a T-test analysis if there is no significant difference between the groups at baseline. The analysis used in this study is statistically rigorous and has reduced the potential impact of the one-year results. However, the linear regression took considerably more of the data into account and showed clearly that the Wii group were significantly better on average throughout the post-intervention phase of the study than the control group. It was always going to be more difficult to find a significant difference in these highly functioning patients.

Most TKR patients receive rehabilitation to a greater or lesser degree. The subjects in this trial were of a high socioeconomic demographic, leading to a high uptake and utilization of rehabilitation services through private insurance. The functional level of the patients at six-week baseline was consequently high, and the mean WOMAC score of the control group was only one point worse than age-matched Australian controls at 18 weeks (15.2 (study control group) vs. 14.2 (age-matched controls)) and 4.7 points better at one-year (9.5 vs. 14.2)(see chapter 2.3.2.1a). The control subjects were rehabilitated to a very high functional level, making the task of improving the function of the intervention group to a significantly higher level than the control group, very difficult.[321] The control patients had already returned to a similar level as measured by pain, knee stiffness, and function as their non-arthritic peers. These normative data are not matched to the same population as that of the surgical patients

in socioeconomic or educational metrics although it was an Australian population sample. They give a good indication of a comparable normative score.

While the minimum measurable difference for the WOMAC score is 1.0, a reduction of two or three points is unlikely to represent a significant clinical improvement. Studies that have investigated MCID give a range of MCID values for WOMAC Total scores representing clinical improvement between 8.2 and 9.1 points on a normalized scale (0-100) or 18% (sd 37%) of the total score.[64, 344] MCID does demonstrate a regression to the mean effect with those scoring below the median baseline WOMAC score showing improvement MCID of 20% versus 17% for those above the median.[344] The TKR-POWER study demonstrated a score improvement of seven points at 18 weeks and six points at one year before adjustment for baseline and 5.0 at 18 weeks and 2.5 at one year after adjustment for baseline. These fall outside of the quoted MCID values of 8.2-9.1 but within the 18% of the total score. The intervention group started at the 6 week baseline with a WOMAC score of 26.3 (12.6) and the control group with 22.5 (12.6) giving an 18% MCID value of 4.73 and 4.05 respectively. Within the varying parameters of the MCID measure, the changes seen in the TKR-POWER study are at the margins of clinically relevant.

The two groups ended up with a similar mean WOMAC total score at the 18-week assessment following the intervention (p=0.054). The study was designed to look at the change scores rather than the scores at one particular time-point. This is of clinical importance because the change score quantifies the individual improvements for each patient over a defined time period after an intervention. By looking at this change in the scores of individuals, they can be analysed for clinical significance. However, whereas it is relatively simple to assess WOMAC scores for statistical significance, clinical significance is more difficult to quantify.

Many of the WOMAC analyses of minimum clinical difference scores have been in the osteoarthritic population, assessing the effect of a single intervention. In the case of rehabilitation studies following surgery, such as in the TKR-POWER study, the patients all had a major intervention in the form of a

knee replacement, which dramatically improved their function and reduced pain. The change score was calculated between the groups who only differed in terms of a secondary intervention, in this case the Wii-enhanced rehabilitation. These change scores are much smaller and at the highly functional end of the WOMAC index.

An appropriate method of analysis is not yet agreed on. It is possible to assess a difference in several ways: the absolute change score, the relative change score, a subset of responders, or a level set by a method such as half of the standard deviation of the baseline scores.

The 20,50,70 method posited by Bellamy and colleagues stratifies patients into those whose WOMAC score had improved by 20, 50 or 70% of their baseline score. The analysis of the TKR-POWER patients showed that a statistically significant proportion of the Wii group had improved their WOMAC score by both 20% and 50% of their baseline scores at 18 weeks and one year compared with the controls.

The high level of function which the control group's WOMAC score represents compared with age-matched controls, the significance of the WOMAC 50% change, and the highly significant regression analysis all point towards the difference between the two groups being clinically important.

In this study, the levels of satisfaction were higher for the Wii group but the difference did not reach significance. Eighty-nine per cent of the Wii group were satisfied with their knee compared with 77% satisfied in the control group. The data were similar for satisfaction with rehabilitation at 92% for Wii vs. 84% for controls. Interestingly, 81% of the Wii group reported that their knee felt normal compared with 68% of the control group. This is perhaps the most important factor in the success of the knee replacement. If the knee feels normal, the patient is likely to be satisfied and happy with the whole process. The delay in collection of satisfaction surveys resulted in lower overall numbers for comparison and left the possibility of a type II error; the satisfaction rates might have been significantly different with larger numbers.

It is important to question the possible reasons why the Wii intervention demonstrated an improvement in functional outcome over the standard exercises given that the other studies have found no particular exercises programs have an advantage. The Wii provided an increase in the variety, the type, and the number of exercises offered. In any exercise-based intervention, it is vital to maintain the motivation of the subject. Offering a larger number of different exercises is likely to decrease boredom.

The Wii-Fit encouraged patients to increase the level of difficulty as their performance improved. This gave them a sense of competition and progress. It also led to a more efficient intervention. For every unit of time spent exercising, they were more likely to be working at an appropriate level of difficulty. This continued increase in the challenge to their aerobic, strength, and balance systems was likely to lead to greater gains than repeating exercises at a similar level of intensity.

The majority of the Wii-Fit exercises incorporate an element of balance training and give balance feedback in real time. Combining a balance challenge and feedback with a strength exercise, for example, increases the challenge to the patient's neural feedback mechanisms including proprioception and conscious balance systems. The dynamic nature of the training is potentially more applicable to function in the real world than the more static exercises of a routine protocol.

The Wii-Fit group performed a greater quantity of exercises per week throughout the study period than the control group. This was an encouraging finding. It is difficult for any population to maintain enthusiasm for exercise protocols. This intervention managed to fulfil enough criteria for adherence to a protocol to sustain interest and participation throughout 12 weeks with no sign of stopping. However, there were still fewer than 60% of patients who completed four exercises a day for four days of the week. This intervention has proved itself to be better than many in the literature in terms of unsupervised adherence but still does not capture all subjects. The diaries were all self-completed, which could have led to bias. However, neither group knew that they were a control or an intervention group. If there were intentions to try to please the investigators with a diary filled with

more ticks than actual exercises performed, the randomization process should have evened this out between the groups. The Wii consoles also store the data for each day's exercise volume. It is a follow-on project to extract those data and compare them with the self-reported diaries and investigate their accuracy.

There are several theories which could explain why the difference in WOMAC score between the two groups remained significant beyond the intervention period up to one year. The intervention may have caused a change in behaviour. The patients who had noticed a greater improvement in function were more inclined to continue their rehabilitation for a longer period of time. Even when the Wii was returned, they could still continue to exercise. It was also known that a number of patients either already had a Wii or bought a Wii at the end of the intervention period, allowing continued use during the post-intervention period. This in itself can be seen as a marker of the success of the intervention as it encouraged continued rehabilitation on a long-term basis, something which has proved so difficult with other interventions.

When the adherence rates were examined in the final study, it was clear that the intervention group performed more exercises per week throughout the study. With each exercise taking a similar amount of time, by extrapolation the intervention group spent more time rehabilitating. The Wii exercise diaries offered many more options for exercises than the control group. This was designed into the study as it reflected real life. The local practice is for formal rehabilitation to cease six to eight weeks after the operation. The patient is left with the instructions they had on finishing their last session. Neither group was limited from performing any other exercise, so they were free to seek extra physiotherapy, exercise tuition, or information and protocols from the internet. This was perceived as making the results more applicable to a clinical setting. It is also worth noting that, regardless of the generous choice of exercises, there was no rigid prescription for what to do. The patients were at liberty to choose. It was interesting that the Wii patients chose to do more Wii exercises. Future work needs to look at the qualitative data on the extra exercises and physiotherapy sought and performed by each group during the intervention period. It may be that the control group

did more exercise which was not recordable in the diaries. Even so, it would show that the Wii offers most of the exercises which are required without the need to seek extra ones and incur additional time and money costs.

6.3 OVERALL SUMMARY OF THE THESIS

To date, there has been no registered, appropriately powered, randomized clinical trial using the Wii as a rehabilitation tool in the TKR population. This study has demonstrated that the Wii could be an effective rehabilitation tool for use at home, particularly through improved adherence to the exercise regimen.

It is important that the results of any clinical study are read in context with the available evidence in the field. It should also be acknowledged that further research should be carried out to confirm these findings in other study populations. This study was conducted among a group with higher than average education and income levels, which may have influenced the uptake of new technology.

It is not clear which components of the Wii schedule were responsible for the improved outcome. The exercises contained a large balance component yet the pilot studies which looked at whether balance is associated with function did not provide a clear association . As the literature review suggested, it is a complicated area and a great many methods are used to measure balance and proprioception. Studies have tried to break the effect of balance down even further into bimodal effects associated with muscle power.[195] This is an area which is clearly too complicated for a simple screening test in the rooms and although the Wii balance board has been shown several times to be a valid and reliable force platform for custom software, the Wii-Fit software is not suitable for that task at the moment.

There was no significant difference in satisfaction between the two groups. The Wii group demonstrated satisfaction levels around 90%, which is the level of satisfaction seen in total hip

replacements. This is an achievement in itself. At the very least, the Wii group were as satisfied with their rehabilitation and the results of their TKR as the controls were.

6.4   FUTURE DIRECTIONS

The performance of a randomized clinical trial in the context of a private orthopaedic practice was a huge undertaking. There was an opportunity to collect data on a wide variety of additional outcomes which remain to be analysed. Future analyses could explore the following.

The associations of patients' psychosocial characteristics such as locus of control, anxiety, and depression with adherence to exercises, pain, and functional outcome warrants further investigation. This is a topic with the potential to interest many clinicians. There is often a belief amongst surgeons and physiotherapists that they can spot certain patients who will not 'do well' after a knee replacement. These patients form a proportion of the dissatisfied 20%. The accuracy of these clinicians' ability to identify these patients has not been tested.

There are patients who, regardless of how they score on objective tests and even some PROMs, will not be satisfied with their knee replacement. If these patients can be identified at the preoperative stage, then measures can be taken to try to improve their outcome. This may involve educational and cognitive strategies to address their knowledge and expectations of this procedure.

If this technology were to be prescribed on a large scale in a clinical setting then the issues of medical device approval and safety would need to be addressed. Although ethical approval was gained to use the device in a research setting, it is likely that as a commercial and clinical device it would need to undergo further approval. There was no statistically significant difference in the number of falls reported during the study and there were no injuries reported from the use of the Wii in the study. As a group of over 60 patients used the device, this suggests it is a safe tool for this population.

This study did not include a formal cost analysis. The additional hardware cost per unit was approximately $300 versus no cost for home exercises. Both groups needed their exercises explained and demonstrated which involves the cost of a therapist's time. The Wii group required a home visit to set the Wii up but it is anticipated that not all patients would require that level of help. The Wii has a simple design and set-up is not complicated. All exercises are well explained and continuous feedback given. It would be an interesting study to investigate the proportion of patients in this age group who could set their Wii up and start their rehabilitation with no home visit. It is planned to follow the cohort beyond the 12 months to report patients' satisfaction and track them longer term to record the rate of revision surgery. It would be feasible to conduct a cost-effectiveness analysis to compare the cost per unit gain on the WOMAC score, even though the study was not powered for this and numbers may be small.

Future analyses will factor in the cost of the intervention and the unit costs for the technology and the time needed for technical support. However, the benefits would be more difficult to cost. Quality-adjusted life years (QALYS) are a well-used method of costing procedures in the health sector. This method would need additional data to those which were collected. Therefore, given the current data, an estimate of the cost benefits of the functional gain would need to be made to offset the hardware costs of the intervention. The hardware costs approximately $300 per unit, which is a relatively small percentage of the overall cost of TKR surgery and rehabilitation programme which run into tens of thousands of dollars per patient.

There is still a large database which will be used for continued analysis to investigate further the effects of the intervention. This will include the associations of outcomes such as psychological assessments and the effect of the intervention on continued exercise out to one year and beyond. Further detailed statistical analyses will be employed to impute missing data and compare the non-responders and those lost to follow up to be sure that they are not confounders to the results.

Future Directions

This study has opened the door to many exciting areas of future study. The ethical approval included a THR-POWER study for looking at total hip replacements in a similar manner. An ACL-POWER study using similar basic methodology to the TKR-POWER study may demonstrate those functional improvements. Even though Baltaci and colleagues found no difference between a control and a Wii-Fit intervention group in an ACL reconstruction population, they did not give the patients the consoles to take home for use on a daily basis and the sample size was small.[267] These studies have the potential to improve outcomes in these related patient populations. The hypothesis is that the Wii offers variety, and increased exercise difficulty as patients progress can lead to increased adherence to a protocol and an overall increase in the frequency and volume of exercise. Further qualitative work with different subsets of patients may be useful to shed more light on this important issue of adherence to exercise regimens.

The TKR population could be tested further. The individual Wii exercises could be tested with regard to the popularity of the specific exercises measured in this study. Even though adherence was better than in the control group, it was still only just over 50%. This leaves room for improvement. Some of this could come from the software itself. It has a number of limitations.

The Wii-Fit software is a generic fitness-based software catering for all age groups and fitness levels. There are timelines and personal messages which are less appropriate for this population. The Wii has a general focus on weight and BMI, so much so that although the Wii-me character representing the patient starts off with average build, once a patient with a high BMI steps on the Wii balance board and they are weighed, the Wii-me grows around the waist in a very visible manner. The console then states that the patient is either overweight or obese. This was uncomfortable for many of the patients.

The measures of balance could be designed to be more valid and comparable to those in the literature. The exercises can generally be used as a replacement for the rehabilitation exercises which are routinely prescribed. However, a software programme which is targeted specifically at this patient population and their personalized recovery pathway would be ideal. It could lead to a faster transition along a protocol for those who are recovering well and encouragement for those who are struggling with the more individualized approach.

This leads us to an area of tele-rehabilitation which could revolutionize the postoperative process for most patients. Appropriate devices could be linked to the internet, allowing surgeons and physiotherapists to track the progress of their patients. Should a patient be obviously struggling or not registering their exercises, an appointment or a phone call could be made to investigate the issue and try to get things back on track. For those who are progressing as planned, an email or on-screen personalized message of encouragement could be sent. This could save on clinic visits, which is another potential cost saving, especially for those living in isolated areas like rural Australia.

The use of www.myclinicaloutcomes.co.uk was an innovation at the time. The study was the first user in Australia and remained the largest user for several years. It has been taken up by several NHS trusts in the UK and allows for vast quantities of PROMs data to be collected from patients followed up for arthroplasty surgery. The website gives the clinician easy access to their patients' scores as an average and on an individual basis. They can compare them with a national average to ensure they are getting the results they are aiming for. Importantly, it is another method of spotting 'red flags'. If a patient's scores suddenly deteriorate or show a downward trend which is not expected, the clinician can be alerted to it and call the patient back for follow-up to investigate the issue. This can save on expensive and time-consuming regular follow-up, which is increasingly being seen as unnecessary for the majority of arthroplasty patients. In this study the website collected PROMS data from more than 50% of the study subjects via email reminders and online data entry.

References

1.   McAlindon TE, Bannuru RR, Sullivan MC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis Cartilage* 2014; **22**(3): 363-88.

2.   Australian Orthopaedic Association National Joint Replacement Registry. Annual Report. Adelaide: AOA, 2013.

3.   Zhang Y, Jordan JM. Epidemiology of osteoarthritis. *Clin Geriatr Med* 2010; (26): 355–69.

4.   Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage* 2010; **18**: 24–33.

