UNIVERSITY OF LIVERPOOL

DEVELOPMENT AND VALIDATION OF A PATIENT REPORTED OUTCOME MEASURE ASSESSING GLOBAL PHYSICAL FUNCTION IN PATIENTS SUFFERING FROM LOWER LIMB OSTEOARTHRITIS

THESIS SUBMITTED IN ACCORDANCE WITH THE REQUIREMENT OF THE UNIVERSITY OF LIVERPOOL FOR THE DEDREE OF MASTER IN PHILOSOPHY

By

Mohammad Abdul Hafidz Haji Yakob

Student Number: 201228806

SUPERVISORS

J. Alsousou M. Roebuck S. Frostick

February 2020

Department of Molecular and Clinical Cancer Medicine

Abstract

Background: Osteoarthritis (OA) is a common musculoskeletal disease that affects multiple joints. Main symptoms include pain and loss of function. The status of a patient can be measured by using patient reported outcome measures (PROMs). PROMs are also used to assess effectiveness of treatment.

Aims: Develop a novel PROM that can assess level of function, pain and overall health in patients suffering from OA of the hip or knee.

Methodology: Development was performed in four stages. For the first stage, 50 patients were recruited for qualitative interviews to identify activities that are affected by the symptoms of OA in these patients. Secondly, a group of 16 participants experienced in treating OA were interviewed and asked to rank items in order of relevance for assessing patients with OA. The third stage consisted of construction of the new PROM and a qualitative pre-test of the prototype with 12 patients. The highest-ranking items from each subdomain were used to create a list of 20 items to include in the questionnaire which was further supplemented by two additional questions: a body map and a VAS. A finalised list of themes (n=70) were ranked within their subdomains to create the new PROM. Finally, the last stage involved recruiting 60 patients for quantitative analysis of the new PROM for validation.

Results: A total of 86 unique themes were identified through qualitative interviews, which were then grouped together as items and categorised into five domains; lower limb function, upper limb function, role limitation, pain symptoms and general health. Additional themes were identified by the expert participants. The finalised PROM consisted of 22 questions containing the five subdomains. Content validity was confirmed by using the methodology outline in the first two stages of this project. The questionnaire showed good internal consistency and reliability. Correlation of the new PROM against the oxford hip score, oxford knee score, WOMAC showed fair, moderate and strong correlations across the related subdomains. The new PROM showed poor correlations overall against the SF12. There were no identified floor or ceiling effects.

Conclusion: From this preliminary study a variety of themes can be identified from patients that suffer from OA. Using themes and input from a panel of experts a prototype of the new PROM has been developed that can assess overall physical function, pain symptoms and general health. The new PROM was validated and has been shown to have good internal consistency and reliability. It has also correlated well against similar validated PROMS and shows no floor and ceiling effects. Further work is still required to complete validate the new PROM, assess its sensitivity to detect change and minimal clinical important difference.

Declarations

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Declaration of Originality

This thesis is a production of my own work in collaboration with the Musculoskeletal Science Research Group, produced during my time at the Department of Molecular and Clinical Cancer Medicine, University of Liverpool between December 2016 and February 2020. The thesis was written by me with guidance from my supervisors Mr Joseph Alsousou, Dr Margaret Roebuck and Professor Simon Frostick.

Acknowledgements

I cannot imagine how I have managed to accomplish what I have without the constant support, guidance and help received along the way. First and foremost, I would like to express my most extreme gratitude to my primary supervisor Professor Simon P Frostick, who took a chance and championed my application to return to the UK to not only complete this MPhil project but also supported me throughout the completion of my training as an Orthopaedic surgeon. Even through his illness he has remained dedicated towards seeing this idea become a reality and I am extremely saddened that he has left.

I would also like to express my thanks to Dr Margaret Roebuck who has mentored me throughout the project, especially during the thesis writing process. Her encouragement, enthusiasm and knowledge have been a constant source of motivation. I do not think I would have even been able to start without her advice and direction.

I am also indebted to Dr Joseph Al Sousou, who has been a great mentor and supervisor. His opinions and criticisms have steered me through the daunting task of completing this body of work. Special thanks also go out to Ms Amanda Wood, for the phenomenal groundwork that she has laid with initiating the project and keeping things in order. I would also like to express special gratitude to the Bruneian Government for and their sponsorship throughout my years in the UK.

This Project would not have been completed without the partnership of my great friend and colleague Dr Nor Azmi Mohammad, with whom I have struggled and rejoiced with side by side since the beginning. And I would also like to express my deepest gratitude to my parents, Haji Yakob Yusof and Khalijah Ali Ahmed, whose prayers and constant support helped me through the toughest times.

And lastly, I leave my final and most heartfelt dedication to my wife, Khairunnisa Isa, and my three children: Hakeem, Zaira and Hamza. They are my inspiration to continuously grow and improve. Without their unwavering patience and encouragement, I would not have been able to press on and without their sacrifices I would not have been able to succeed.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Osteoarthritis

1.1.1 Definition

Osteoarthritis (OA) is a debilitating joint disease that commonly affects individuals (Litwic, et al., 2013). It is hallmarked by damage and degeneration of cartilage in the affected joint, with associated periarticular changes which include remodelling of subchondral bone, ligament laxity, osteophyte formation, synovitis and muscle weakness (Hutton, 1989). The most common joints affected are the knees, hips, and small joints in the hand (Prieto-Alhambra, et al., 2014). Pain is the predominant symptom of OA and it imposes a biopsychosocial burden on the individual (Neogi, 2013). Management of pain is the main driving force for individuals to seek medical care (Hadler, 1992).

1.1.2 Epidemiology

In the UK, over 6.5 million people have sought help for OA from their general practitioners (GP), and it is estimated that in patients over the age of 45 1 in 5 suffer from OA of the knee (Arthritis research UK). Based on work by Yu et al (2015) the annual consultation incidence of OA in persons ≥ 15 years of age was 8.6/1000 in men and 10.6/1000 in women. The study also demonstrated an increase in incidence with increasing age that peaks in the 75-84 age group. Based on this data it was concluded that per year a new diagnosis of OA in at risk groups was 9 in 1000. Globally OA of the knee is responsible for approximately 85% of the burden of OA worldwide (Global burden of disease study, 2016).

1.1.3 Pathophysiology

Historically, it was believed that OA is caused by 'wear-and-tear', however we now understand it to be a dynamic process due to imbalance between destruction and repair of joint tissues (Hunter, et al., 2014). This complex pathogenesis entails the interplay of inflammatory, metabolic and mechanical factors which ultimately leads to destruction of the joint (Hunter, et al., 2014). OA starts with alteration in cartilage composition and loss of integrity making it more susceptible to damage, which will eventually affect the deep layers of calcified cartilage (Leoser, et al., 2016). In an attempt to repair this damage, cells that maintain cartilage (chondrocytes) will increase their synthetic activity, leading to a pro-inflammatory response that affects all the joint tissues eventually leading to the degenerative changes (Hsia, et al., 2018).

1.1.4 Diagnosis

The standard for diagnosis of OA is through clinical assessment which includes obtaining history of symptoms and physical examination of clinical signs. Diagnostic criteria are available from the American College of Rheumatism (ACR) and the European League Against Rheumatism (EULAR) (Altman, et al., 1986) (Zhang, et al., 2010). Radiographic criteria have also been created by Kellgren and Lawrence (1963), however there is limited correlation between clinical and radiographic signs (Hannan, et al., 2000).

1.1.5 Treatment

According to guidance from the National Institute of Clinical Excellence (NICE), a comprehensive clinical history and physical examination is recommended to assess the effects of OA on function, quality of life, social interaction, leisure activities and mood (NICE, 2014). Management of OA consists of a stepwise approach with non-pharmacological

treatment preceding pharmacological therapy as the recommended first line treatment (Nelson, et al., 2014). Non-pharmacological methods that have proven to be beneficial include physical exercise (Fransen, et al., 2015) and weight loss (Messier, et al., 2004). Pharmacological treatment which is recommended for OA include paracetamol and Non-steroidal anti-inflammatory drugs (NSAIDs) (Nelson, et al., 2014).

After exhausting non-operative treatment, patients with severe OA experiencing a decline in quality of life are recommended to be referred to Orthopaedic surgeons for joint replacement surgery (NICE, 2014). A review article by Wood et al (2013) on the management of osteoarthritis concluded that joint replacement surgery of the hip and knee leads to better quality of life, higher patient satisfaction and longevity. A way of collecting and monitoring data on joint replacement surgeries is through national joint registries (NJR). In the UK, patients undergoing elective hip or knee replacement operations funded by the NHS are asked to complete patient reported outcome measure (PROM) questionnaires before and after surgery (NHS UK, 2019). This is done to assess improvement in health from the patient's perspective.

1.2 Impact of OA

The burden of OA on a person includes pain, poor quality of life, limitations in daily activity and disability (van der Waal, et al., 2005). The World Health Organisation has defined disability as an umbrella term to encompass impairment in body function or structure; limitation of activity in performing a task or action; and restriction in participation of a person in life situations (World Health Organisation). There is also a significant economic cost to OA for the individual, society and health care providers (Hunter & Bierma-Zeinstra, 2019).

1.2.1 Daily activity and disability

OA is the most common musculoskeletal (MSK) disorder affecting elderly patients (Felson, et al., 1987). MSK disorders together form the 4th most common cause of morbidity (Prince, et al., 2015). Pain and loss of function from OA can lead to inactivity, eventually causing loss of quality of life (Clynes, et al., 2019). Pain, swelling and stiffness can also cause impairment in activities of daily living (ADL) like using computers, driving cars, walking, using stairs and carrying objects (Marshall, et al., 2018).

Clynes et al (2019) have demonstrated that a diagnosis of lower limb OA in older patients is associated with difficulties in mobility, ability to self-care and perform ADL. Zhang et al (2002) found that when comparing persons with and without a diagnosis of hand OA there was a significant difference in functional limitation in activities involving the hand such as writing, gripping and using small objects.

1.2.2 Social participation

Social participation includes interaction in social activities, going to work and managing one's home. Participation restriction is the limitation of this interaction and is proven to be related to poor health outcomes (Wilkie, et al., 2007). Joint pain, depressive symptoms and environmental barriers are associated with participation restriction reinforcing the biopsychosocial effects of OA (Theis, et al., 2013).

1.2.3 Depression and anxiety

A review by Stubbs et al (2016) looking at studies that investigated depression and anxiety in people with OA showed a pooled prevalence of 19.9% for depression and 21.3% for anxiety symptoms. The relative risk for depression and anxiety compare to those without OA were 1.17 and 1.35 respectively. The study concluded that a fifth of patients with OA suffer from

depression and anxiety, however it was unclear whether this proportion is increased compared to people without OA.

1.2.4 Economic cost

The economic costs of OA can be viewed as the direct cost of medical treatment and the indirect cost to the individual and healthcare service. It is estimated that the cost of treatment of OA in high-income countries like USA, UK, Canada and Australia ranges between 1 to 2.5% of their gross domestic product (GDP) (Hunter, et al., 2014). The indirect cost due to loss of productivity and loss of income are also substantial and are often overlooked when considering the burden of OA (Gupta, et al., 2005). In the most recent public health bulletin by Arthritis Research UK, the total cost of performing hip and knee replacements for Liverpool local authority was £5.7 million in 2011-12 (Arthritis research UK, 2019). This cost is likely to have increased given the trend of increasing incidence of OA.

1.2.5 Kinetic Chain

The kinetic chain system is a concept borrowed from engineering that is widely accepted in the realms of physical rehabilitation (Butler & Major, 2003). The human body can be likened to a linked system of rigid segments (limbs and spine) jointed together by pin joints (joints) (Karandikar & Vargas, 2011). When both ends of this system are fixed in place no movement can happen, application of an external force at one end causes transfer of force to each adjacent segment. The kinetic chain concept was popularised by Steindler who categorised a chain as either being open or closed, depending whether or not the terminal segment is fixed or free (Karandikar & Vargas, 2011).