5.   Lementowski PW, Zelicof SB. Obesity and osteoarthritis. *Am J Orthop* 2008; **37**: 148–51.

6.   Reid CR, Bush PM, Cummings NH, McMullin DL, Durrani SK. A review of occupational knee disorders. *J Occup Rehabil* 2010; **20**: 489–501.

7.   Baker PN, van der Meulen JH, Lewsey J, Gregg PJ, National Joint Registry for E, Wales. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. *J Bone Joint Surg Br* 2007; **89**(7): 893-900.

8.   Robertsson O, Dunbar M, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. *Acta Orthop Scand* 2000; **71**(3): 262-7.

9.   Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res* 2006; **452**: 35-43.

10.   Gandhi R, Davey JR, Mahomed NN. Predicting patient dissatisfaction following joint replacement surgery. *J Rheumatol* 2008; **35**(12): 2415-8.

11.   Kim TK, Chang CB, Kang YG, Kim SJ, Seong SC. Causes and predictors of patient's dissatisfaction after uncomplicated total knee arthroplasty. *J Arthroplasty* 2009; **24**(2): 263-71.

12.   Bourne RB, Chesworth B, Davis A, Mahomed N, Charron K. Comparing patient outcomes after THA and TKA: is there a difference? *Clin Orthop Relat Res* 2010; **468**(2): 542-6.

13.   Fortin PR, Clarke AE, Joseph L, et al. Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. *Arthritis Rheum* 1999; **42**(8): 1722-8.

14.   Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for

osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ open* 2012; **2**(1): e000435.

15.	ACORN - Arthroplasty Clinical Outcomes Registry 2013 Annual report, 2013.

16.	Lange B, Flynn S, Proffitt R, Chang CY, Rizzo AS. Development of an interactive game-based rehabilitation tool for dynamic balance training. *Top Stroke Rehabil* 2010; **17**(5): 345-52.

17.	Deutsch JE, Borbely M, Filler J, Huhn K, Guarrera-Bowlby P. Use of a low-cost, commercially available gaming console (Wii) for rehabilitation of an adolescent with cerebral palsy. *Phys Ther* 2008; **88**(10): 1196-207.

18.	Bieryla KA, Dold NM. Feasibility of Wii Fit training to improve clinical measures of balance in older adults. *Clin Interv Aging* 2013; **8**: 775-81.

19.	Holmes JD, Jenkins ME, Johnson AM, Hunt MA, Clark RA. Validity of the Nintendo Wii balance board for the assessment of standing balance in Parkinson's disease. *Clinical Rehabilitation* 2013; **27**(4): 361-6.

20.	Baltaci G, Harput G, Haksever B, Ulusoy B, Ozer H. Comparison between Nintendo Wii Fit and conventional rehabilitation on functional performance outcomes after hamstring anterior cruciate ligament reconstruction: prospective, randomized, controlled, double-blind clinical trial. *Knee Surg Sports Traumatol Arthrosc* 2013; **21**(4): 880-7.

21.	Fung V, Ho A, Shaffer J, Chung E, Gomez M. Use of Nintendo Wii Fit in the rehabilitation of outpatients following total knee replacement: a preliminary randomised controlled trial. *Physiotherapy* 2012; **98**(3): 183-8.

22.	Medical Advisory SecretariatOntario Ministry of H, Long-Term C. Physiotherapy rehabilitation after total knee or hip replacement: an evidence-based analysis (Structured abstract). *Toronto: Medical Advisory Secretariat, Ontario Ministry of Health and Long-Term Care (MAS)*, 2005. http://www.mrw.interscience.wiley.com/cochrane/clhta/articles/HTA-32005001179/frame.html (accessed.

23.	Ruiz D, Jr., Koenig L, Dall TM, et al. The direct and indirect costs to society of treatment for end-stage knee osteoarthritis. *J Bone Joint Surg Am* 2013; **95**(16): 1473-80.

24.	Dakin H, Gray A, Fitzpatrick R, Maclennan G, Murray D, Group KATT. Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ open* 2012; **2**(1): e000332.

25.	Krummenauer F, Guenther KP, Kirschner S. Cost effectiveness of total knee arthroplasty from a health care providers' perspective before and after introduction of an interdisciplinary clinical pathway--is investment always improvement? *BMC Health Serv Res* 2011; **11**: 338.

26.	Higashi H, Barendregt JJ. Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis: discrete-event simulation model. *PloS one* 2011; **6**(9): e25403.

27.	Segal L, Day SE, Chapman AB, Osborne RH. Can we reduce disease burden from osteoarthritis? *Med J Aust* 2004; **180**(5 Suppl): S11-7.

28.	Dowsey MM, Liew D, Choong PF. Economic burden of obesity in primary total knee arthroplasty. *Arthritis Care Res (Hoboken)* 2011; **63**(10): 1375-81.

29.	TIme spent in hospitals and emergency departments. 2014. http://www.myhospitals.gov.au/hospital/1153B2180/royal-north-shore-hospital/length-of-stay-selected-surgical-procedures.

30.	Laubenthal KN, Smidt GL, Kettelkamp DB. A quantitative analysis of knee motion during activities of daily living. *Phys Ther* 1972; **52**(1): 34-43.

31.	Australian Orthopaedic Association National Joint Replacement Registry. Annual Report. Adelaide: AOA, 2014.

32.	Frost H, Lamb SE, Robertson S. A randomized controlled trial of exercise to improve mobility and function after elective knee arthroplasty. Feasibility, results and methodological difficulties. *Clin Rehabil* 2002; **16**(2): 200-9.

33.	Shepperd S, Harwood D, Gray A, Vessey M, Morgan P. Randomised controlled trial comparing hospital at home care with inpatient hospital care. II: cost minimisation analysis. *BMJ* 1998; **316**(7147): 1791-6.

34.	Kramer JF, Speechley M, Bourne R, Rorabeck C, Vaz M. Comparison of clinic- and home-based rehabilitation programs after total knee arthroplasty. *Clin Orthop Relat Res* 2003; (410): 225-34.

35.	Walsh MB, Herbold J. Outcome after rehabilitation for total joint replacement at IRF and SNF: a case-controlled comparison. *American Journal of Physical Medicine & Rehabilitation* 2006; **85**(1): 1-5.

36.	Attfield SF, Wilton TJ, Pratt DJ, Sambatakakis A. Soft-tissue balance and recovery of proprioception after total knee replacement. *J Bone Joint Surg Br* 1996; **78**(4): 540-5.

37.	Minns Lowe CJ, Barker KL, Dewey M, Sackley CM. Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2007; **335**(7624): 812.

38.	Mak JC, Fransen M, Jennings M, et al. Evidence-based review for patients undergoing elective hip and knee replacement. *ANZ journal of surgery* 2014; **84**(1-2): 17-24.

39.	Moffet H, Collet JP, Shapiro SH, Paradis G, Marquis F, Roy L. Effectiveness of intensive rehabilitation on functional ability and quality of life after first total knee arthroplasty: A single-blind randomized controlled trial. *Arch Phys Med Rehabil* 2004; **85**(4): 546-56.

40.	Lin C-WC, March L, Crosbie J, et al. Maximum recovery after knee replacement--the MARKER study rationale and protocol. *BMC Musculoskeletal Disorders* 2009; **10**: 69.

41.	Moffet H, Collet J-P, Shapiro SH, Paradis G, Marquis F, Roy L. Effectiveness of intensive rehabilitation on functional ability and quality of life after first total knee arthroplasty: A single-blind randomized controlled trial. *Arch Phys Med Rehabil* 2004; **85**(4): 546-56.

42.	Lingard EA, Berven S, Katz JN, Kinemax Outcomes G. Management and care of patients undergoing total knee arthroplasty: variations across different health care settings. *Arthritis Care Res* 2000; **13**(3): 129-36.

43.	Mahomed NN, Davis AM, Hawker G, et al. Inpatient compared with home-based rehabilitation following primary unilateral total hip or knee replacement: a randomized controlled trial. *J Bone Joint Surg Am* 2008; **90**(8): 1673-80.

44.   Coyte P, Young W, Croxford R. Costs and outcomes associated with alternative discharge strategies following joint replacement surgery: analysis of an observational study using a propensity score. *Journal of Health Economics* 2000; **19**: 907-29.

45.   Kauppila AM, Sintonen H, Aronen P, Ohtonen P, Kyllonen E, Arokoski JP. Economic evaluation of multidisciplinary rehabilitation after primary total knee arthroplasty based on a randomized controlled trial. *Arthritis Care Res (Hoboken)* 2011; **63**(3): 335-41.

46.   Lavernia CJ, D'Apuzzo MR, Hernandez VH, Lee DJ, Rossi MD. Postdischarge costs in arthroplasty surgery. *J Arthroplasty* 2006; **21**(6 Suppl 2): 144-50.

47.   Naylor J, Harmer A, Fransen M, Crosbie J, Innes L. Status of physiotherapy rehabilitation after total knee replacement in Australia. *Physiother Res Int* 2006; **11**(1): 35-47.

48.   NIH. National Institutes of Health: Consensus statement on total knee replacement. . NIH Consensus development Conference on TKR; 2003 December 8-10; 2003.

49.   Khan F, Ng L, Gonzalez S, Hale T, Turner-Stokes L. Multidisciplinary rehabilitation programmes following joint replacement at the hip and knee in chronic arthropathy. *Cochrane Database Syst Rev* 2008; (2): CD004957.

50.   Westby MD, Brittain A, Backman CL. Expert consensus on best practices for post-acute rehabilitation after total hip and knee arthroplasty: a Canada and United States Delphi study. *Arthritis Care Res (Hoboken)* 2014; **66**(3): 411-23.

51.   Black N. Patient reported outcome measures could help transform healthcare. *Bmj* 2013; **346**(jan28 1): f167-f.

52.   Dunbar M, Robertsson O, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty. Results of a survey of 3600 patients from The Swedish Knee Arthroplasty Registry. *J Bone Joint Surg [Br]* 2001; **83**: 339 - 44.

53.   Alviar MJ, Olver J, Brand C, Hale T, Khan F. Do patient-reported outcome measures used in assessing outcomes in rehabilitation after hip and knee arthroplasty capture issues relevant to patients? Results of a systematic review and ICF linking process. *Journal of Rehabilitation Medicine* 2011; **43**(5): 374-81.

54.   Hrobjartsson A, Thomsen AS, Emanuelsson F, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ* 2013; **185**(4): E201-11.

55.   Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988; **15**(12): 1833-40.

56.   Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997; **24**(4): 799-802.

57.   Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003; **106**(3): 337-45.

58.   Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005; **113**(1-2): 9-19.

59.   Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. *J Rheumatol* 2002; **29**(1): 131-8.

60.   Bellamy N, Kean WF, Buchanan WW, Gerecz-Simon E, Campbell J. Double blind randomized controlled trial of sodium meclofenamate (Meclomen) and diclofenac sodium (Voltaren): post validation reapplication of the WOMAC Osteoarthritis Index. *J Rheumatol* 1992; **19**(1): 153-9.

61.   Bellamy N. The WOMAC Knee and Hip Osteoarthritis Indices: development, validation, globalization and influence on the development of the AUSCAN Hand Osteoarthritis Indices. *Clin Exp Rheumatol* 2005; **23**(5 Suppl 39): S148-53.

62.   Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation of WOMAC. Following THR or TKR in OA. . *J Orth Rheum* 1988; **1**: 95-108.

63.   Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. *Clin Rheumatol* 1986; **5**(2): 231-41.

64.   Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis* 2005; **64**(1): 29-33.

65.   Minns Lowe CJ, Barker KL, Holder R, Sackley CM. Comparison of postdischarge physiotherapy versus usual care following primary total knee arthroplasty for osteoarthritis: an exploratory pilot randomized clinical trial. *Clin Rehabil* 2012; **26**(7): 629-41.

66.   Bellamy N, Wilson C, Hendrikz J. Population-based normative values for the Western Ontario and McMaster (WOMAC) osteoarthritis index and the Australian/Canadian (AUSCAN) hand osteoarthritis index functional subscales. *Inflammopharmacology* 2010; **18**(1): 1-8.

67.   Fortin PR, Penrod JR, Clarke AE, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. *Arthritis Rheum* 2002; **46**(12): 3327-30.

68.   Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. *J Rheumatol* 1995; **22**(6): 1193-6.

69.   Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol* 1997; **50**(3): 239-46.

70.   Kantz ME, Harris WJ, Levitsky K, Ware JE, Jr., Davies AR. Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. *Med Care* 1992; **30**(5 Suppl): MS240-52.

71.   Wells G, Beaton D, Shea B, et al. Minimal clinically important differences: review of methods. *J Rheumatol* 2001; **28**(2): 406-12.

72.   Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled clinical trials* 1989; **10**(4): 407-15.

73.   Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002; **14**(2): 109-14.

74.   Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal Clinically Important Difference. *J Rheumatol* 2001; **28**(2): 400-5.

75.   Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; **27**(3 Suppl): S178-89.

76.   Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Controlled clinical trials* 1991; **12**(4 Suppl): 142S-58S.

77.   Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? *Ann Rheum Dis* 2007; **66 Suppl 3**: iii40-1.

78.   Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; **41**(5): 582-92.

79.   Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res* 2004; **4**(5): 581-5.

80.   Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol* 2000; **27**(11): 2635-41.

81.   Bellamy N, Bell MJ, Goldsmith CH, et al. Evaluation of WOMAC 20, 50, 70 response criteria in patients treated with hylan G-F 20 for knee osteoarthritis. *Ann Rheum Dis* 2005; **64**(6): 881-5.

82.   Bellamy N, Bell MJ, Pericak D, et al. BLISS index for analyzing knee osteoarthritis trials data. *J Clin Epidemiol* 2007; **60**(2): 124-32.

83.   Bellamy N, Bell MJ, Goldsmith CH, et al. BLISS index using WOMAC index detects between-group differences at low-intensity symptom states in osteoarthritis. *J Clin Epidemiol* 2010; **63**(5): 566-74.

84.   Kirwan JR. Minimum clinically important difference: the crock of gold at the end of the rainbow? *J Rheumatol* 2001; **28**(2): 439-44.

85.   Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perception of patients about total knee replacement. *J Bone Joint Surg [Br]* 1998; **80-B**: 63 - 9.

86.   Murray DW, Fitzpatrick R, Rogers K, et al. The use of the Oxford hip and knee scores. *JBJS Br* 2007; **89**(8): 1010.

87.   Johnston L, MacLennan G, McCormack K, Ramsay C, Walker A. The Knee Arthroplasty Trial (KAT) design features, baseline characteristics, and two-year functional outcomes after alternative approaches to knee replacement. *J Bone Joint Surg Am* 2009; **91**(1): 134-41.

88.   Garratt AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: a structured review. *Rheumatology (Oxford)* 2004; **43**(11): 1414-23.

89.   Mohtadi N. Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency. *Am J Sports Med* 1998; **26**(3): 350-9.

90.   Bremner-Smith AT, Ewings P, Weale AE. Knee scores in a 'normal' elderly population. *Knee* 2004; **11**(4): 279-82.

91.   Clement ND, MacDonald D, Simpson AH. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2014; **22**(8): 1933-9.

92.   Harcourt WG, White SH, Jones P. Specificity of the Oxford knee status questionnaire. The effect of disease of the hip or lumbar spine on patients' perception of knee disability. *J Bone Joint Surg Br* 2001; **83**(3): 345-7.

93.   Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. *Rheumatology (Oxford)* 1999; **38**(9): 870-7.

94.   Jette D, Jette A. Physical therapy and health outcomes in patients with knee impairments. *Phys Ther* 1996; **76**: 1178 - 87.

95.   Kiebzak G, Campbell M, Mauerhan D. The SF-36 general health status survey documents the burden of osteoarthritis and the benefits of total joint arthroplasty: but why should we use it? *Am J Manag Care* 2002; **8**: 463 - 74.

96.   Lingard E, Katz J, Wright R, Wright E, Sledge C. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with the SF-36 and WOMAC. *J Bone Joint Surg [Am]* 2001; **83**: 1856.

97.   Davies AP. Rating systems for total knee replacement. *Knee* 2002; **9**(4): 261-6.

98.   Bachmeier CJ, March LM, Cross MJ, et al. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. *Osteoarthritis Cartilage* 2001; **9**(2): 137-46.

99.   Wilson D, Tucker G, Chittleborough C. Rethinking and rescoring the SF-12. *Sozial- und Praventivmedizin* 2002; **47**(3): 172-7.

100.   The SF-12: An even shorter health survey. http://www.sf-36.org/tools/sf12.shtml.

101.   Wallenstein SL. Neurophysiological Correlates of Pain. NewYork: Elsevier; 1984.

102.   Collins SL, Moore AM, McQuay HJ. The visual analogue pain intensity scale: What is moderate pain in millimetres? *Pain* 1997; **72**(1-2): 95-7.

103.   Morris LS, Schulz RM. Patient compliance--an overview. *J Clin Pharm Ther* 1992; **17**(5): 283-95.

104.   Berg JS, Dischler J, Wagner DJ, Raia JJ, Palmer-Shevlin N. Medication compliance: a healthcare problem. *Ann Pharmacother* 1993; **27**(9 Suppl): S1-24.

105.   Haynes RB, McKibbon KA, Kanani R. Systematic review of randomised trials of interventions to assist patients to follow prescriptions for medications. *Lancet* 1996; **348**(9024): 383-6.

106.   Donovan JL, Blake DR. Patient non-compliance: deviance or reasoned decision-making? *Soc Sci Med* 1992; **34**(5): 507-13.

107.   Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. *Arthritis Rheum* 1995; **38**(11): 1541-6.

108.   Adams S, Pill R, Jones A. Medication, chronic illness and identity: the perspective of people with asthma. *Soc Sci Med* 1997; **45**(2): 189-201.

109.   Campbell R, Evans M, Tucker M, Quilty B, Dieppe P, Donovan JL. Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee. *J Epidemiol Community Health* 2001; **55**(2): 132-8.

110.   Salmon P, Hall GM, Peerbhoy D, Shenkin A, Parker C. Recovery from hip and knee arthroplasty: Patients' perspective on pain, function, quality of life, and well-being up to 6 months postoperatively. *Arch Phys Med Rehabil* 2001; **82**(3): 360-6.

111.   Jones CA, Beaupre LA, Johnston DW, Suarez-Almazor ME. Total joint arthroplasties: current concepts of patient outcomes after surgery. *Rheum Dis Clin North Am* 2007; **33**(1): 71-86.

112.   Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am* 2004; **86-A**(5): 963-74.

113.   Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. *J Rheumatol* 2000; **27**(7): 1745-52.

114.   Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW. Functional outcome and patient satisfaction in total knee patients over the age of 75. *J Arthroplasty* 1996; **11**(7): 831-40.

115.   Hawker G, Wright J, Coyte P, et al. Health-related quality of life after knee replacement. *J Bone Joint Surg Am* 1998; **80**(2): 163-73.

116.   Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clin Orthop Relat Res* 2010; **468**(1): 57-63.

117.   Kennedy DM, Stratford PW, Riddle DL, Hanna SE, Gollish JD. Assessing recovery and establishing prognosis following total knee arthroplasty. *Phys Ther* 2008; **88**(1): 22-32.

118.   Nilsdotter AK, Toksvig-Larsen S, Roos EM. A 5 year prospective study of patient-relevant outcomes after total knee replacement. *Osteoarthritis & Cartilage* 2009; **17**(5): 601-6.

119.   Westby MD, Backman CL. Patient and health professional views on rehabilitation practices and outcomes following total hip and knee arthroplasty for osteoarthritis:a focus group study. *BMC Health Services Research* 2010; **10**: 119.

120.   Mahomed NN, Liang MH, Cook EF, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol* 2002; **29**(6): 1273-9.

121.   Woolhead G, Donovan J, Dieppe P. Patient expectations and total joint arthroplasty. *Journal of Rheumatology* 2003; **30**(7): 1656-7; author reply 7.

122.   Lochman JE. Factors related to patients' satisfaction with their medical care. *J Community Health* 1983; **9**(2): 91-109.

123.   Linder-Pelz SU. Toward a theory of patient satisfaction. *Soc Sci Med* 1982; **16**(5): 577-82.

124.   Ross CK, Sinacore JM, Stiers W, Budiman-Mak E. The role of expectations and preferences in health care satisfaction of patients with arthritis. *Arthritis Care Res* 1990; **3**(2): 92-8.

125.	Haas M. The relationship between expectations and satisfaction: a qualitative study of patients' experiences of surgery for gynaecological cancer. *Health Expect* 1999; **2**(1): 51-60.

126.	Fitzpatrick R, Hopkins A. Problems in the conceptual framework of patient satisfaction research: an empirical exploration. *Sociol Health Illn* 1983; **5**(3): 297-311.

127.	Vissers MM, de Groot IB, Reijman M, Bussmann JB, Stam HJ, Verhaar JA. Functional capacity and actual daily activity do not contribute to patient satisfaction after total knee arthroplasty. *BMC Musculoskeletal Disorders* 2010; **11**: 121.

128.	Buffart LM, van den Berg-Emons RJ, van Meeteren J, Stam HJ, Roebroeck ME. Lifestyle, participation, and health-related quality of life in adolescents and young adults with myelomeningocele. *Developmental medicine and child neurology* 2009; **51**(11): 886-94.

129.	Turner AP, Kivlahan DR, Haselkorn JK. Exercise and quality of life among people with multiple sclerosis: looking beyond physical functioning to mental health and participation in life. *Arch Phys Med Rehabil* 2009; **90**(3): 420-8.

130.	Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology and comparison of two theoretical models. *Management Science* 1989; **15**(8): 982-1003.

131.	Bailey JE, Pearson SW. Development of a tool for measuring and analyzing computer user satisfaction. *Management Science* 1983; **29**(5): 530-45.

132.	Coffin R, MacIntyre PD. Motivational influences on computer-related affective states. . *Computers in Human Behavior* 1999; **15**: 549-69.

133.	Swanson EB. Measuring user attitudes in MIS research. A review. *The International Journal of Management Science* 1982; **10**(2): 157-65.

134.	Lucas HCJ. Why information systems fail. New York: Columbia University Press; 1985.

135.	Fuerst WL, Cheney PH. Factors effecting the perceived utilization of computer based decision support systems. *Decision Sciences* 1982; **13**: 554-69.

136.	Shaft TM, Sharfman MP, Wu WW. Reliability assessment of the attitude towards computers instrument. *Computers in Human Behavior* 2004; **20**: 661-89.

137.	Schwartz J. The computer market. *American demographics* 1988; **10**(9): 38-41.

138.	Internet Access - Households and Individuals, 2013. United Kingdom: Office for National Statistics, 2013.

139.	Loyd BH, Loyd DE. The reliability and validity of an instrument for the assessment of computer attitudes. *Educational and Psychological Measurement* 1985; **45**(4): 903-8.

140.	Nickell GS, Pinto JN. The Computer attitude scale. *Computers in Human Behavior* 1986; **2**: 301-6.

141.	Abdel-Gaid S, Trueblood CR, Shrigley RL. A systematic procedures for contructing a valid microcomputer attitude scale. *Journal of Research in Science Teaching* 1986; **23**: 823-39.

142.	Knezek G, Miyashita K. A preliminary study of the Computer Attitude Questionnaire. Denton, TX: Texas Center for Educational Technology, 1994.

143. Popovich PM, Hyde KR, Zakrajsek T, Blumer C. The development of the attitudes toward computer usage scale. *Educational and Psychological Measurement* 1987; **47**: 261.

144. Jay GM, Willis SL. Influence of direct computer experience on older adults' attitudes toward computers. *J Gerontol* 1992; **47**(4): P250-7.

145. Kraus IK, Hoyer WJ. Technology and the older person: Age, sex and experience as moderators of attitudes toward computers. New York: Plenum Press; 1984.

146. Regan DT, Fazio R. On the consistency between attitudes and bahavior: Look to the method of attitude formation. *Journal of Experimental Social Psychology* 1977; **13**: 28-45.

147. Fishbein MC, Ajzen I. Beleif, attitde, intention and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley; 1975.

148. McGuire WJ. Attitudes and attitude change. New York: Addison-Wesley; 1985.

149. Fitzgerald GK, Piva SR, Irrgang JJ, Bouzubar F, Starz TW. Quadriceps activation failure as a moderator of the relationship between quadriceps strength and physical function in individuals with knee osteoarthritis. *Arthritis Rheum* 2004; **51**(1): 40-8.

150. Stratford PW, Kennedy D, Pagura SM, Gollish JD. The relationship between self-report and performance-related measures: questioning the content validity of timed tests. *Arthritis Rheum* 2003; **49**(4): 535-40.

151. Konig A, Walther M, Kirschner S, Gohlke F. Balance sheets of knee and functional scores 5 years after total knee arthroplasty for osteoarthritis: a source for patient information. *J Arthroplasty* 2000; **15**(3): 289-94.

152. Walsh M, Woodhouse LJ, Thomas SG, Finch E. Physical impairments and functional limitations: a comparison of individuals 1 year after total knee arthroplasty with control subjects. *Phys Ther* 1998; **78**(3): 248-58.

153. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res* 1989; (248): 13-4.

154. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural balance in an elderly population. *Arch Phys Med Rehabil* 1992; **73**(11): 1073-80.

155. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. *N Engl J Med* 1988; **319**(26): 1701-7.

156. Luukinen H, Koski K, Hiltunen L, Kivela SL. Incidence of rate of falls in an aged population in northern Finland. *J Clin Epidemiol* 1994; **47**: 843-50.

157. Nevitt MC, Cummings SR, Hudes ES. Risk factors for injurious falls: a prospective study. *J Gerontol* 1991; **46**(5): M164-70.

158. Miyamoto RG, Kaplan KM, Levine BR, Egol KA, Zuckerman JD. Surgical management of hip fractures: an evidence-based review of the literature. I: femoral neck fractures. *J Am Acad Orthop Surg* 2008; **16**(10): 596-607.

159. Tinetti ME. Clinical practice. Preventing falls in elderly persons. *N Engl J Med* 2003; **348**(1): 42-9.

160. Guidelines for the prevetnion of falls in older persons. American Geriatrics society, British Geriatrics Society, and American Academy of Ortthopaedic Surgeons Panel on Falls Prevention. *J Am Geriatr Soc* 2001; **49**: 664-72.

161.	Baloh RW, Corona S, Jacobson KM, Enrietto JA, Bell T. A prospective study of posturography in normal older people. *J Am Geriatr Soc* 1998; **46**(4): 438-43.

162.	Thapa PB, Gideon P, Brockman KG, Fought RL, Ray WA. Clinical and biomechanical measures of balance as fall predictors in ambulatory nursing home residents. *J Gerontol A Biol Sci Med Sci* 1996; **51**(5): M239-46.

163.	Lord SR, Sherrington C, Menz H. Risk factors and strategies for prevention. Cambridge: Cambridge University Press; 2001.

164.	Overstall PW, Exton-Smith AN, Imms FJ, Johnson AL. Falls in the elderly related to postural imbalance. *Br Med J* 1977; **1**(6056): 261-4.

165.	Wolf SL, Barnhart HX, Kutner NG, McNeely E, Coogler C, Xu T. Selected as the best paper in the 1990s: Reducing frailty and falls in older persons: an investigation of tai chi and computerized balance training. *J Am Geriatr Soc* 2003; **51**(12): 1794-803.

166.	Wild D, Nayak US, Isaacs B. Prognosis of falls in old people at home. *J Epidemiol Community Health* 1981; **35**(3): 200-4.

167.	Nichols DS, Glenn TM, Hutchinson KJ. Changes in the Mean Centre of Balance during balance testing in young adults. . *Phys Therap* 1995; **75**(8): 699-706.

168.	Gustavsen MA, G. Mengshoel, AM. Measuring balance in sub-acute stroke rehabilitation. *Advances in Physiotherapy* 2006; **8**(1): 15-22.

169.	Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. *Can J Public Health* 1992; **83 Suppl 2**: S7-11.

170.	Berg KO, Wood-Dauphinee SL, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiotherapy Canada* 1989; **41**: 304-11.

171.	Raiche M, Hebert R, Prince F, Corriveau H. Screening older adults at risk of falling with the Tinetti balance scale. *Lancet* 2000; **356**(9234): 1001-2.

172.	Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc* 1986; **34**(2): 119-26.

173.	Tinetti ME, Williams TF, Mayewski R. Fall risk index for elderly patients based on number of chronic disabilities. *Am J Med* 1986; **80**(3): 429-34.

174.	Haas BM, Burden AM. Validity of weight distribution and sway measurements of the Balance Performance Monitor. *Physiotherapy Research International*; **5**(1): 19-32.

175.	Piirtola M, Era P. Force Platform Measurements as Predictors of Falls among Older People - A Review. *Gerontology* 2006; **52**(1): 1-16.

176.	Goldie PA, Bach TM, Evans OM. Force platform measures for evaluating postural control: reliability and validity. *Arch Phys Med Rehabil* 1989; **70**(7): 510-7.

177.	Winter DA, Patla AE, Frank JS. Assessment of balance contol in humans. *Medical progress through technology* 1990; **16**: 31-51.

178.	Haas BM, Burden AM. Validity of weight distribution and sway measurements of the Balance Performance Monitor. *Physiother Res Int* 2000; **5**(1): 19-32.

179.	Geurts AC, Nienhuis B, Mulder TW. Intrasubject variability of selected force-platform parameters in the quantification of postural control. *Arch Phys Med Rehabil* 1993; **74**(11): 1144-50.

180.   Clark RA, Bryant AL, Pua Y, McCrory P, Bennell K, Hunt M. Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. *Gait Posture* 2010; **31**(3): 307-10.

181.   Tucker MG, Kavanagh JJ, Morrison S, Barrett RS. What Are the Relations Between Voluntary Postural Sway
Measures and Falls-History Status in Community-Dwelling
Older Adults? . *Arch Phys Med Rehabil* 2010; **91**(May): 750-8.

182.   Tucker MG, Kavanagh JJ, Barrett RS, Morrison S. Age-related differences in postural reaction time and coordination during voluntary sway movements. *Hum Mov Sci* 2008; **27**(5): 728-37.

183.   Teasdale N, Simoneau M. Attentional demands for postural control: the effects of aging and sensory reintegration. *Gait Posture* 2001; **14**(3): 203-10.

184.   Melzer I, Benjuya N, Kaplanski J. Age-related changes of postural control: effect of cognitive tasks. *Gerontology* 2001; **47**(4): 189-94.

185.   Maki BE, Holliday PJ, Topper AK. A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population. *J Gerontol* 1994; **49**(2): M72-84.

186.   Topper AK, Maki BE, Holliday PJ. Are activity-based assessments of balance and gait in the elderly predictive of risk of falling and/or type of fall? *J Am Geriatr Soc* 1993; **41**(5): 479-87.

187.   Boulgarides LK, McGinty SM, Willett JA, Barnes CW. Use of clinical and impairment-based tests to predict falls by community-dwelling older adults. *Phys Ther* 2003; **83**(4): 328-39.

188.   Bergland A, Jarnlo GB, Laake K. Predictors of falls in the elderly by location. *Aging clinical and experimental research* 2003; **15**(1): 43-50.

189.   Stel VS, Smit JH, Pluijm SM, Lips P. Balance and mobility performance as treatable risk factors for recurrent falling in older persons. *J Clin Epidemiol* 2003; **56**(7): 659-68.

190.   Sihvonen S, Sipila S, Taskinen S, Era P. Fall incidence in frail older women after individualized visual feedback-based balance training. *Gerontology* 2004; **50**(6): 411-6.