Clinical applications of the kinetic chain concept include the improvement of core stability to provide a more stable base for distal function and ambulation (Dillman, et al., 1994).

Understanding this helps explain how pathology in the hip or knee could eventually lead to shoulder impingement or scapular dyskinesia (Burkhart, et al., 2003). Due to this, OA of the hip and knee may also have an impact to overall function of a patient.

1.3 Joint arthroplasty performance monitoring

1.3.1 National joint registry

The first joint registry that was establish was in Sweden in 1976 for knee replacements (Herberts, et al., 1989). This led the way in NJR and evidence-based surgery. Countries now with their own NJR include UK, France, Germany, Norway, Denmark and USA.

The NJR in the UK was started in April 2003, however submission of data for NHS organisations only became mandatory as of April 2011. Data is collected from England, Wales Northern Ireland and the Isle of Man. (NJR, 2020)

1.3.2 NJR purpose

The NJR collects data on the type of surgery performed, implants used and method of implantation, PROM data and final outcome data. Currently the NJR collects data on hip, knee, ankle, elbow and shoulder joint replacements. This data helps to monitor the performance of implants and effectiveness of different operations, with the aim to improve clinical standards, including raising the alarm for any early warning issues relating to patient safety. (NJR, 2020)

1.3.3 NJR Data from the UK

NJR published the 16th annual report in 2019. Since data collection started, there are now 2.85 million recorded entries in the registry making it the largest and most comprehensive national joint registry in the world. It showed that primary joint replacements were

performed for OA in 92% of hip replacements and 96% of knee replacements and that the majority of operations were performed in patients over the age of 75 years (NJR Annual Report, 2019). Compared to the previous year findings included a fall in revision surgery and an increase in the number of hip, ankle, elbow and shoulder replacements (NJR Annual Report, 2019). Table 1 provides a break down of total number of joint replacements completed list by joint.

Table 1: Total number of replacements recorded by joint

Joint replaced	Total number in registry
Hip	1,091,892
Knee	1,193,830
Ankle	5,587
Shoulder	37,916
Elbow	3,573
Total	2,332,798*
	*only includes entries with usable data

1.3.4 PROM data

Data collected by the NJR pertaining to the aspects of the operation are important however this data does not capture the outcomes of surgery from the patient perspective. This why the NJR also collects PROM data to obtain valuable information about the pain, mobility and general health of the patient following surgery. The NHS collects this PROM data before surgery and at six months after surgery only. But as patients' condition may continue to change over a longer term the NJR has permission to continue to collect this data at one, three and five years after surgery. PROMs currently used in the NJR are the EuroQol 5D (EQ-5D), EQ Visual analogue scale (EQ VAS), Oxford Hip Score (OHS) and Oxford Knee Score (OKS). (NJR, 2020)

1.4 Patient reported outcome measures

1.4.1 Definition

PROMs are validated questionnaires that are completed by patients themselves and are used to measure their functional status or wellbeing (Dawson, et al., 2010). PROMs collect information from the patient point of view, and is useful in monitoring patient status over time, as in the case of joint replacement surgery. A patient would complete a PROM by responding to each question according to their perceived status. The responses are usually assigned a numerical rating or score which can then be combined to represent an overall score for the concept (for example pain level) being measured (Dawson, et al., 2010).

1.4.2 PROM uses

PROMs can be used in different ways as the information obtained is relevant to the clinicians, healthcare services, researchers, and patients themselves. Originally PROMs were designed to assess outcomes in clinical trials, they are now also used to monitor patient progress and evaluate treatment efficacy. PROMs are also used as a tool to measure quality improvement, benchmark service providers, evaluate cost effectiveness and monitor performance. As we have seen in the case of NJR data. For patients, PROMs help them to gain insight into their own condition and provides information to allow them to make informed choices in selecting treatment options or service providers. (Nelson, et al., 2015)

1.4.3 NHS initiative

In April 2009, the NHS initiated the *National PROMs Programme* to collect information on how well the health service was treating patients (NHS England, 2018). Four operations were included in this programme: total hip replacement; total knee replacement; varicose veins;

groin hernia surgery. As of October 2017, PROMS collection for varicose veins and hernia surgery have been stopped.

1.4.4 PROM types

PROMs can either be generic questionnaires measuring quality of life or health in general, for example the Short form 36 (SF-36) which has eight subdomains that pertain to general health (Ware-JE & Sherbourne, 1992). And in the case of the NJR the EQ-5D.

PROMs can also be more specific to a condition or a joint, as with the use of the Oxford hip score and Oxford knee scores in the NJR.

1.4.5 PROM development principles

The consensus-based standards for the selection of health measurement instruments initiative (COSMIN) is an international multidisciplinary team of researchers with expertise in development and evaluation of outcome measures. The COSMIN Study Design checklist (Mokkink, et al., 2019) is a guide to ensure that PROM development is to a recommended standard. This checklist covers general recommendations for designing a PROM and standards to evaluate specific measurement properties or psychometric criteria (see Appendix A).

Development of a robust questionnaire is a multi-step process. Firstly, the intended population the PROM will be used for must be defined, for example all adult patients with OA. Secondly, the questions used to obtain information from this group must be relevant and useful. Thirdly, the PROM must be tested for its validity.

As PROMs are meant to capture information from the patient point of view, one way to understand what is relevant for them is to collect qualitative data. This form of data can

provide a list of activities that can be converted into concepts, for example the ability to walk. From these concepts, questions or items can be formulated and included into a questionnaire. Often groups of questions will ask similar features about a common concept and these questions become a subdomain within the questionnaire. An example of a subdomain is the pain subdomain of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

Once the items to be include in the PROM are decided and questions have been designed a duty falls on the researcher to ensure that the questionnaire is valid for use in the intended population (Fitzpatrick, et al., 1998)

1.5 PROM examples

1.5.1 Joint specific PROMs

1.5.1.1 OHS

The OHS was developed in 1996 with the intention of being a patient centred outcome measure to assess the benefit of treatment of hip OA (Dawson, et al., 1996). It was developed specifically to assess the benefit of total hip replacement in patients suffering from hip OA. The developers of the OHS believed that a shorter questionnaire would be more practical and simpler to use, and were able to demonstrate that the OHS was more sensitive in detecting clinical change (Dawson, et al., 1996).