191.   Sihvonen SE, Sipila S, Era PA. Changes in postural balance in frail elderly women during a 4-week visual feedback training: a randomized controlled trial. *Gerontology* 2004; **50**(2): 87-95.

192.   Era P, Avlund K, Jokela J, et al. Postural balance and self-reported functional ability in 75-year-old men and women: a cross-national comparative study. *J Am Geriatr Soc* 1997; **45**(1): 21-9.

193.   Bascuas I, Tejero M, Monleon S, Boza R, Muniesa JM, Belmonte R. Balance 1 year after TKA: correlation with clinical variables. *Orthopedics* 2013; **36**(1): e6-12.

194.   Levinger P, Menz HB, Wee E, Feller JA, Bartlett JR, Bergman NR. Physiological risk factors for falls in people with knee osteoarthritis before and early after knee replacement surgery. *Knee Surgery, Sports Traumatology, Arthroscopy* 2011; **19**(7): 1082-9.

195.   Schwartz I, Kandel L, Sajina A, Litinezki D, Herman A, Mattan Y. Balance is an important predictive factor for quality of life and function after primary total knee replacement. *Journal of Bone & Joint Surgery - British Volume* 2012; **94**(6): 782-6.

196. Jarnlo GB. Functional balance tests related to falls among elderly people living in the community. *Eur J Geriatr* 2003; **5**(1): 7-14.

197. Medicine ACoS. ACSM's Guidelines for Exercise Testing and Prescription. (7th Ed). 2005.

198. Brill PA, Macera CA, Davis DR, Blair SN, Gordon N. Muscular strength and physical function. *Med Sci Sports Exerc* 2000; **32**(2): 412-6.

199. Buchman AS, Wilson RS, Boyle PA, Tang Y, Fleischman DA, Bennett DA. Physical activity and leg strength predict decline in mobility performance in older persons. *J Am Geriatr Soc* 2007; **55**(10): 1618-23.

200. Hughes VA, Frontera WR, Wood M, et al. Longitudinal muscle strength changes in older adults: influence of muscle mass, physical activity, and health. *J Gerontol A Biol Sci Med Sci* 2001; **56**(5): B209-17.

201. Barrett DS, Cobb AG, Bentley G. Joint proprioception in normal, osteoarthritic and replaced knees. *J Bone Joint Surg Br* 1991; **73**(1): 53-6.

202. Swanik CB, Lephart SM, Rubash HE. Proprioception, kinesthesia, and balance after total knee arthroplasty with cruciate-retaining and posterior stabilized prostheses. *J Bone Joint Surg Am* 2004; **86-A**(2): 328-34.

203. Wada M, Kawahara H, Shimada S, Miyazaki T, Baba H. Joint proprioception before and after total knee arthroplasty. *Clin Orthop Relat Res* 2002; (403): 161-7.

204. Howe TE, Rochester L, Jackson A, Banks PM, Blair VA. Exercise for improving balance in older people. *Cochrane Database Syst Rev* 2007; (4): CD004963.

205. Piva SR, Gil AB, Almeida GJM, DiGioia AM, 3rd, Levison TJ, Fitzgerald GK. A balance exercise program appears to improve function for patients with total knee arthroplasty: a randomized clinical trial. *Physical Therapy* 2010; **90**(6): 880-94.

206. Fitzgerald GK, Childs JD, Ridge TM, Irrgang JJ. Agility and perturbation training for a physically active individual with knee osteoarthritis. *Phys Ther* 2002; **82**(4): 372-82.

207. Webster KE, Feller JA, Wittwer JE. Balance confidence and function after knee-replacement surgery. *J Aging Phys Activity* 2006; **14**(2): 181-91.

208. Swinkels A, Newman JH, Allain TJ. A prospective observational study of falling before and after knee replacement surgery. *Age and Ageing* 2008; **38**(2): 175-81.

209. Connelly DM, Vandervoort AA. Effects of detraining on knee extensor strength and functional mobility in a group of elderly women. *J Orthop Sports Phys Ther* 1997; **26**(6): 340-6.

210. Gibbs J, Hughes S, Dunlop D, Singer R, Chang RW. Predictors of change in walking velocity in older adults. *J Am Geriatr Soc* 1996; **44**(2): 126-32.

211. Lord SR, Rogers MW, Howland A, Fitzpatrick R. Lateral stability, sensorimotor function and falls in older people. *J Am Geriatr Soc* 1999; **47**(9): 1077-81.

212. Moxley Scarborough D, Krebs DE, Harris BA. Quadriceps muscle strength and dynamic stability in elderly persons. *Gait Posture* 1999; **10**(1): 10-20.

213. Rejeski WJ, Mihalko SL. Physical activity and quality of life in older adults. *J Gerontol A Biol Sci Med Sci* 2001; **56 Spec No 2**: 23-35.

214. Slemenda C, Heilman DK, Brandt KD, et al. Reduced quadriceps strength relative to body weight: a risk factor for knee osteoarthritis in women? *Arthritis Rheum* 1998; **41**(11): 1951-9.

215.   Brandt KD, Heilman DK, Slemenda C, et al. Quadriceps strength in women with radiographically progressive osteoarthritis of the knee and those with stable radiographic changes. *J Rheumatol* 1999; **26**(11): 2431-7.

216.   McAlindon TE, Cooper C, Kirwan JR, Dieppe PA. Determinants of disability in osteoarthritis of the knee. *Ann Rheum Dis* 1993; **52**(4): 258-62.

217.   Stevens JE, Mizner RL, Snyder-Mackler L. Quadriceps strength and volitional activation before and after total knee arthroplasty for osteoarthritis. *J Orthop Res* 2003; **21**(5): 775-9.

218.   Huang CH, Cheng CK, Lee YT, Lee KS. Muscle strength after successful total knee replacement: a 6- to 13-year followup. *Clin Orthop Relat Res* 1996; (328): 147-54.

219.   Silva M, Shepherd EF, Jackson WO, Pratt JA, McClung CD, Schmalzried TP. Knee strength after total knee arthroplasty. *J Arthroplasty* 2003; **18**(5): 605-11.

220.   Maffiuletti NA. Assessment of hip and knee muscle function in orthopaedic practice and research. *J Bone Joint Surg Am* 2010; **92**(1): 220-9.

221.   Rossi MD, Hasson S. Lower-limb force production in individuals after unilateral total knee arthroplasty. *Arch Phys Med Rehabil* 2004; **85**(8): 1279-84.

222.   Mizner RL, Stevens JE, Snyder-Mackler L. Voluntary activation and decreased force production of the quadriceps femoris muscle after total knee arthroplasty. *Phys Ther* 2003; **83**(4): 359-65.

223.   Ishii Y, Terajima K, Koga Y, Takahashi HE, Bechtold JE, Gustilo RB. Gait analysis after total knee arthroplasty. Comparison of posterior cruciate retention and substitution. *Journal of orthopaedic science : official journal of the Japanese Orthopaedic Association* 1998; **3**(6): 310-7.

224.   Ouellet D, Moffet H. Locomotor deficits before and two months after knee arthroplasty. *Arthritis Rheum* 2002; **47**(5): 484-93.

225.   Cress ME, Buchner DM, Questad KA, Esselman PC, deLateur BJ, Schwartz RS. Continuous-Scale Physical Functional performance in healthy older adults: A validation study. *Arch Phys Med Rehabil* 1996; **77**: 1243-50.

226.   Lingard EA, Katz JN, Wright EA, Sledge CB, Kinemax Outcomes G. Predicting the outcome of total knee arthroplasty. *Journal of Bone & Joint Surgery - American Volume* 2004; **86-A**(10): 2179-86.

227.   Mizner RL, Petterson SC, Stevens JE, Axe MJ, Snyder-Mackler L. Preoperative quadriceps strength predicts functional ability one year after total knee arthroplasty. *Journal of Rheumatology* 2005; **32**(8): 1533-9.

228.   Jones CA, Voaklander DC, Suarez-Alma ME. Determinants of function after total knee arthroplasty. *Physical Therapy* 2003; **83**(8): 696-706.

229.   Zeni JA, Jr., Snyder-Mackler L. Early postoperative measures predict 1- and 2-year outcomes after unilateral total knee arthroplasty: importance of contralateral limb strength. *Physical Therapy* 2010; **90**(1): 43-54.

230.   Slemenda C, Brandt KD, Heilman DK, et al. Quadriceps weakness and osteoarthritis of the knee. *Ann Intern Med* 1997; **127**(2): 97-104.

231.   Valtonen A, Poyhonen T, Heinonen A, Sipila S. Muscle deficits persist after unilateral knee replacement and have implications for rehabilitation. *Phys Ther* 2009; **89**(10): 1072-9.

232.   Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "get-up and go" test. *Arch Phys Med Rehabil* 1986; **67**(6): 387-9.

233.   Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991; **39**(2): 142-8.

234.   Rikli R, Jones, CJ. Development and validation of a functional fitness test for community-residing older adults. *J Ageing Phy Act* 1999; **7**: 129-61.

235.   Hallage T, Krause MP, Haile L, et al. The effects of 12 weeks of step aerobics training on functional fitness of elderly women. *J Strength Cond Res* 2010; **24**(8): 2261-6.

236.   Gunter KB, White KN, Hayes WC, Snow CM. Functional mobility discriminates nonfallers from one-time and frequent fallers. *J Gerontol A Biol Sci Med Sci* 2000; **55**(11): M672-6.

237.   Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther* 2000; **80**(9): 896-903.

238.   Rose DJ, Jones J, Lucchese N. Predicting the probability of falls in adults using 8-foot get up and go. *J Ageing and Phys Act* 2002; **10**: 466-75.

239.   Austin N, Devine A, Dick I, Prince R, Bruce D. Fear of falling in older women: a longitudinal study of incidence, persistence, and predictors. *J Am Geriatr Soc* 2007; **55**(10): 1598-603.

240.   Shigematsu R, Chang M, Yabushita N, et al. Dance-based aerobic exercise may improve indices of falling risk in older women. *Age Ageing* 2002; **31**(4): 261-6.

241.   Artz N, Dixon S, Wylde V, Beswick A, Blom A, Gooberman-Hill R. Physiotherapy provision following discharge after total hip and total knee replacement: a survey of current practice at high-volume NHS hospitals in England and wales. *Musculoskelet* 2013; **11**(1): 31-8.

242.   Roos EM. Effectiveness and practice variation of rehabilitation after joint replacement. *Curr Opin Rheumatol* 2003; **15**(2): 160-2.

243.   Oldmeadow LB, McBurney H, Robertson VJ. Hospital stay and discharge outcomes after knee arthroplasty: implications for physiotherapy practice. *Aust J Physiother* 2002; **48**(2): 117-21.

244.   Waddell J, Johnson K, Hein W, Raabe J, FitzGerald G, Turibio F. Orthopaedic practice in total hip arthroplasty and total knee arthroplasty: results from the Global Orthopaedic Registry (GLORY). *Am J Orthop (Belle Mead NJ)* 2010; **39**(9 Suppl): 5-13.

245.   Westby MD, Kennedy D, Jones D, Jones A, Doyle-Waters MM, Backman C. Post-acute physiotherapy for primary total knee arthroplasty (protocol). *Cochrane Database of Systematic Reviews* 2008; (2).

246.   Maravic M, Landais P. Usefulness of a national hospital database to evaluate the burden of primary joint replacement for coxarthrosis and gonarthrosis in patients aged over 40 years. *Osteoarthritis Cartilage* 2006; **14**(6): 612-5.

247.   Ko V, Naylor J, Harris I, Crosbie J, Yeo A, Mittal R. One-to-one therapy is not superior to group or home-based therapy after total knee arthroplasty: a randomized, superiority trial. *Journal of Bone & Joint Surgery - American Volume* 2013; **95**(21): 1942-9.

248. Kramer JF, Speechley M, Bourne R, Rorabeck C, Vaz M. Comparison of clinic- and home-based rehabilitation programs after total knee arthroplasty. *Clin Orthop* 2003; (410): 225-34.

249. Madsen M, Larsen K, Madsen IK, Soe H, Hansen TB. Late group-based rehabilitation has no advantages compared with supervised home-exercises after total knee arthroplasty. *Dan Med J* 2013; **60**(4): A4607.

250. Mitchell C, Walker J, Walters S, Morgan AB, Binns T, Mathers N. Costs and effectiveness of pre- and post-operative home physiotherapy for total knee replacement: randomized controlled trial. *J Eval Clin Pract* 2005; **11**(3): 283-92.

251. Rajan RA, Pack Y, Jackson H, Gillies C, Asirvatham R. No need for outpatient physiotherapy following total knee arthroplasty: a randomized trial of 120 patients. *Acta Orthop Scand* 2004; **75**(1): 71-3.

252. Vuorenmaa M, Ylinen J, Piitulainen K, et al. Efficacy of a 12-month, monitored home exercise programme compared with normal care commencing 2 months after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med* 2014; **46**(2): 166-72.

253. Valtonen A, Poyhonen T, Sipila S, Heinonen A. Effects of aquatic resistance training on mobility limitation and lower-limb impairments after knee replacement. *Archives of Physical Medicine & Rehabilitation* 2010; **91**(6): 833-9.

254. LaStayo PC, Meier W, Marcus RL, Mizner R, Dibble L, Peters C. Reversing muscle and mobility deficits 1 to 4 years after TKA: a pilot study. *Clinical Orthopaedics & Related Research* 2009; **467**(6): 1493-500.

255. Mockford BJ, Beverland DE. Daoes an outpatient physiotherapy regime improve the range of knee motion after total knee arthroplasty: a prospective study. *JBJS Br* 2004; **86-B**(Supp II 12B).

256. Piva SR, Gil AB, Almeida GJ, DiGioia AM, 3rd, Levison TJ, Fitzgerald GK. A balance exercise program appears to improve function for patients with total knee arthroplasty: a randomized clinical trial. *Phys Ther* 2010; **90**(6): 880-94.

257. Pagnacco G, Oggero E, Wright CH. Biomedical instruments versus toys:a preliminary comparison of force platforms and the nintendo wii balance board - biomed 2011. *Biomed Sci Instrum* 2011; **47**: 12-7.

258. Negus JJ, Mani BA, Scholes CJ, March L, Parker DA. Is the Wii Fit for purpose? Australian and New Zealand Orthopaedic Research Society Annual meeting. QUT, Brisbane QLD; 2011.

259. Jorgensen MG, Laessoe U, Hendriksen C, Nielsen OBF, Aagaard P. Intrarater reproducibility and validity of Nintendo Wii balance testing in community-dwelling older adults. *Journal of aging and physical activity* 2014; **22**(2): 269-75.

260. Wikstrom EA. Validity and reliability of Nintendo Wii Fit balance scores. *Journal of athletic training* 2012; **47**(3): 306-13.

261. Laver K, Ratcliffe J, George S, Burgess L, Crotty M. Is the Nintendo Wii Fit really acceptable to older people? A discrete choice experiment. *BMC geriatrics* 2011; **11**: 64.

262. Young W, Ferguson S, Brault S, Craig C. Assessing and training standing balance in older adults: a novel approach using the 'Nintendo Wii' Balance Board. *Gait Posture* 2011; **33**(2): 303-5.

263.   Lewis GN, Woods C, Rosie JA, McPherson KM. Virtual reality games for rehabilitation of people with stroke: perspectives from the users. *Disability and rehabilitation Assistive technology* 2011; **6**(5): 453-63.

264.   Williams MA, Soiza RL, Jenkinson AM, Stewart A. EXercising with Computers in Later Life (EXCELL) - pilot and feasibility study of the acceptability of the Nintendo(R) WiiFit in community-dwelling fallers. *BMC research notes* 2010; **3**: 238.

265.   Prosperini L, Fortuna D, Gianni C, Leonardi L, Marchetti MR, Pozzilli C. Home-based balance training using the Wii balance board: a randomized, crossover pilot study in multiple sclerosis. *Neurorehabilitation and neural repair* 2013; **27**(6): 516-25.

266.   Nitz JC, Kuys S, Isles R, Fu S. Is the Wii Fit a new-generation tool for improving balance, health and well-being? A pilot study. *Climacteric* 2010; **13**(5): 487-91.

267.   Baltaci G, Harput G, Haksever B, Ulusoy B, Ozer H. Comparison between Nintendo Wii Fit and conventional rehabilitation on functional performance outcomes after hamstring anterior cruciate ligament reconstruction: prospective, randomized, controlled, double-blind clinical trial. *Knee Surgery, Sports Traumatology, Arthroscopy* 2013; **21**(4): 880-7.

268.   Goble DJ, Cone BL, Fling BW. Using the Wii Fit as a tool for balance assessment and neurorehabilitation: the first half decade of "Wii-search". *Journal of neuroengineering and rehabilitation* 2014; **11**: 12.

269.   Kennedy LG, Newman JH, Ackroyd CE, Dieppe PA. When should we do knee replacements? *Knee* 2003; **10**(2): 161-6.

270.   Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. The effect of age on pain, function, and quality of life after total hip and knee arthroplasty. *Arch Intern Med* 2001; **161**(3): 454-60.

271.   Gakhar H, McConnell B, Apostolopoulos AP, Lewis P. A pilot study investigating the use of at-home, web-based questionnaires compiling patient-reported outcome measures following total hip and knee replacement surgeries. *Journal of long-term effects of medical implants* 2013; **23**(1): 39-43.

272.   SOCRATES Orthopaedic Outcomes Software. http://www.socratesortho.com.

273.   Watkins MA, Riddle DL, Lamb RL, Personius WJ. Reliability of goniometric measurements and visual estimates of knee range of motion obtained in a clinical setting. *Phys Ther* 1991; **71**(2): 90-6; discussion 6-7.

274.   Busse ME, Hughes G, Wiles CM, Rosser AE. Use of hand-held dynamometry in the evaluation of lower limb muscle strength in people with Huntington's disease. *J Neurol* 2008; **255**(10): 1534-40.

275.   Fleiss J. The design and analysis of clinical experiments. New York, USA: John Wiley & Sons; 1986.

276.   Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; **59**(1): 12-9.

277.   Marsh AP, Rejeski WJ, Lang W, Miller ME, Messier SP. Baseline balance and functional decline in older adults with knee pain: the Observational Arthritis Study in Seniors. *J Am Geriatr Soc* 2003; **51**(3): 331-9.

278.	Wegener L, Kisner C, Nichols D. Static and dynamic balance responses in persons with bilateral knee osteoarthritis. *J Orthop Sports Phys Ther* 1997; **25**(1): 13-8.

279.	Hinman RS, Bennell KL, Metcalf BR, Crossley KM. Balance impairments in individuals with symptomatic knee osteoarthritis: a comparison with matched controls using clinical tests. *Rheumatology (Oxford)* 2002; **41**(12): 1388-94.

280.	Petrella RJ, Lattanzio PJ, Nelson MG. Effect of age and activity on knee joint proprioception. *Am J Phys Med Rehabil* 1997; **76**(3): 235-41.

281.	Lord SR, Clark RD, Webster IW. Physiological factors associated with falls in an elderly population. *J Am Geriatr Soc* 1991; **39**(12): 1194-200.

282.	Warren PJ, Olanlokun TK, Cobb AG, Bentley G. Proprioception after knee arthroplasty. The influence of prosthetic design. *Clin Orthop Relat Res* 1993; (297): 182-7.

283.	Koralewicz LM, Engh GA. Comparison of proprioception in arthritic and age-matched normal knees. *J Bone Joint Surg Am* 2000; **82-A**(11): 1582-8.

284.	Gillespie LD, Gillespie WJ, Cumming R, Lamb SE, Rowe BH. Interventions for preventing falls in the elderly. *Cochrane Database Syst Rev* 2000; (2): CD000340.

285.	Campbell AJ, Robertson MC. Rethinking individual and community fall prevention strategies: a meta-regression comparing single and multifactorial interventions. *Age Ageing* 2007; **36**(6): 656-62.

286.	Cachupe WJC, Shifflett W, Kahanov L, Wughalter EH. Reliability of Biodex Balance System Measures. *Measurement in Physical Education and Exercise Science* 2001; **5**(2): 97-108.

287.	Parraca JA, Olivares PR, Carbonell-Baeza A, Aparicio VA, Adsuar JC, Gusi N. Test-Retest reliability of Biodex Balance SD on physically active old people. *Journal of Human Sport and Exercise* 2011; **6**(2): 444-51.

288.	Clark RA, Howells B, Feller J, Whitehead T, Webster KE. Clinic-based assessment of weight-bearing asymmetry during squatting in people with anterior cruciate ligament reconstruction using Nintendo Wii Balance Boards. *Arch Phys Med Rehabil* 2014; **95**(6): 1156-61.

289.	Gil-Gomez JA, Llorens R, Alcaniz M, Colomer C. Effectiveness of a Wii balance board-based system (eBaViR) for balance rehabilitation: a pilot randomized clinical trial in patients with acquired brain injury. *J Neuroeng Rehabil* 2011; **8**: 30.

290.	Puh U, Majcen N, Hlebs S, Rugelj D. Effects of Wii balance board exercises on balance after posterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc* 2014; **22**(5): 1124-30.

291.	Huurnink A, Fransz DP, Kingma I, van Dieen JH. Comparison of a laboratory grade force platform with a Nintendo Wii Balance Board on measurement of postural control in single-leg stance balance tasks. *Journal of Biomechanics* 2013; **46**(7): 1392-5.

292.	Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**(8476): 307-10.

293.	Springer BA, Marin R, Cyhan T, Roberts H, Gill NW. Normative values for unipedal stance. . *J Geriatric Phy Ther* 2007; **30**(1): 8.

294.	Bohannon RW, Larkin PA, Cook AC, Gear J, Singer J. Decrease in timed balance test scores with aging. *Phys Ther* 1984; **64**(7): 1067-70.

295.   Felson DT, Naimark A, Anderson J, Kazis L, Castelli W, Meenan RF. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum* 1987; **30**(8): 914-8.

296.   Felson DT, Zhang Y, Hannan MT, et al. The incidence and natural history of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum* 1995; **38**(10): 1500-5.

297.   Guccione AA, Felson DT, Anderson JJ, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *Am J Public Health* 1994; **84**(3): 351-8.

298.   Davis MA, Ettinger WH, Neuhaus JM, Mallon KP. Knee osteoarthritis and physical functioning: evidence from the NHANES I Epidemiologic Followup Study. *J Rheumatol* 1991; **18**(4): 591-8.

299.   Bennell KL, Hinman RS, Metcalf BR, et al. Relationship of knee joint proprioception to pain and disability in individuals with knee osteoarthritis. *J Orthop Res* 2003; **21**(5): 792-7.

300.   Barrack RL, Skinner HB, Cook SD, Haddad RJ, Jr. Effect of articular disease and total knee arthroplasty on knee joint-position sense. *J Neurophysiol* 1983; **50**(3): 684-7.

301.   Kaplan FS, Nixon JE, Reitz M, Rindfleish L, Tucker J. Age-related changes in proprioception and sensation of joint position. *Acta Orthop Scand* 1985; **56**(1): 72-4.

302.   Sharma L, Pai YC, Holtkamp K, Rymer WZ. Is knee joint proprioception worse in the arthritic knee versus the unaffected knee in unilateral knee osteoarthritis? *Arthritis Rheum* 1997; **40**(8): 1518-25.

303.   Pai YC, Rymer WZ, Chang RW, Sharma L. Effect of age and osteoarthritis on knee proprioception. *Arthritis Rheum* 1997; **40**(12): 2260-5.

304.   Nordesjo L, Nordgren B, Wigren A. Isometric strength and endurance in patients with severe rheumatoid arthritis or osteoarthritis in the knee joint. *Scand J Rheumatol* 1983; **12**: 152-6.

305.   Fuchs S, Thorwesten L, Niewerth S. Proprioceptive function in knees with and without total knee arthroplasty. *Am J Phys Med Rehabil* 1999; **78**(1): 39-45.

306.   Gillespie LD, Gillespie WJ, Robertson MC, Lamb SE, Cumming RG, Rowe BH. Interventions for preventing falls in elderly people. *Cochrane Database Syst Rev* 2003; (4): CD000340.

307.   Tinetti ME. Instability and falling in elderly patients. *Seminars in neurology* 1989; **9**(1): 39-45.

308.   McChesney JW, Woollacott MH. The effect of age-related declines in proprioception and total knee replacement on postural control. *J Gerontol A Biol Sci Med Sci* 2000; **55**(11): M658-66.

309.   Pua YH, Liang Z, Ong PH, Bryant AL, Lo NN, Clark RA. Associations of knee extensor strength and standing balance with physical function in knee osteoarthritis. *Arthritis care & research* 2011; **63**(12): 1706-14.

310.   Hassan BS, Mockett S, Doherty M. Static postural sway, proprioception, and maximal voluntary quadriceps contraction in patients with knee osteoarthritis and normal control subjects. *Ann Rheum Dis* 2001; **60**(6): 612-8.

311. Hurley MV, Scott DL, Rees J, Newham DJ. Sensorimotor changes and functional performance in patients with knee osteoarthritis. *Ann Rheum Dis* 1997; **56**(11): 641-8.

312. Nuesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Juni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011; **342**: d1165.

313. Levinger P, Menz HB, Wee E, Feller JA, Bartlett JR, Bergman NR. Physiological risk factors for falls in people with knee osteoarthritis before and early after knee replacement surgery. *Knee Surg Sports Traumatol Arthrosc* 2011; **19**(7): 1082-9.

314. Rudolph KS, Schmitt LC, Lewek MD. Age-related changes in strength, joint laxity, and walking patterns: are they related to knee osteoarthritis? *Phys Ther* 2007; **87**(11): 1422-32.

315. Cenciarini M, Loughlin PJ, Sparto PJ, Redfern MS. Stiffness and damping in postural control increase with age. *IEEE transactions on bio-medical engineering* 2010; **57**(2): 267-75.

316. van der Esch M, Steultjens M, Harlaar J, Knol D, Lems W, Dekker J. Joint proprioception, muscle strength, and functional ability in patients with osteoarthritis of the knee. *Arthritis Rheum* 2007; **57**(5): 787-93.

317. Minor MA, Hewett JE, Webel RR, Dreisinger TE, Kay DR. Exercise tolerance and disease related measures in patients with rheumatoid arthritis and osteoarthritis. *J Rheumatol* 1988; **15**(6): 905-11.

318. Menegoni F, Galli M, Tacchini E, Vismara L, Cavigioli M, Capodaglio P. Gender-specific effect of obesity on balance. *Obesity (Silver Spring)* 2009; **17**(10): 1951-6.

319. Hue O, Simoneau M, Marcotte J, et al. Body weight is a strong predictor of postural stability. *Gait Posture* 2007; **26**(1): 32-8.

320. Rougier PR. Relative contribution of the pressure variations under the feet and body weight distribution over both legs in the control of upright stance. *J Biomech* 2007; **40**(11): 2477-82.

321. Bellamy N, Wilson C, Hendrikz J. Population-Based normative values for the Western Ontario and McMaster (WOMAC) Osteoarthritis Index: Part I. *Semin Arthritis Rheum* 2011; **41**: 139-48.

322. Negus JJ, Mani BA, Scholes CJ, March L, Parker DA. Is the Wii Fit for purpose? British Orthopaedic Research Society Annual Meeting. Cambridge, UK; 2011.

323. Rissanen P, Aro S, Slatis P, Sintonen H, Paavolainen P. Health and quality of life before and after hip or knee arthroplasty. *J Arthroplasty* 1995; **10**(2): 169-75.

324. van Essen GJ, Chipchase LS, O'Connor D, Krishnan J. Primary total knee replacement: short-term outcomes in an Australian population. *J Qual Clin Pract* 1998; **18**(2): 135-42.

325. Yoo JH, Chang CB, Kang YG, Kim SJ, Seong SC, Kim TK. Patient expectations of total knee replacement and their association with sociodemographic factors and functional status. *Journal of Bone & Joint Surgery - British Volume* 2011; **93**(3): 337-44.

326. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clinical Orthopaedics & Related Research* 2010; **468**(1): 57-63.

327.   Dickstein R, Heffes Y, Shabtai EI, Markowitz E. Total knee arthroplasty in the elderly: patients' self-appraisal 6 and 12 months postoperatively. *Gerontology* 1998; **44**(4): 204-10.

328.   Williams JI, Llewellyn Thomas H, Arshinoff R, Young N, Naylor CD. The burden of waiting for hip and knee replacements in Ontario. Ontario Hip and Knee Replacement Project Team. *J Eval Clin Pract* 1997; **3**(1): 59-68.

329.   March LM, Cross MJ, Lapsley H, et al. Outcomes after hip or knee replacement surgery for osteoarthritis. A prospective cohort study comparing patients' quality of life before and after surgery with age-related population norms. *Med J Aust* 1999; **171**(5): 235-8.

330.   Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. *Arch Phys Med Rehabil* 1999; **80**(5): 572-9.

331.   Murray MP, Gore DR, Laney WH, Gardner GM, Mollinger LA. Kinesiologic measurements of functional performance before and after double compartment Marmor knee arthroplasty. *Clin Orthop Relat Res* 1983; (173): 191-9.

332.   Skinner HB. Pathokinesiology and total joint arthroplasty. *Clin Orthop Relat Res* 1993; (288): 78-86.

333.   Roush SE. Patient-perceived functional outcomes associated with elective hip and knee arthroplasties. *Phys Ther* 1985; **65**(10): 1496-500.

334.   Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c332.

335.   Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. *Osteoarthritis Cartilage* 2007; **15**(3): 273-80.

336.   Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty. Results of a survey of 3600 patients from The Swedish Knee Arthroplasty Registry. *J Bone Joint Surg Br* 2001; **83**(3): 339-44.

337.   Carpenter MR, Carpenter RL, Peel J, Zukley LM. The reliability of isokinetic and isometric leg strength measures among individuals with symptoms of mild osteoarthritis. *J Sports Med Phys Fitness* 2006; **46**: 1-5.

338.   Stickles B, Phillips L, Brox WT, Owens B, Lanzer WL. Defining the relationship between obesity and total joint arthroplasty. *Obesity research* 2001; **9**(3): 219-23.

339.   Lamb SE, Frost H. Recovery of mobility after knee arthroplasty: expected rates and influencing factors. *J Arthroplasty* 2003; **18**(5): 575-82.

340.   Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am* 2004; **86-A**(10): 2179-86.

341.   Ritter MA, Wing JT, Berend ME, Davis KE, Meding JB. The clinical effect of gender on outcome of total knee arthroplasty. *J Arthroplasty* 2008; **23**(3): 331-6.

342.   Hidding A, van Santen M, De Klerk E, et al. Comparison between self-report measures and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis and fibromyalgia. *J Rheumatol* 1994; **21**(5): 818-23.

343.   Wittink H, Rogers W, Sukiennik A, Carr DB. Physical functioning: self-report and performance measures are related but distinct. *Spine (Phila Pa 1976)* 2003; **28**(20): 2407-13.

344.   Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for

required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. *Arthritis Rheum* 2001; **45**(4): 384-91.

345.   Wobig M, Bach G, Beks P, et al. The role of elastoviscosity in the efficacy of viscosupplementation for osteoarthritis of the knee: a comparison of hylan G-F 20 and a lower-molecular-weight hyaluronan. *Clinical therapeutics* 1999; **21**(9): 1549-62.