The OHS consists of 12 questions (see Appendix B). It assesses pain and function of the hip with regards to daily activities. Examples of concepts it investigates include walking, climbing stairs and using transportation. There is a recall period of 4 weeks and each item has five

possible responses based on the level of difficulty the patient has (score 1-5). The scores can be added together to give an overall score. (Wylde, et al., 2005)

In its original form the OHS used lower scores to reflect better function. However, with its widespread use in the orthopaedic world there has been some considerable confusion in its application. In light of this, the originators of the OHS published recommendations of its usage in 2007 stating that the OHS should be scored from 0 – 48, with higher scores being better function. (Murray, et al., 2007)

The main strengths of the OHS are that it is a short PROM to complete and has good reliability (Wylde, et al., 2005). It is also good at detecting change, and is more sensitive in this regard when compared to the SF-36 and the Arthritis Impact Measurement Scale (AIMS) in patients undergoing total hip replacements (Dawson, et al., 1996). The OHS has also been translated and validated into other languages including Italian (Martinelli, et al., 2011) and Danish (Paulsen, et al., 2012).

1.5.1.2 OKS

The OKS was developed in 1998 by the same research group as the OHS to assess the benefits of undergoing a total knee replacement (TKR) (Dawson, et al., 1998). The development team used the same design philosophies to develop the OKS and found that it was sensitive in detecting clinical change and had good reliability (Dawson, et al., 1998).

The OKS is also a 12 item questionnaire with similar questions as the OHS and covers aspects of pain and level of function in regard to knee OA (see Appendix C). The recall period is similarly 4 weeks. As in the OHS it was originally scored on a scale of 12 – 60 but as per the recommendations from the original development team scoring the OKS from 0 – 48 is recommended. (Murray, et al., 2007)

Although originally validated for patients undergoing TKR, the OKS has since been used as an outcome measure for other clinical studies to assess different forms of treatment for OA including intra-articular injection of joint lubricating supplements (Clarke, et al., 2005), knee realignment surgery as in high tibial osteotomy (Weale, et al., 2001) and following unicompartmental knee replacement (Reilly, et al., 2005). The OKS has also been translated for use in Sweden as part of their nation joint registry (Dunbar, et al., 2000).

Both the OHS and the OKS are popular joint specific PROMs which are short and are sensitive in detecting clinical change. They are frequently used as outcome measures in clinical trials, clinical audits and as one of the PROMs in the NJR in the UK. (Murray, et al., 2007)

1.5.2 Condition specific

1.5.2.1 WOMAC

The WOMAC is a PROM used to evaluate patients with hip or knee OA. Bellamy et al published two validation studies to assess the measurement properties of the WOMAC. One to establish its use as a PROM to assess the effects of joint replacement for hip or knee OA (Bellamy, et al., 1988) and the other to assess non-operative treatment of OA with NSAIDs (Bellamy, et al., 1988).

WOMAC has 24 questions that evaluate three subdomains: pain, stiffness and physical function (see Appendix D). The pain subdomain has 5 questions, stiffness has 2 and physical function has 17 questions. Each item in WOMAC has 5 responses (0 – 4) scaling in severity with higher scores indicating higher levels of symptoms or physical disability (McConnell, et al., 2001). The sum of the scores from each subscale can be used separately or totalled to derive an index score. WOMAC has also been validated so that the responses use a visual analogue scale (VAS) instead of discrete values (McConnell, et al., 2001).

The WOMAC is another popular PROM to assess many treatments for patients with OA. Harris et al (2016) systematically reviewed PROMs used in hip and knee OA and concluded that WOMAC was the second most promising PROM based on the evidence available testing its psychometric properties as well as its development process. WOMAC has also now been used to assess patients with other conditions aside from OA including patients with lower back pain, rheumatoid arthritis and fibromyalgia (Wolfe & Kong, 1999). It has also been translated into other languages and validated (Wolfe & Kong, 1999).

1.5.2.1 Michigan body map

The Michigan body map (MBM) was developed because of inclusion of pain location the in Fibromyalgia Survey Criteria in 2011. The Fibromyalgia Survey Criteria is a screening tool used to diagnose fibromyalgia, a chronic pain syndrome. The MBM's aim was to further refine the Fibromyalgia Survey Criteria for localising areas of pain that a patient has. A point of interest of the MBM is that it is a graphical depiction of the human body and the patient is instructed to select the parts of the body (35 sites) where they experience pain. The MBM has demonstrated good ease of use and straightforward scoring. (Brummet, et al., 2016)

1.5.3 Generic PROMs

1.5.3.1 SF-36

The SF-36 is a 36-item questionnaire developed to measure general health and quality of life. It was developed by the RAND Corporation by selecting items from a much longer questionnaire used in the medical outcomes study, a large multi-site study conducted in more than 22,000 patients to explain variations in patient outcomes (RAND Corp., 2019). SF-36 was validated initially among general medical populations and was found to have good reliability

between patient groups with similar health (Brazier, et al., 1992). Three aspects of general health that are covered by SF-36 are: functional ability, wellbeing and overall health (Patel, et al., 2007).

Eight domains are covered in the SF-36: bodily pain, physical function, role limitations due to physical health problems, role limitations due to emotional problems, energy levels, social functioning, emotional well being and general health perception (Patel, et al., 2007). Scoring instructions for each domain can be obtained from the RAND website, which gives a weighted score or each domain from 0 – 100, with higher scores indicating better health (RAND Corp). A modified scoring method is also possible which derives a physical component summary (PCS) and a mental component summary (MCS) from the 8 domain scores, but this method requires ample population data as it is adjusted by the population mean and standard deviation (Patel, et al., 2007).

SF-36 is used across many different medical conditions including angina and renal dialysis (Marquis, et al., 1995) (Meyer, et al., 1994). SF-36 has also been used assess health status, determine treatment effectiveness, compare different treatments for different types of orthopaedics conditions (Patel, et al., 2007).