346.   van den Akker-Scheek I, Stevens M, Groothoff JW, Bulstra SK, Zijlstra W. Preoperative or postoperative self-efficacy: which is a better predictor of outcome after total hip or knee arthroplasty? *Patient Educ Couns* 2007; **66**(1): 92-9.

347.   Ko V, Naylor J, Harris I, Crosbie J, Yeo A, Mittal R. One-to-one therapy is not superior to group or home-based therapy after total knee arthroplasty: a randomized, superiority trial. *J Bone Joint Surg Am* 2013; **95**(21): 1942-9.

APPENDIX

Medline Strategy

Rehabilitation in TKR

1. exp Arthroplasty, Replacement, Knee/

2. exp Knee Prosthesis/

3. exp Knee Joint/

4. exp Joint Prosthesis/

5. knee.mp.

6. (replace$ or arthroplast$ or implant$ or endoprosthe$ or prosthe$).mp.

7. exp "Prostheses and Implants"/

8. 3 or 5

9. or/4,6-7

10. 8 and 9

11. or/1-2,10

12. exp REHABILITATION/

13. exp REHABILITATION CENTERS/

14. exp REHABILITATION NURSING/

15. rehab$.mp.

16. exp Patient Care Team/

17. multidisciplinar$.tw.

18. interdisciplinar$.tw.

19. multiprofessional$.tw.

20. multimodal$.tw.

21. exp Patient Care Management/

22. exp Occupational Therapy/

23. occupational therap$.tw.

24. exp Physical Therapy Techniques/

25. exp "Physical Therapy (Specialty)"/

26. exp Physical Therapy Department, Hospital/

27. physical therap$.tw.

28. physiotherap$.tw.

29. (early adj1 (mobil$ or discharg$ or ambulat$)).tw.

30. exp Critical Pathways/

31. exp Therapy, Computer-Assisted/

32. exp Exercise Therapy/

33. (exercis$ adj3 therap$).tw.

34. or/12-33

35. 11 and 34

36. functional outcome$.tw

37. subjective outcome$.tw

38. objective outcome$.tw

39. tgug$.tw

40. tug$.tw

41. ug$.tw

42. timed and up and go$.tw

43. dynamometer$.tw

44. handheld dynan$.tw

45. isometric strength$.tw

46. quadriceps$.tw

47. exp knee joint

48. exp balance

49. force plate$.tw

50. stabilometry$.tw

51. cent$ of pressure$.tw

52. sway$.tw

53. womac$.tw

54. western Ontario mcmaster$.tw

55. oks$.tw

56. oxford knee score$.tw

57. sf-12$.tw

58. hads$.tw

59. American knee society score$.tw

60. kscrs$.tw

61. computer efficacy$.tw

62. computer awareness$.tw

63. global health measure$.tw

64. disease specific$.tw

65. exp osteoarthritis

66. exp compliance

67. exp Adherence

68. exercise diary$.tw

69. Wii$.tw

70. X-box$.tw

71. Kinect$.tw

72. Playstation$.tw

73. Ps3$.tw

74. virtual$.tw

75. Telemedicine$.tw

76. Computer assisted$.tw

77. Or/36-76

78. 35 and 77

ii) WOMAC index of osteoarthritis

Patient Name: _____

Patient ID: _____   Side: _____

Date of Review: ___ / ___ / ___   (complete either the date of review or the follow up period below)

Follow up period: Pre Op   OR   _____   Weeks / Months / Years   (add the delay and circle one)

<u>**Patients - please place an X in one box on each line to indicate your response to that question.**</u>

These questions concern the amount of pain you are currently experiencing due to arthritis in your knees. For each situation please enter the amount of pain recently experienced.

How much pain you have ..

1. Walking on a flat surface?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

2. Going up or down stairs ?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

3. At night while in bed ?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

4. Sitting or lying ?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

5. Standing upright ?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

These questions concern the amount of joint stiffness (not pain) you are currently experiencing due to arthritis in your knees. Stiffness is a sensation of restriction or slowness in the use with which you move your joints.

6. How severe is your stiffness after first wakening in the morning ?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

7. How severe is your stiffness after sitting, lying, or resting later in the day ?

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

265

These questions concern your physical function.

By this we mean your ability to move around and to look after yourself. For each of the following activities, please indicate the degree of difficulty you are currently experiencing due to arthritis.

What degree of difficulty do you have with ...

8. Descending stairs

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

9. Ascending stairs

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

10. Rising from sitting

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

11. Standing

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

12. Bending to floor

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

13. Walking on flat

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

14. Getting out / in of car

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

15. Going shopping

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

16. Putting on socks/stockings

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

17. Rising from bed

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

18. Taking off socks/stockings

☐ None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme

These questions concern your physical function.

By this we mean your ability to move around and to look after yourself. For each of the following activities, please indicate the degree of difficulty you are currently experiencing due to arthritis.

What degree of difficulty do you have with ...

19. Lying in a bed

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

20. Getting in / out a bath or shower

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

21. Sitting

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

22. Getting on /off toilet

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

23. Heavy domestic duties

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

24. Light domestic duties

☐ None   ☐ Mild   ☐ Moderate   ☐ Severe   ☐ Extreme

267

iii) Oxford Knee Score

Oxford 12-item Knee Questionnaire

SOCRATES

ICRS

eForm v 2.1

Patient Name: _____

Side: ☐ Left

Patient ID: _____

☐ Right

Date of review: ☐☐ ☐☐ ☐☐☐☐ (complete either the date of review or the follow up period below)

Follow up period: Pre Op OR _____ Weeks / Months / Years (add the delay and circle one)

Patients - please place an X in one box on each line to indicate your response to that question.

During the past four weeks

1. How would you describe the pain you usually had from your knee ?

☐ None ☐ Very mild ☐ Mild ☐ Moderate ☐ Severe

2. Have you had any trouble with washing and drying yourself (all over) because of your knee ?

☐ No trouble at all ☐ Very little trouble ☐ Moderate trouble ☐ Extreme difficulty ☐ Unable to do

3. Have you had any trouble getting in and out of a car or using public transport because of your knee ? (whichever you tend to use)

☐ No trouble at all ☐ Very little trouble ☐ Moderate trouble ☐ Extreme difficulty ☐ Unable to do

4. For how long have you been able to walk before the pain from your knee became severe ? (with or without a stic

☐ No pain / >30 minutes ☐ 16 to 30 minutes ☐ 5 to 15 minutes ☐ Around the house only ☐ Not at all

5. After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your knee ?

☐ Not at all painful ☐ Slightly painful ☐ Moderately painful ☐ Very painful ☐ Unbearable

6. Have you been limping when walking, because of your knee ?

☐ Rarely / never ☐ Sometimes or just at first ☐ Often, not just at first ☐ Most of the time ☐ All of the time

During the past four weeks

7. Could you kneel down and get up again afterwards ?

| Yes, easily | With little difficulty | With moderate difficulty | With extreme difficulty | No, unable |

8. Have you been troubled by pain from your knee in bed at night ?

| No nights | Only 1 or 2 nights | Some nights | Most nights | Every night |

9. How much has pain from your knee interfered with your usual work (including housework) ?

| Not at all | A little bit | Moderately | Greatly | Totally |

10. Have you felt that your knee might suddenly give away or let you down ?

| Rarely / never | Sometimes or just at first | Often, not just at first | Most of the time | All of the time |

11. Could you do the household shopping on your own ?

| Yes, easily | With little difficulty | With moderate difficulty | With extreme difficulty | No, unable |

12. Could you walk down a flight of stairs ?

| Yes, easily | With little difficulty | With moderate difficulty | With extreme difficulty | No, unable |

270

iv) SF-12

SOCRATES

I C R S
International Cartilage Repair Society

eForm v 2.1

Patient Name: _____   **Patient ID:** _____

Reviewer Name: _____   **Side:** ☐ Left ☐ Right

Filled in by: ☐ Operating Dr. ☐ Other MD ☐ Research Assistant ☐ Questionnaire ☐ Other

Date of review: ☐☐☐☐☐☐☐☐   **Next visit date:** ☐☐☐☐☐☐☐☐

Study Name: _____   **Study Number:** _____

1. In general, would you say your health is:

   ☐ Excellent ☐ Very good ☐ Good ☐ Fair ☐ Poor

   The following items are about activities you might do during a typical day. Does your health now limit you in these activities ?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
2. Moderate activities:moving a table, pushing a vacuum cleaner, bowling, playing golf	☐	☐	☐
3. Climbing several flights of stairs	☐	☐	☐

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities...
... as a result of your physical health ?

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
4. Accomplished less than you would like	☐	☐	☐	☐	☐
5. Were limited in the kind of work or other activities	☐	☐	☐	☐	☐

...as a result of any emotional problems (such as feeling depressed or anxious)

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
6. Accomplished less than you would like	☐	☐	☐	☐	☐
7. Didn't do work or other activities as carefully as usual	☐	☐	☐	☐	☐

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)

   ☐ Not at all ☐ A little bit ☐ Moderately ☐ Quite a bit ☐ Extremely

These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks...

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
9. Have you felt calm and peaceful ?	☐	☐	☐	☐	☐
10. Did you have a lot of energy ?	☐	☐	☐	☐	☐
11. Have you felt downhearted and blue ?	☐	☐	☐	☐	☐
12. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc) ?	☐	☐	☐	☐	☐

SF12 : Health Status Questionnaire v2 Page 1 of 1

272

v) Visual analogue scale pain score

VAS Pain Assessment Score

SOCRATES

I C R S

eform v 3.1

Patient Name: _____   **Side:** ☐ Left

Patient ID: _____   ☐ Right

Date of review: ☐☐☐☐☐☐☐☐   (complete either the date of review or the follow up period below)

Follow up period: Pre Op   OR   _____   Weeks / Months / Years   (add the delay and circle one)

Simply place a vertical line at the position on the line below that corresponds accurately with your perception of your answer to the question. Please ensure that your line crosses the horizontal line, inside the shaded area.

The further to the left you put your line, the less you experience that symptom, the further to the right you put your line, the more you experience that symptom.

NO PAIN ——————————————————————— EXTREME PAIN

VAS Pain Assessment Score

STANDARD QUESTION ASKED –

'How much pain are you feeling in the operated knee on an average day within the

last week?'

vi) Satisfaction assessment

Subject ID : _____

Randomization number : _____

18 week or 1 year : _____

			Very dissatisfied	Dissatisfied	Satisfied	Very Satisfied
Q1	How satisfied are you with your operated knee(s)?		1	2	3	4
Q2	How satisfied are you with your rehabilitation?					

			YES	NO		
Q3	Does the operated knee (s) feel normal?		1	0		

Tuesday, 16 October 12

vii) CIGAR

CIGAR - Computers, Internet, Gaming in Rehabilitation

1. ID: ____Surname_____ preop ___ 6w __ 18w ___6m __ 1y___

٭ 1. How comfortable do you feel using computers in general?

◯ Very comfortable

◯ Happy to use them for most things

◯ I use them if I have to

◯ I am uncomfortable using computers and avoid if possible

٭ 2. Do you have family of friends that help you with technical and computing issues?

◯ Yes

◯ No

◯ Don't need help

٭ 3. How often do you check your emails?

◯ Daily or more

◯ Weekly

◯ Monthly or less

◯ I don't have an email address

٭ 4. Do you check your emails on a computer, a smart phone or both?

◯ Computer

◯ Smart Phone

◯ Both

◯ I don't use email

٭ 5. If you have internet at home, what type do you have?

☐ No internet

☐ Dial-up

☐ Broadband (wired, wireless or mobile)

٭ 6. How often do you use the internet?

◯ Daily

◯ Weekly

◯ Monthly or less

◯ I don't use it at all

CIGAR - Computers, Internet, Gaming in Rehabilitation

2. ID: ____Surname_____ preop ___ 6w __ 18w ___6m __ 1y___

✱7. How comfortable are you using the following -

	Have never or will not use	Not comfortable but will use if have to	Happy to use occasionally	Comfortable using regualarly
A Desktop computer	◯	◯	◯	◯
Internet clothes shopping	◯	◯	◯	◯
Internet music / book shopping	◯	◯	◯	◯
Internet grocery shopping	◯	◯	◯	◯
Internet banking	◯	◯	◯	◯
Nintendo Wii / Wii-fit	◯	◯	◯	◯
Playstation 3 / X-Box 360	◯	◯	◯	◯
Nintendo DS i.e. Braintrainer	◯	◯	◯	◯
Facebook	◯	◯	◯	◯
Twitter	◯	◯	◯	◯

viii) Exercise diaries for Wii group

ID : XXX

Wii Diary 1- Week 7 onwards

Week	Beginning Monday	Click on Wii Board in locker room		BALANCE GAMES	YOGA			MUSCLE WORKOUTS				AEROBIC		Training plus
		Stillness - Needs unlocking	Basic balance test	Table tilt	Deep breathing	Half Moon	Warrior	Torso twists	Rowing Squat	Parallel stretch *	Jackknife *	Basic Step	Hula Hoop	Cycling
7	3rd Sep													
8	10th Sep													
9	17th Sep													
10	24th Sep													
11	1st Oct													
12	8th Oct													
13	15th Oct													
14	22nd Oct													
15	29th Oct													
16	5th Nov													
17	12th Nov													
18	19th Nov													

Place a tick in each box every time you perform that exercise in that week.
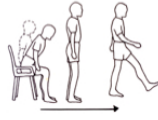
eg: the exercise was done 4 times this week   * = optional

ID : XXX

Wii Diary 2 - Week 11 onwards

Week	Beginning Monday	BALANCE GAMES			YOGA				MUSCLE WORKOUTS				AEROBIC			Training Plus	
		Ski slalom	Balance bubble	Penguin slide	Tree	Triangle	Sun salutation	Chair	Bridge	Crocodile Twist	Sideways leg lift	Side lunge	Jogging	Step plus	Super hula	Tilt city	Driving range
11	1st Oct																
12	8th Oct																
13	15th Oct																
14	22nd Oct																
15	29th Oct																
16	5th Nov																
17	12th Nov																
18	19th Nov																

Place a tick in each box every time you perform that exercise in that week.

✓ ✓
✓ ✓   eg: the exercise was done 4 times this week

ID : XXX

Wii Diary 3 - Week 15 onwards

Week	Beginning Monday	BALANCE TESTS & GAMES					YOGA					MUSCLE WORKOUTS					AEROBIC		
		Single leg balance test	Ski jump	Tilt city / Table tilt plus	Balance bubble plus	Snow board	Standing knee	Palm tree	King of the dance	Spine extension	Down ward dog	Single leg extension	Single leg twist	Balance bridge	Lunge	Single leg reach	Basic Step	Hula Hoop	Cycling (Wii Fit Plus)
15	29th Oct																		
16	5th Nov																		
17	12th Nov																		
18	19th Nov																		

Place a tick in each box every time you perform that exercise in that week.

✓ ✓
✓ ✓   eg: the exercise was done 4 times this week

282

ix) Exercise instructions and diary for both groups

Rehabilitation Study Exercises

Heel Raises

Stand tall
Push into balls of feet
Raise heels as high as you can
Slowly lower

Sit to Stand

Bend knees, feet back and hands on the chair
With weight through both legs, lean forward "nose over toes"
Assist with upper body as you push evenly through both legs to stand
Stand tall with hips and knees straight
Return to sitting at a controlled pace

½ Squat at Wall

Feet 2-3 inches from wall
Buttocks on wall, hands on thighs
Keeping your chest up, bend your knees and slide down as far as you can control
Push through legs to return to standing

Stairs

a) Foot placement

Stand close to step
Feet shoulder width apart
Stand tall
Bend operated leg hip and knee up to place your foot on the step
Try not to hitch your hip
Return to foot to the floor

b) Step Up

Place operated leg on step
Use handrail for support as needed
"Nose over toes" body forward onto step with other leg
Return to standing

c) Step Down

Standing on the step facing down
Use handrail for support as needed
Bend operated knee to lower the other leg onto the floor

Balance Exercises

a) Step Stance

Weight shift onto front/operated leg
Tense thigh muscles
Tighten core trunk muscles by exhaling as you draw navel up
Squeeze buttocks
Lift heel of the un-operated leg
Hold for 5-10 seconds
10 reps

b) Stride Stance

Weight shift onto operated leg,
Lift heel of un-operated leg
Stand tall, don't lean
Tighten buttocks
Hold for 5-10 seconds
10 reps

c) Single Leg Balance

Do not risk falling, stand close to a stable support
Stand tall
Keep knees relaxed
Draw in navel
Squeeze buttocks
Move weight over onto operated leg, lifting other leg
Hold

Exercise Diary

ID : XXX

Week	Beginning Monday	BALANCE			STRENGTH			STAIRS		
		Step stance	Stride stance	Single leg balance	Heel Raises	Sit to stand	1/2 squat at wall	Foot placement	Step up	Step down
7	10th Sep									
8	17th Sep									
9	24th Sep									
10	1st Oct									
11	8th Oct									
12	15th Oct									
13	22nd Oct									
14	29th Oct									
15	5th Nov									
16	12th Nov									
17	19th Nov									
18	26th Nov									

Place a tick in each box every time you perform that exercise in that week.

eg: the exercise was done 4 times this week

x) Publications included in the thesis

Patient outcomes using Wii-enhanced rehabilitation after total knee replacement – The TKR-POWER study

Q1 J.J. Negus [a,*], D. Cawthorne [b,1], J.S. Chen [c,2], C.J. Scholes [a,3], D.A. Parker [a,4], L.M. March [c,5]

[a] Sydney Orthopaedic Research Institute, Level 1 The Gallery, 445 Victoria Avenue, Chatswood NSW 2067, Australia
[b] Orthopaedic Department, Royal North Shore Hospital, St Leonard's, NSW, Australia
[c] University of Sydney Institute of Bone and Joint Research and Department of Rheumatology, Royal North Shore Hospital, St Leonard's, NSW, Australia

ARTICLE INFO

ABSTRACT

Background: Home-based rehabilitation following total knee replacement surgery can be as effective as clinic-based or in-patient rehabilitation. The use of the Nintendo Wii has been postulated as a novel rehabilitation tool that adds an additional focus on balance and proprioception into the recovery protocol. The aim of the proposed clinical trial is to investigate the effectiveness of this novel rehabilitation tool, used at home for three months after total knee replacement surgery and to assess any lasting improvements in functional outcome at one year.

Methods/Design: This will be a randomised controlled trial of 128 patients undergoing primary total knee replacement. The participants will be recruited preoperatively from three surgeons at a single centre. There will be no change to the usual care provided until 6 weeks after the operation. Then participants will be randomised to either the Wii-Fit group or usual rehabilitative care group. Outcomes will be assessed preoperatively, a 6-week post surgery baseline and then at 18 weeks, 6 months and 1 year. The primary outcome is the change in self-reported WOMAC total score from week 6 to 18 weeks. Secondary outcomes include objective measures of strength, function and satisfaction scores.

Discussion: The results of this clinical trial will be directly relevant for implementation into clinical practice. If beneficial, this affordable technology could be used by many patients to rehabilitate at home. Not only could it optimize the outcomes from their total knee replacement surgery but decrease the need for clinic-based or outpatient therapy for the majority.

Trial Registration: (ACTRN12611000291987).

© 2014 Published by Elsevier Inc.

1. Background

The use of video game systems in the rehabilitation of patients has been increasing in the medical and allied health community in the last decade. These gaming systems have been shown to be more fun than conventional exercise in a rehabilitation setting [1–3]. The Nintendo Wii™ and Wii-Fit™ (Nintendo of America, Redmond, WA, USA) is a popular, motion controlled gaming system used across a broad age range and available to the general public. It is easy to use and can deliver a varied exercise routine, with a focus on balance. The Wii has been used to rehabilitate patients with neurological injuries [4–7] but there have been few quality studies describing the use of the Wii to rehabilitate orthopaedic patients [8,9]. A feasibility study using the Wii as an adjunct to physiotherapy concluded that it has potential for patients following TKR, but the study was underpowered to find a difference. The study also used the

* Corresponding author. Tel.: +61 406 919 574.
E-mail addresses: jonathan.negus@cantab.net (J.J. Negus),
donald.cawthorne@gmail.com (D. Cawthorne), jschen@med.usyd.edu.au
(J.S. Chen), cscholes@sori.com.au (C.J. Scholes), dparker@sydneyortho.com.au
(D.A. Parker), lyn.march@sydney.edu.au (L.M. March).
[1] Tel.: +61 421 372 148.
[2] Tel.: +61 2 9463 1890.
[3] Tel.: +61 411 032 987.
[4] Tel.: +61 2 9904 7182.
[5] Tel.: +61 2 9463 1890.

Wii in the physiotherapy rooms during a supervised session, not as a home-based device [9].

The number of total knee replacements (TKR) surgeries performed annually is increasing across the developed world. The Australian joint registry has recorded over 40,000 TKRs performed each year in Australia, increasing by approximately 6000 each year. [10] There is good evidence that a TKR leads to the relief of pain but studies consistently report that 20% of patients remain dissatisfied after their surgery [11–14].

Health related quality of life scores, especially physical function, remain lower after TKR surgery than an age-matched population, despite significant improvements from their pre-operative levels [12,15–19]. Functional deficits remain in most patients at one year after TKR surgery and beyond [20–23]. It is not known if these deficits are the inevitable result of disease and surgery or if they can be improved with targeted rehabilitation. The WOMAC index is a validated subjective outcome measure for TKR arthroplasty patients and includes pain, stiffness and functional subscales making it ideal for a rehabilitation study.

Home-based exercise can be as effective as physiotherapist led rehabilitation [24–29]. Applying this gaming technology as a rehabilitation tool at home is likely to be of equal benefit to a supervised setting and may be more cost-effective.

Delaying the intensive phase of rehabilitation until six weeks post operation has been shown to be effective in the short term [30]. This could be due to swelling, pain and even anaemia delaying exercise at an intensity that would contribute significantly to muscle strength [31].

The aim of this proposed clinical trial is to investigate the effectiveness of using the Nintendo Wii-Fit for rehabilitation at home after a primary TKR. It is to be used daily for 3 months, starting at six weeks post surgery with the primary outcome being change in WOMAC total score from the six week baseline to 18 weeks and 1 year. We hypothesise that it will improve subjective and objective functional outcomes and increase adherence to recommended rehabilitation programs when compared to usual care in patients following unilateral primary total knee replacement surgery.

2. Methods / Design

The TKR-POWER study (Total Knee Replacement –Patient Outcomes Using Wii-Enhanced Rehabilitation) is a randomized controlled trial that will be conducted from one research Institute in Australia studying patients operated on at two large private hospitals by three surgeons. Participants will be required to provide written informed consent prior to starting the study. Ethics approval was obtained from the Northern Sydney Human Ethics Review Committee (HREC) and the University of Sydney HREC using the National Ethics Application form http://www.neaf.gov.au/ The TKR-POWER study has been registered at the Australian and New Zealand Clinical Trials Registry. (ACTRN12611000291987).

Following a routine decision made by the orthopaedic surgeon and patient to undergo total knee replacement surgery all potential participants will be invited to participate and undergo initial screening for eligibility prior to signing the informed consent and baseline assessments at the rooms of the three operating surgeons. Inclusion criteria are 1) Primary unilateral or bilateral TKR, 2) English speaking. Exclusion criteria are 1) Inability to return for all extra follow up visits 2) Medical conditions severely affecting their balance.

There will be no change to the usual medication or rehabilitative care during the peri-operative inpatient or postoperative period up until 6 weeks after the operation. While this may be a confounder, the groups are randomised. The study population will all have private health insurance giving them access to rehabilitation in the first 6 weeks. Previous studies that have tried to change the rehabilitation in this period have struggled with recruitment due to patient choice on early rehabilitation. The quantity and location of rehabilitation in the first 6 weeks will be recorded and analysed as a covariate.

At the 6-week visit, patients will be randomized into the Wii group or the usual rehabilitative care specific to the treating surgeon (Control group). Randomization schedules will be generated a priori using a computer generated random number sequence by an independent researcher with no participant contact. Separate randomization schedules will be made for unilateral and bilateral TKRs. Allocations will be sealed in opaque and consecutively numbered envelopes with a clear audit trail.

The control group will have their exercises demonstrated in the research institute. The Wii group will have a single home visit to ensure correct set-up of the Wii and demonstration of the exercises. Both groups will have the lead investigators telephone number for any queries on exercises.

2.1. Interventions

2.1.1. Control group

Participants will be given written descriptions and diagrams of the exercises to perform at home between 6 weeks and 18 weeks post-op. They will be given a diary to record their daily exercise. The diary will also serve as the instructions for when to perform the different exercises.

The exercises were taken from the standard rehabilitative exercise sheet given to patients by one of the major hospitals in which the majority of the participants had surgery and early rehabilitation. The advice will be to perform exercises such as squats, step-ups and calf raises for a total of 30 minutes each day (see Appendix 1). The prescription will be for 3 sets of 10 repetitions of each exercise.

2.1.2. Wii Group

Participants will receive the same exercise sheet and diary as the control group and will be asked to record any of these standard rehabilitative exercises that they perform.

Participants will also receive a second set of diaries that detail three stages of Wii exercises that increased in difficulty every 4 weeks. They will also be given a Wii console and balance board with Wii-Fit software to take home. The Wii-Fit software splits its exercises into yoga, strength training, aerobic and balance categories. The software explains the exercises and gives feedback in real time. They will be asked to perform a total of 30 minutes each day, using the Wii preferentially over the routine exercises (See Appendix 2). The prescription will be for 1 set of each exercise as the Wii-Fit automatically builds in repetitions to each set. There will be three Wii diaries corresponding to three phases of four weeks each. In each phase, new and often more challenging exercises will be introduced.

WOMAC total score at week 18 between the two study groups. Appropriate transformations of the WOMAC total score will be considered if needed to meet the statistical assumptions of the regression model.

The same analysis will be applied to the WOMAC total score at one year as a secondary analysis. In addition to these analyses, adjusting for any differences in potential confounders at baseline will be considered as well as a mixed model using all repeated measures. Appropriate statistical methods will be used for the secondary outcomes (see Table 1).

3. Summary

The aim of the proposed trial is to investigate the effectiveness of using the Wii-Fit as a novel home-based rehabilitation tool to improve function in patients recovering from primary TKR surgery. The functional improvement will be assessed by the change in their WOMAC score, from the start of intervention at 6 weeks to the end at 18 weeks. The change score at 1 year will assess any lasting effect. The comparison group will receive usual rehabilitative care. This is based on local practice where most TKR patients are discharged from supervised rehabilitation at six to eight weeks from their operation and given written instructions for ongoing exercises. There is no additional information given from that point, unless the patient actively seeks it out.

The WOMAC index was chosen as it is a subjective assessment of the patients' own recovery, and it has been well validated in a TKR population. The pain, stiffness and function subscales allow for analysis of specific areas of improvement. Secondary outcome measures such as balance and timed up and go provide further information on functional aspects of recovery but are too specific a primary outcome for this study. They are objective scores and as such, they do not take the patient's subjective assessment of their own results into account. Subjective outcomes are commonly assessed as an important marker of the success of TKR surgery, from the patient's perspective.

Current evidence supports the practice of rehabilitation at home following the initial recovery period after a TKR [26,27,41], exercises focusing on balance as well as strength and flexibility can improve self-reported outcomes [42] and intensive therapy may well provide more benefit if initiated 6-8 weeks after the surgery [30,31]. As it is not necessary to be in a supervised setting for rehabilitation after TKR, any attempts to improve results from rehabilitation can be targeted at the home environment. The use of the Wii offers the opportunity for guidance, instruction, motivation and variety on a daily basis without the need for the financial and time costs of transport to a rehabilitation facility. It may be that just increasing the volume and chronicity of exercise by using the Wii at home leads to an improved outcome. This age group is likely to benefit from a little and often approach to exercise rather than several hours of concentrated exercise two to three times a week.

This study is likely to provide a number of challenges. The delivery and set-up of over 60 Wii consoles in the subjects' homes will be time consuming with significant associated costs. There is the potential for extra home visits should the subjects encounter technical difficulties that cannot be managed over the phone. Although studies have shown this age group to be capable of using this technology, these studies have usually been carried out in a supervised setting. Our subjects will be using them unsupervised at home. The Nintendo Wii-fit has been shown to have variable levels of acceptability in the older patient population [6,43]. This could lead to subjects dropping out of the study if they find that they can't cope with or just don't like using the technology.

This is a large study with many data points to collect over the course of a year. There will be significant logistical challenges in organization of assessments and subsequent data collection. The assessments will all take place in an established research institute that is located next to the surgeons' offices. This minimizes transport issues and additional visits over and above the routine surgical follow-up visits.

There are a number of decisions that had to be made when setting this study up. During the time period between the operation and week 6, it has been determined to leave the subjects to their usual care. This means that they will attend rehabilitation or physiotherapy as deemed suitable by their surgeons and rehabilitation staff. While it would be neater for the study to standardize this period, the study population will all be private patients with access to private rehabilitation. Previous studies had difficulties recruiting from this population when they thought that they may lose their choice of post-operative rehabilitation, due to them via their insurance. Therefore, the subjects will be randomised at week 6 to minimize the effect of any differences in this period. The duration of any in-patient and outpatient rehabilitation will be recorded for future analysis as a covariate.

The Wii group will have a larger choice of exercises with the option of increasing exercise difficulty as the intervention progresses. This is different to the control group whose exercises are limited and don't change. However, this does mirror current local practice, where after six to eight weeks, most patients are discharged from formal rehabilitation with a sheet of exercises and no more rehabilitative input. Both groups will be free to perform any additional exercises of their choosing through the study period and qualitative data will be recorded at each assessment.

The Wii is hypothesised to improve subjective and functional outcome for a number of reasons. It has a significant balance component to most of the exercises. Improving the patients' balance may be a factor in their subjective assessment of recovery. The Wii produces performance scores for each exercise, which demonstrates improvements over time, generating interest and competition as well as a sense of achievement. This is hypothesised to improve compliance with a given program.

We made the assumption that for an intervention to be most feasible and cost-effective it should be home-based, intensive and should begin 6 weeks after TKR surgery. This allows for more intensive exercise once the pain and swelling have settled. The lack of supervision could offset the initial hardware costs of the console and balance board we used.

The TKR-POWER study is designed with sufficient power to detect a clinically meaningful benefit should it exist. The Wii equipment is widely available in retail outlets for a modest cost and can be reused or shared once patients have finished with it. If effective, it could optimize outcomes for patients and further reduce the need for costly outpatient

Both groups will be free to perform any other exercises prescribed or otherwise including physiotherapy, hydrotherapy, walking and sports.

2.2. Outcome Assessments

Assessments will be conducted preoperatively either when the participant is first recruited or planned for a future date, if it is more than 3 months from recruitment to the date of the operation. Assessments continue at 6 weeks (baseline), 18 weeks, 6 months and 12 months from the operation (see Table 1). Research institute staff will perform all assessments and will be blinded to the treatment allocation. Demographic information collected at baseline will include socio-economic status, comorbidities, height, weight, history of lower limb arthritis, trauma and previous orthopaedic surgery. Questionnaire data will be collected online when possible using www.myclinicaloutcomes.co.uk, otherwise directly into a database in the institute.