1.5.3.2 SF-12

The SF-12 uses a subset of questions from SF-36 to create a shorter 12-item questionnaire that is believed to derive similar measures of PCS and MCS while taking a shorter amount of time to complete (Ware, et al., 1996). Two items from the physical function, mental health and the physical and mental role limitation domains of the SF-36 are used to make up 8 questions of the SF-12, while one item from the remaining four domains (body pain, energy

levels, social function and general health) make up the remaining four (see Appendix E) (Ware, et al., 1996).

SF-12 uses proprietary owned scoring algorithm to obtain the PCS and MCS, also based on population mean scores (Ware, et al., 1995). A study by Jenkinson et al (1997) demonstrated that the PCS and MCS scores of SF-12 and SF-36 in three separate population groups (heart failure, sleep apnoea and inguinal hernia surgery) were virtually identical. Based on this the SF-12 may be the PROM of choice in certain scenarios given its shorter completion time, however, as only one or two items are used for each original subscale this limits the precision of the SF-12 (Patel, et al., 2007).

1.5.3.3 EQ-5D

The EQ-5D was established by the EuroQol Group in 1987 as a generic outcome measure to assess health-related quality of life in a patient. It consists of five domains which are: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each domain can be answered using a three-level system: no problems, some problems, severe problems. The combination of responses from the five domains allows for 243 unique health states. There is also a visual analogue scale for self-assessment of one's state of health on a scale from 0 – 100. The EQ-5D was developed as a supplementary tool to be used alongside other PROM. (Rabin & De Charro, 2001)

1.6 Questionnaire validation

Validation of a questionnaire ensures that the questionnaire measures what it intends to measure (Polit & Beck, 2006). As reviewed by Gagnier et al. (2017), a PROM should satisfy certain psychometric criteria to confirm its validity. There are different types of validity which include content validity, reliability and accuracy, and criterion validity.

1.6.1 Content validity

For content validity, qualitative analysis of structured interviews with patients and opinions from a panel of experts in the field can generate a list of relevant items. The items selected by this process would confirm the content validity of the questionnaire.

1.6.2 Reliability

To be trusted a questionnaire needs to be reliable. This means that repeating the questionnaire within a suitable timeframe, for instance within two weeks, should produce the same results providing that there have not been any significant changes with the condition of the patient. To assess reliability the correlation or level of agreement between the results of two questionnaires filled in by the same patient at different times can be measured (Bolarinwa, 2015).

1.6.3 Internal consistency

Within the questionnaire, items grouped together should all aim to measure the same concept. This can be confirmed by measuring the internal consistency of the questionnaire or its individual subdomains. A method of measuring this is to use Cronbach's alpha coefficient, which is presented as a value between 0 and 1. The closer the value is to 1 the better the

internal consistency. If internal consistency is found to be low then items can be sequentially removed until the best Cronbach's alpha coefficient is derived. (Cho, 2016)

1.6.4 Criterion validity

Criterion validity aims to confirm whether information obtained from a PROM can predict the true status of a patient (Liang, et al., 2014). In other words, this is the accuracy of a questionnaire. This form of validity requires a 'gold standard' to measure against. Criterion validity in the clinical setting is difficult to measure as the perception of pain and physical function are subjective concepts with no true gold standard test to compare against. Accepted methods off confirming criterion validity are to measure correlations against other PROMs which are intended to measure the same or similar concepts. If the new PROM can demonstrate good correlation with other established PROMs, then it can be considered to have satisfactory criterion validity.

Subtypes of criterion validity include convergent and divergent validity, which aim to confirm correlation between similar concepts and no correlation between unrelated concepts respectively.

1.6.5 Floor and ceiling effects

The total scores from a questionnaire may sometimes demonstrate floor and/or ceiling effects. This means that a higher than acceptable proportion of patients are reporting either the maximum or minimum score. This could indicate that the patient is not able to express their level of status which may exist beyond the confines of the questionnaire. Acceptable limits of floor and ceiling effects are 15% of total scores. This can be calculated by measuring the frequency of maximum or minimum scores within the validation sample. (Lim, et al., 2015)

1.6.6 Sensitivity to detect change

When a patient is subjected to an intervention, their status may change. A change in status can be reflected by a difference in scores of a questionnaire completed before and after the intervention. For example, the difference in pain scores before and after joint replacement surgery should be statistically different (McQuay, 2004). In validating a questionnaire for sensitivity, the effect size can be measured statistically. The minimum clinically important difference is a concept of measuring sensitivity in clinical practise (King, 2011).

1.6.7 Patient burden

Aside from satisfying the above psychometric criteria, other aspects that make a questionnaire suitable for use include its ease of delivery, time required to complete, and level of comprehension needed to answer the questions. All these factors reflect a burden on the patient. For a questionnaire to be effortless to complete it should be simple in its understanding and presentation.

1.7 Research Methods

As already mentioned, the development of a PROM is a multi-step process. To that end different types of research methods must be used for different stages of the development process. Broadly, there are two general methods of obtaining data; qualitative and quantitative. There is also a mixed method approach which incorporates both elements today. In terms of this project, either qualitative or quantitative methods will be used exclusively for a particular stage.

1.7.1 Qualitative Research

The process of designing a questionnaire requires understanding of the population being observed. Qualitative research is a useful method of obtaining data pertaining to the patient experience, through a systematic analysis of textual material obtained from talk or observations (Malterud, 2001). By exploring these experiences, the researcher can begin to develop an understanding behind the motivations and beliefs of the observed population, and through interpretation of these findings one is able extract points of view that may be more common to that population (Malterud, 2001).

One way to obtain data through qualitative research is through interviews with the study participants. Common themes and underlying ideas and experiences can then be observed from analysis of these interviews. There are several different types of established qualitative research approaches; case-study, ethnography, grounded theory, and phenomenology (Grossoehme, 2014). This study will apply phenomenology research as its qualitative research approach to explore patient experiences living with OA and the limitations they observe in their daily lives, with the disease of OA being the phenomenon being experienced by the individual patients.

The common themes extracted through this qualitative research will be used as the basis of a conceptual framework to guide the development process of designing a novel PROM. This PROM will then be assessed for its validity through psychometric analysis.