2.3. Primary outcomes

The primary outcome will be self-reported total score on the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index using a 5-point Likert scale (WOMAC LK3.1). The WOMAC consists of 24 items covering three subscales: pain (5 items), stiffness (2 items) and physical function (17 items). Each question has a 5 option Likert score from 0–4 giving a score ranging from 0 (no pain, stiffness or difficulty) to 96 (maximum pain, stiffness or difficulty). The WOMAC is a widely used questionnaire specifically designed to evaluate knee and hip OA. Its clinometric properties have been validated to be responsive to change in patients undergoing therapeutic interventions for OA knee including TKR [32–35].

The occurrence of any adverse events will be sought from participants during each follow-up assessment. Adverse events will be defined as musculoskeletal or cardiovascular events resulting in hospitalization and any falls for whatever reason.

2.4. Sample size

The sample size was calculated to detect a minimal clinically important difference in the WOMAC total score at 18 weeks after total knee replacement between the control and Wii groups. We defined the minimal clinically important change according to criteria suggested by Bellamy et al.; A standard deviation of 14 for calculating sample size with use of this scale, and a change of at least 7.4 points following a weak intervention, such as low dose of Voltaren (Diclofenac sodium). We calculated a sample size to detect a difference of 8 with a standard deviation of 14, a ratio of treatment to control of 1:1, a type-I error of 5% and 80% power to be 49 in each group. After allowing for a loss to follow-up of at least 20%, we calculated a sample size of 128 (64 in each group).

2.5. Statistical analysis

Data will be analysed according to intention to treat principles according to the original group assignment. A linear regression model including the WOMAC total score at week 6 as a covariate will be used to compare difference in the

Table 1
Secondary outcomes and their statistical test.

Test	Description	Statistical test
Measured at all Assessment Time points		
Oxford knee score	12 item self-reported functional score for TKR. Range 0–48.	Unpaired t-test of change score
SF-12	12-item Short-form Health Survey. Data from the 12 items will be used to construct the physical and mental component summary scores	Unpaired t-test
VAS – pain score	The patients bisect a 10 cm line with 0 being no pain and 10 being worst pain imaginable.	Unpaired t-test
HADS – Hospital Anxiety and Depression Score	Data from the 14 questions will be used to construct an anxiety and a depression summary score.	Unpaired t-test of each score
Locus of Control	Data from the 18 questions will be used to construct summary scores in – Internal, Doctors, Chance and Other people.	Unpaired t-test of each of 4 loci
AKCRS [39,40]	The objective score – Includes knee ROM & alignment. The ROM will be measured with a goniometer. Knee flexion and extension will be measured as an active movement while supine.	Unpaired t-test of ROM and score
Balance [7,36,37]	Static standing balance tests will be performed using the Wii-Fit and the Biodex balance platform with eyes open and closed with both double leg and single leg stances.	Unpaired t-test
Strength [38,39]	Maximal Isometric knee flexion and extension strength for both legs – Isometric muscle strength assessed with a hand held dynamometer seated for Quadriceps and prone for hamstrings. The best of three attempts will be recorded.	Unpaired t-test
Up and Go [40]	This is the shortened version of the Timed up and Go test. It is the time taken for the participant to get up from a chair, walk 2.44 m, turn around a cone and return to the chair to sit down.	Unpaired t-test
End of Intervention period – 18 weeks		
Exercise diary	Participants will record their exercise frequency on a written diary. The Wii group will also have their exercises recorded by the Wii console.	Unpaired t-test of mean number of exercises per week
At 18 weeks and 1 year only		
Patient satisfaction	A 3 question self-reported score.	Chi Squared

rehabilitation in this group. It is also possible that due to the interactive and personalized nature of the intervention, it will improve compliance and adherence with prescribed exercise treatment.

The study is anticipated to be completed and report on its findings in early 2015.

List of abbreviations

TKR   total knee replacement

Competing interests

The authors declare that they have no competing interests

Authors' contributions

JN conceived and designed the study. He will participate in data acquisition and analysis as well as manuscript drafting.

LM participated in methodological planning, proposed data analysis and manuscript drafting.

DP participated in the study conception and design, patient recruitment as well as giving clinical and surgical advice

CS participated in the methodological planning, data analysis and manuscript drafting.

JSC participated in study design, sample size estimations and statistical analysis advice and will help perform data analysis, interpretation and manuscript drafting.

DC will perform data acquisition and interpretation

Appendix 1

Rehabilitation Study Exercises

Heel Raises

Stand tall
Push into balls of feet
Raise heels as high as you can
Slowly lower

Sit to Stand

Bend knees, feet back and hands on the chair
With weight through both legs, lean forward "nose over toes"
Assist with upper body as you push evenly through both legs to stand
Stand tall with hips and knees straight
Return to sitting at a controlled pace

½ Squat at Wall

Feet 2-3 inches from wall
Buttocks on wall, hands on thighs
Keeping your chest up, bend your knees and slide down as far as you can control
Push through legs to return to standing

Stairs

a) Foot placement

Stand close to step
Feet shoulder width apart
Stand tall
Bend operated leg hip and knee up to place your foot on the step
Try not to hitch your hip
Return to foot to the floor

b) Step Up

Place operated leg on step
Use handrail for support as needed
"Nose over toes" body forward onto step with other leg
Return to standing

c) Step Down

Standing on the step facing down
Use handrail for support as needed
Bend operated knee to lower the other leg onto the floor

Balance Exercises

a) Step Stance

Weight shift onto front/operated leg
Tense thigh muscles
Tighten core trunk muscles by exhaling as you draw navel up
Squeeze buttocks
Lift heel of the un-operated leg
Hold for 5-10 seconds
10 reps

b) Stride Stance

Weight shift onto operated leg,
Lift heel of un-operated leg
Stand tall, don't lean
Tighten buttocks
Hold for 5-10 seconds
10 reps

c) Single Leg Balance

Do not risk falling, stand close to a stable support
Stand tall
Keep knees relaxed
Draw in navel
Squeeze buttocks
Move weight over onto operated leg, lifting other leg
Hold

290

Exercise Diary

Sydney Orthopaedic Research Institute

ID : XXX

Week	Beginning Monday	BALANCE			STRENGTH			STAIRS		
		Step stance	Stride stance	Single leg balance	Heel Raises	Sit to stand	1/2 squat at wall	Foot placement	Step up	Step down
7	10th Sep									
8	17th Sep									
9	24th Sep									
10	1st Oct									
11	8th Oct									
12	15th Oct									
13	22nd Oct									
14	29th Oct									
15	5th Nov									
16	12th Nov									
17	19th Nov									
18	26th Nov									

Place a tick in each box every time you perform that exercise in that week.

eg: the exercise was done 4 times this week

384   **Appendix 2**

385

Sydney Orthopaedic Research Institute

ID : XXX

Wii Diary 1 - Week 7 onwards

Week	Beginning Monday	Click on Wii Board in locker room		BALANCE GAMES	YOGA				MUSCLE WORKOUTS				AEROBIC		Training plus
		Stillness - Needs unlocking	Basic balance test	Table tilt	Deep breathing	Half Moon	Warrior	Torso twists	Rowing Squat	Parallel stretch *	Jackknife *	Basic Step	Hula Hoop	Cycling	
7	3rd Sep														
8	10th Sep														
9	17th Sep														
10	24th Sep														
11	1st Oct														
12	8th Oct														
13	15th Oct														
14	22nd Oct														
15	29th Oct														
16	5th Nov														
17	12th Nov														
18	19th Nov														

Place a tick in each box every time you perform that exercise in that week.

eg: the exercise was done 4 times this week   * = optional

291

References

[1] Graves L, Stratton G, Ridgers ND, Cable NT. Energy expenditure in adolescents playing new generation computer games. Br J Sports Med 2008;42(7):592–4.

[2] Nitz JC, Kuys S, Isles R, Fu S. Is the Wii Fit a new-generation tool for improving balance, health and well-being? A pilot study. Climacteric 2010;13(5):487–91.

[3] Williams MA, Soiza RL, Jenkinson AM, Stewart A. EXercising with Computers in Later Life (EXCELL) — pilot and feasibility study of the acceptability of the Nintendo(R) WiiFit in community-dwelling fallers. BMC Res Notes 2010;3:238.

[4] Lange B, Flynn S, Proffitt R, Chang CY, Rizzo AS. Development of an interactive game-based rehabilitation tool for dynamic balance training. Top Stroke Rehabil 2010;17(5):345–52.

[5] Deutsch JE, Borbely M, Filler J, Huhn K, Guarrera-Bowlby P. Use of a low-cost, commercially available gaming console (Wii) for rehabilitation of an adolescent with cerebral palsy. Phys Ther 2008;88(10):1196–207.

[6] Bieryla KA, Dold NM. Feasibility of Wii Fit training to improve clinical measures of balance in older adults. Clin Interv Aging 2013;8:775–81.

[7] Holmes JD, Jenkins ME, Johnson AM, Hunt MA, Clark RA. Validity of the Nintendo Wii balance board for the assessment of standing balance in Parkinson's disease. Clin Rehabil 2013;27(4):361–6.

[8] Baltaci G, Harput G, Haksever B, Ulusoy B, Ozer H. Comparison between Nintendo Wii Fit and conventional rehabilitation on functional performance outcomes after hamstring anterior cruciate ligament reconstruction: prospective, randomized, controlled, double-blind clinical trial. Knee Surg Sports Traumatol Arthrosc 2013;21(4):880–7.

[9] Fung V, Ho A, Shaffer J, Chung E, Gomez M. Use of Nintendo Wii Fit in the rehabilitation of outpatients following total knee replacement: a preliminary randomised controlled trial. Physiotherapy 2012;98(3):183–8.

[10] Australian Orthopaedic Association National Joint Replacement Registry. Annual report. Adelaide: AOA; 2013.

[11] Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57–63.

[12] Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. J Rheumatol 2000;27(7):1745–52.

[13] Dickstein R, Heffes Y, Shabtai EI, Markowitz E. Total knee arthroplasty in the elderly: patients' self-appraisal 6 and 12 months postoperatively. Gerontology 1998;44(4):204–10.

[14] Hawker G, Wright J, Coyte P, et al. Health-related quality of life after knee replacement. J Bone Joint Surg Am 1998;80(2):163–73.

[15] Salmon P, Hall GM, Peerbhoy D, Shenkin A, Parker C. Recovery from hip and knee arthroplasty: Patients' perspective on pain, function, quality of life, and well-being up to 6 months postoperatively. Arch Phys Med Rehabil 2001;82(3):360–6.

[16] van Essen GJ, Chipchase LS, O'Connor D, Krishnan J. Primary total knee replacement: short-term outcomes in an Australian population. J Qual Clin Pract 1998;18(2):135–42.

[17] Williams JI, Llewellyn Thomas H, Arshinoff R, Young N, Naylor CD. The burden of waiting for hip and knee replacements in Ontario. Ontario Hip and Knee Replacement Project Team. J Eval Clin Pract 1997;3(1):59–68.

[18] March LM, Cross MJ, Lapsley H, et al. Outcomes after hip or knee replacement surgery for osteoarthritis. A prospective cohort study comparing patients' quality of life before and after surgery with age-related population norms. Med J Aust 1999;171(5):235–8.

[19] Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. Arch Phys Med Rehabil 1999;80(5):572–9.

[20] Murray MP, Gore DR, Laney WH, Gardner GM, Mollinger LA. Kinesiologic measurements of functional performance before and after double compartment Marmor knee arthroplasty. Clin Orthop Relat Res 1983;173:191–9.

[21] Walsh M, Woodhouse LJ, Thomas SG, Finch E. Physical impairments and functional limitations: a comparison of individuals 1 year after total knee arthroplasty with control subjects. Phys Ther 1998;78(3):248–58.

[22] Skinner HB. Pathokinesiology and total joint arthroplasty. Clin Orthop Relat Res 1993;288:78–86.

[23] Roush SE. Patient-perceived functional outcomes associated with elective hip and knee arthroplasties. Phys Ther 1985;65(10):1496–500.

[24] Madsen M, Larsen K, Madsen IK, Soe H, Hansen TB. Late group-based rehabilitation has no advantages compared with supervised home-exercises after total knee arthroplasty. Dan Med J 2013;60(4):A4607.

[25] Mitchell C, Walker J, Walters S, Morgan AB, Binns T, Mathers N. Costs and effectiveness of pre- and post-operative home physiotherapy for total knee replacement: randomized controlled trial. J Eval Clin Pract 2005;11(3):283–92.

[26] Mahomed NN, Davis AM, Hawker G, et al. Inpatient compared with home-based rehabilitation following primary unilateral total hip or knee replacement: a randomized controlled trial. J Bone Joint Surg Am 2008;90(8):1673–80.

[27] Kramer JF, Speechley M, Bourne R, Rorabeck C, Vaz M. Comparison of clinic- and home-based rehabilitation programs after total knee arthroplasty. Clin Orthop Relat Res 2003;410:225–34.

[28] Rajan RA, Pack Y, Jackson H, Gillies C, Asirvatham R. No need for outpatient physiotherapy following total knee arthroplasty: a randomized trial of 120 patients. Acta Orthop Scand 2004;75(1):71–3.

[29] Ko V, Naylor J, Harris I, Crosbie J, Yeo A, Mittal R. One-to-one therapy is not superior to group or home-based therapy after total knee arthroplasty: a randomized, superiority trial. J Bone Joint Surg Am 2013;95(21):1942–9.

[30] Moffet H, Collet JP, Shapiro SH, Paradis G, Marquis F, Roy L. Effectiveness of intensive rehabilitation on functional ability and quality of life after first total knee arthroplasty: A single-blind randomized controlled trial. Arch Phys Med Rehabil 2004;85(4):546–56.

[31] Lin C-WC, March L, Crosbie J, et al. Maximum recovery after knee replacement–the MARKER study rationale and protocol. BMC Musculoskelet Disord 2009;10:69.

[32] Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis Cartilage 2007;15(3):273–80.

[33] Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation of WOMAC. Following THR or TKR in OA. J Orthop Rheumatol 1988;1:95–108.

[34] Bellamy N, Kean WF, Buchanan WW, Gerecz-Simon E, Campbell J. Double blind randomized controlled trial of sodium meclofenamate (Meclomen) and diclofenac sodium (Voltaren): post validation reapplication of the WOMAC Osteoarthritis Index. J Rheumatol 1992;19(1):153–9.

[35] Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. J Rheumatol 1995;22(6):1193–6.

[36] Clark RA, Bryant AL, Pua Y, McCrory P, Bennell K, Hunt M. Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. Gait Posture 2010;31(3):307–10.

[37] Howe TE, Rochester L, Jackson A, Banks PM, Blair VA. Exercise for improving balance in older people. Cochrane Database Syst Rev 2007;4:CD004963.

[38] Zeni Jr JA, Snyder-Mackler L. Early postoperative measures predict 1- and 2-year outcomes after unilateral total knee arthroplasty: importance of contralateral limb strength. Phys Ther 2010;90(1):43–54.

[39] Carpenter MR, Carpenter RL, Peel J, Zukley LM. The reliability of isokinetic and isometric leg strength measures among individuals with symptoms of mild osteoarthritis. J Sports Med Phys Fitness 2006;46:1–5.

[40] Rikli R, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act 1999;7:129–61.

[41] Minns Lowe CJ, Barker KL, Dewey M, Sackley CM. Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials. BMJ 2007;335(7624):812.

[42] Piva SR, Gil AB, Almeida GJM, DiGioia III AM, Levison TJ, Fitzgerald GK. A balance exercise program appears to improve function for patients with total knee arthroplasty: a randomized clinical trial. Phys Ther 2010;90(6):880–94.

[43] Laver K, Ratcliffe J, George S, Burgess L, Crotty M. Is the Nintendo Wii Fit really acceptable to older people? A discrete choice experiment. BMC Geriatr 2011;11:64.