1.7.2 Quantitative Research

Quantitative research involves the gathering and analysis of numerical data. There are four different types of quantitative research: descriptive, correlational, casual-comparative and experimental. (Creswell, 2009)

The quantitative research involved in validation of a PROM will mainly fall into the descriptive and correlational types, as the data obtained will be used to test the robustness of the PROM through different statistical tests. The validation of the PROM has already been discussed in section 1.6.

1.8 Project Rationale

PROMs have their merit in the clinical setting in facilitating our understanding of certain conditions from the point of view of the patients. As it currently stands the NJR in the UK is actively utilising PROMs to capture patient reported data to measure outcomes of joint arthroplasty. As already mentioned, patient undergoing hip or knee arthroplasty are given a battery of PROMs to complete in one sitting (NJR, 2020), this can become a burden for the patient due to the time it takes to complete these PROMs.

Part of the reason why multiple PROMs are used is due to the inherent limitations for the joint specific PROMs to capture data regarding comorbidities that way be a contributing factor in a patient's loss of function or disability (Lim, et al., 2006). It has also been recognised that aside from the physical limitations due to pain, OA can also have social and psychological impacts on patients (Wilkie, et al., 2007) (Theis, et al., 2013).

This project proposes to design a novel PROM that will aim to reduce patient burden by creating a condition specific PROM that will assess patients' overall level of function in the context of suffering from OA. The project will also attempt to incorporate alternative methods of questioning (VAS and a body map) to gauge a patient's insight into their level of function and to graphically localise where they are having pain. For such a PROM to be suitable for use in the scope of our clinical practise it should then be validated to satisfy the previously mentioned psychometric criteria (Pynsent, 2004).

Project aim

"To design a patient reported outcome measure designed for patients with lower limb osteoarthritis that can assess their level of global function, level of pain and overall general health which satisfies psychometric criteria."

Project objectives

- Define the population of patients the PROM will be developed for.
- Recruit patients to qualitatively explore symptoms, aspects and activities that are affected by OA.
- Recruit expert participants to qualitatively assess features of OA and the functional limitations that are caused by it.
- Design a novel PROM based on information obtained from recruited participants.
- Preliminary test of PROM to qualitatively ascertain ease of use.
- Quantitative research to validate PROM using mentioned psychometric criteria.

Chapter 2: Methodology

2.1 PROM design outline

Development of the PROM was performed in four stages.

First stage

Derive a list of themes related to the effects of osteoarthritis in the target population which would then become the basis for creating questions (items) in the new PROM.

Second Stage

Recruit a panel of experts to provide additional themes that can potentially be included into the questionnaire and to ask them to rank all the included themes in order of priority.

Third Stage

Construction of the new PROM by selecting themes to create items and to include a body map of joint pain and a VAS for patient satisfaction. A qualitative pre-test will be conducted in this stage before the final version of the PROM is validated.

Fourth Stage

Validation of the questionnaire according to the various psychometric criteria already mentioned. A flow chart of the stages is depicted in the following figure (Figure 1)

All statistical calculations were performed using IBM SPSS Statistics for Windows (Version 25.0., Released 2017, Armonk, NY: IBM Corp.).

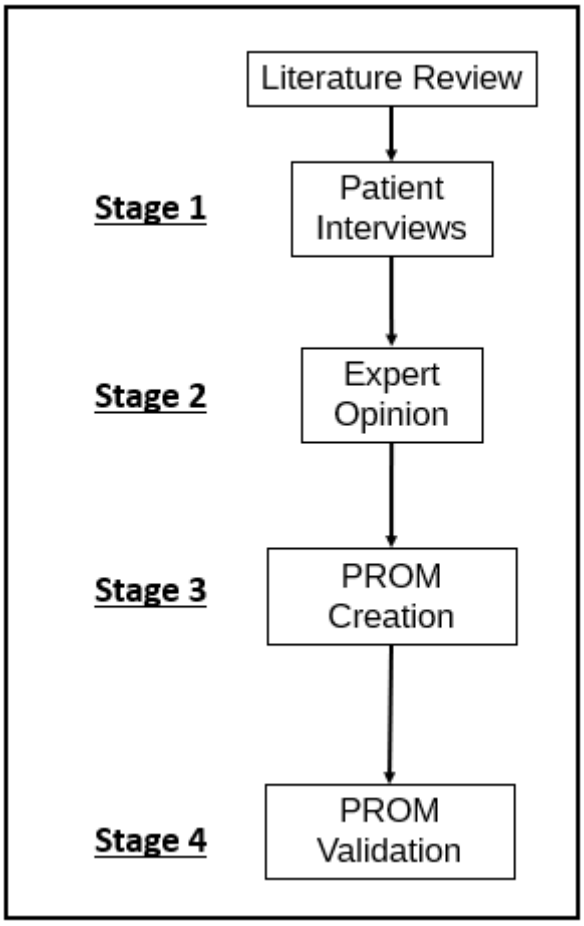


Figure 1: Flow chart of questionnaire creation

Each stage was regarded as an independent step in the project. A new sample of participants were recruited to complete each stage. Each participant was informed about the project objectives in general and more specifically about the stage they would be involved in. Participation was voluntary and informed consent was obtained from each participant. A summary of the number of participants recruited is shown in the following figure (n = number of participants). Detailed demographics of patient participants are included in the following sections.

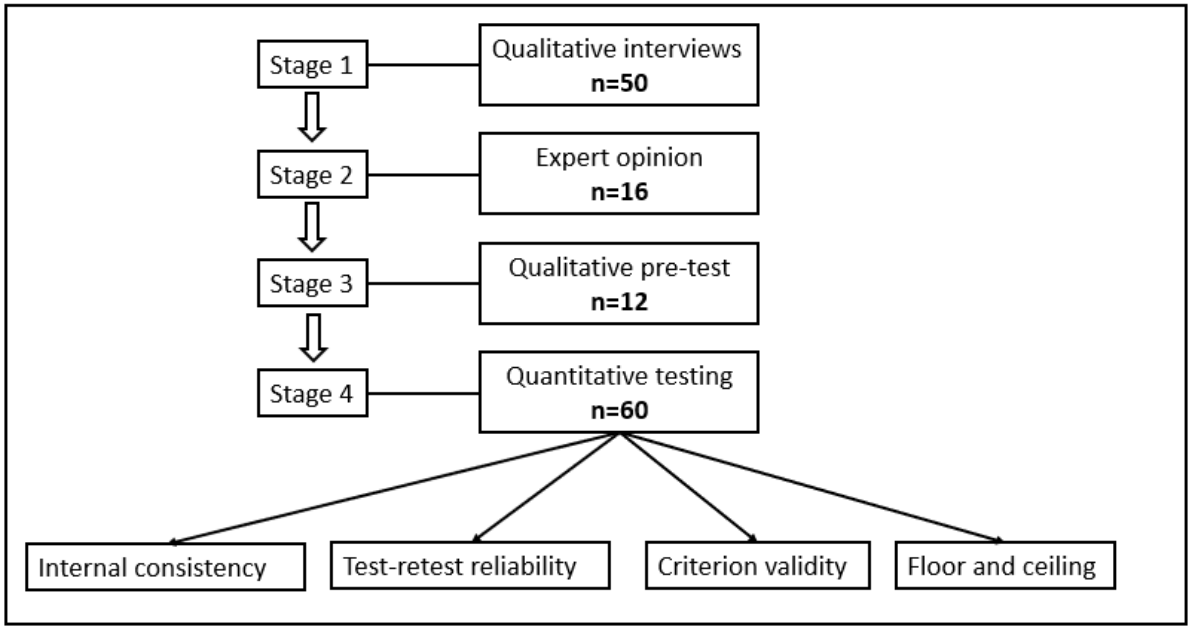


Figure 2: Summary of participant recruitment

2.2 Patient recruitment

The patients were recruited from orthopaedic clinics and wards at Broadgreen Hospital (Royal Liverpool and Broadgreen University Hospitals Trust). The inclusion criteria were adult patients suffering from OA of the hip or knee that were being offered joint replacement surgery. The reason for selecting these patients was that end stage OA would represent a relatively stable state of the disease where there would be limited variability in day to day symptoms. Patients were recruited for the item generation, prototype pretesting and validation stages of the PROM development. Patients that participated were given printed information sheets about the stage of the project (see Appendix F) and written consent was obtained and recorded on consent form (see Appendix G).

2.3 Health professional recruitment

The health professionals consisted of orthopaedic consultants, senior orthopaedic trainees, and physiotherapists with experience in managing patients with OA. This group formed a panel of experts that were recruited for the stage of item selection for the PROM. All participants were employed by the Royal Liverpool and Broadgreen University Hospitals Trust and worked in the same clinics that the patients were recruited from. Information about the project was given and consent was obtained for their participation (see Appendix H). Expert participation was voluntary.

2.4 Ethical approval

Ethical approval was obtained from the Health Research Authority (NHS England) to undertake this project (Appendix I). This project was sponsored by the University of Liverpool. The researchers involved in this project have completed an *Introduction to Good Clinical Practice (GCP) eLearning module* organised by the National Institute for Health Research, Clinical Research Network (Appendix J).

2.4 Development stage 1

2.4.1 Patient Demographics

The inclusion criteria for the participants of this stage of the project were:

1. Patient over the age of 18 years old
2. Has an established diagnosis of OA of at least one hip or knee knee

The following table (Table 2)summarises details of the patient participants recruited for this stage of the project.

Table 2: Summary statistics for participants in stage 1

Summary Statistic	Value
Number recruited	50
Mean age (years) (sd)	66 (11.39)
Male:Female ratio	23:27
BMI (Kg/M^2) (sd)	31 (5.3)

The following table (Table 3) summarises the primary joints affected by OA in this sample of patients. The patients have been listed based on the total number of joints affected by OA regardless to treatment that has be given. Objective information regarding the joints involved was obtained from electronic clinical records of the patients and recorded on the participant data collection sheets.

Table 3: Summary of joints affected by OA

Joints affected	Number of Patients
Both hips and knees	18
Bilateral hip joints	8
Single hip joint	9
Bilateral knee joints	10
Single knee joint	5
Total	50

2.4.2 Qualitative data collection

Semi-structured interviews were conducted by two researchers (HY & AM). An interview guide was used to maximise diversity in patient responses (Appendix K). The minimum number of patients to be recruited for this stage was set at 50. By interviewing this number, it was believed that the majority of aspects pertaining to the effects of OA would be explored and this would ensure questions included in the PROM would satisfy content validity. Patient responses were transcribed onto data collection sheets during the interview (see Appendix L).

2.4.3 Qualitative data analysis

Information collected from each interview was analysed and recorded in Excel (Microsoft). Findings that described a reduction of functional ability or activities that were limited due to patients' symptoms were identified and highlighted as themes. Other concepts that were considered important and were also included were patient symptoms, general health, and social interactions. These themes were then grouped into categories which were then used to create subdomains for the PROM.

2.5 Development stage 2

Initially, the panel of experts were interviewed to identify concepts that they felt suitably measured the level of physical function from a clinician's perspective. This was to confirm that the themes identified were relevant for the clinician and to ascertain if any other concepts were missed from the interviews with the patients.

To rank the themes, the panel of experts completed questionnaires which listed all the themes grouped in their subdomains. The participants were asked to comment on how strongly they agreed or disagreed on the importance of each being used as a concept to assess the subdomain it belonged to. Responses were recorded using a Likert scale which were given numerical values for statistical analysis (value = 1-6) (see Appendix M). Two rounds of questionnaires were conducted. Themes with the low average scores were eliminate in between the first and second rounds. The expert panel were made aware of the results of the first round prior to conducting the second round. The highest scoring themes from the second round were considered for the final version of the questionnaire.

2.6 Development stage 3

2.6.1 Item selection

In creating the PROM from the list of themes the first thing to consider was the length of the questionnaire. To reduce patient burden only 20 questions were included. The selection of items for the final questionnaire was made by the research team as guided by the opinions of the expert panel. All of the five subdomains identified were included.

The number of items for each subdomain were not the same and more questions were included for the subdomains pertaining to physical function. This was done intentionally to allow for a PROM that was weighted towards assessing physical function but was still able to provide a reasonable reflection of the patient's social role limitation, pain symptoms and general health.

With regards to the design of each question it was felt that using a Likert type scale with five responses would be the most appropriate in terms of ease of answering the questions. For each item, the research team assigned a score from zero to four, higher values indicating better levels of functional ability. These values were assigned to aid in statistical analysis and to provide an overall score for the questionnaire.

2.6.2 Body map and VAS

A body map was created by the researchers using Powerpoint (Microsoft) as an original graphic to represent the human body. Boxes were included over the main joints for patient to place numbers in accordance with severity of pain experienced in that joint. The VAS scale was constructed as a horizontal line (10 cm) with equidistant markings at each 1cm to represent 10%. Patients would be asked rate their level of satisfaction with their overall level

of function on this scale. The researcher is then able to measure where the mark in placed in mm to obtain the patients' level of satisfaction as a percentage.

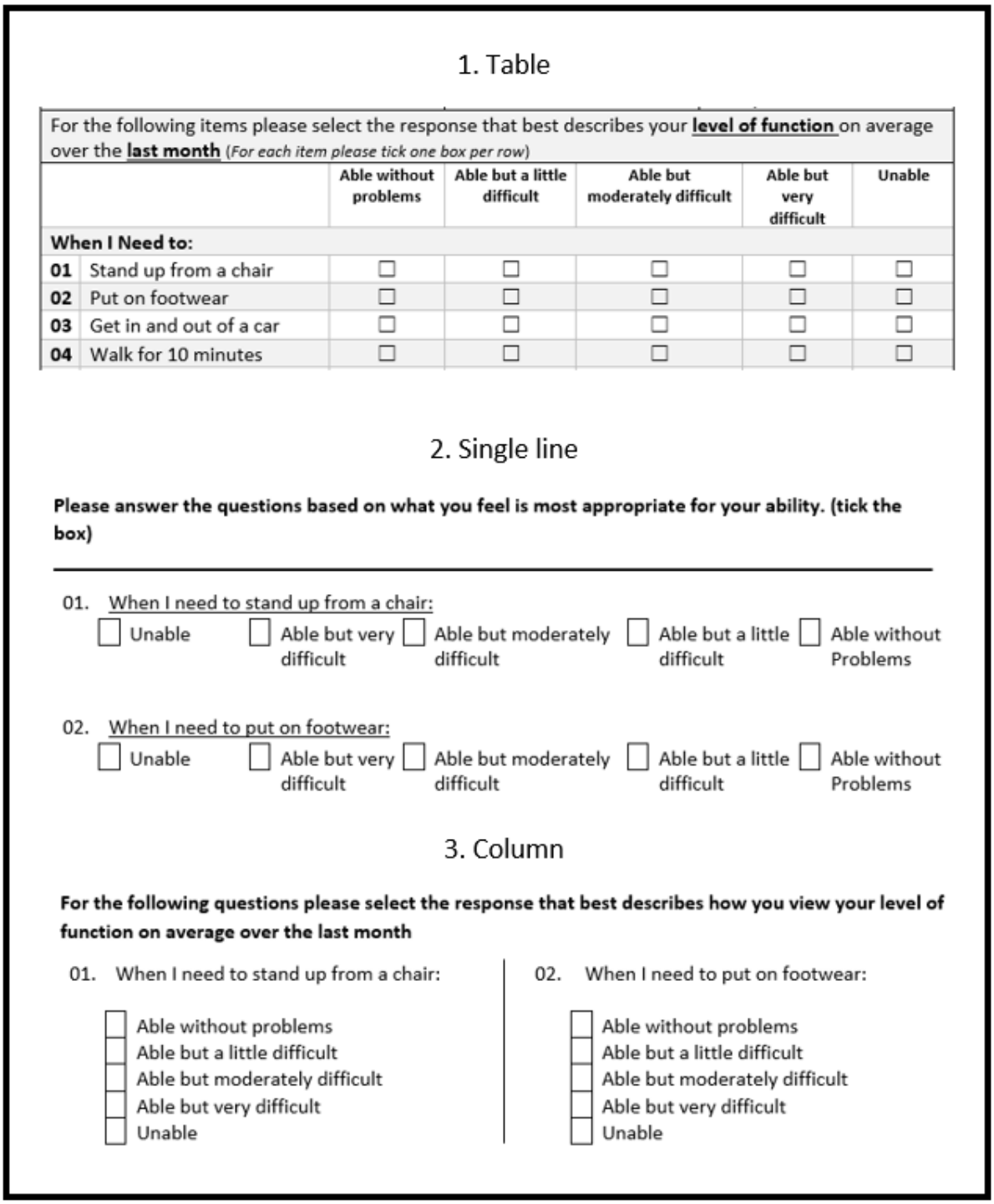
2.6.3 Qualitative pre-test

Table 4: Summary Statistic for participants in Stage 3

Summary Statistic	Value
Number recruited	12
Mean age (years) (sd)	70 (7.15)
Male:Female ratio	4:8

Patients were given the prototype questionnaire to complete, and a short interview was conducted afterwards. Patient were recruited from the Orthopaedic wards on the day there were to have their joint arthroplasty surgery. The follow table summarises the participant demographics. An interview guide (Appendix N) was used to ensure that the research team enquired about the clarity of instructions, how easy it was to complete, and whether there were any questions that were unnecessary or left out from the questionnaire. The aim was to gather opinion on whether the PROM covered all relevant aspects to assess physical function. Patients were also encouraged to comment on anything that was unclear or confusing.

Patients were also presented with three formatting styles for the questionnaire and were asked to choose a preference. The style that received the most votes would be used as the style for the finalised questionnaire. Figure 3 below provides a sample of each formatting style.



1. Table

For the following items please select the response that best describes your **level of function** on average over the **last month** *(For each item please tick one box per row)*

		Able without problems	Able but a little difficult	Able but moderately difficult	Able but very difficult	Unable
When I Need to:						
01	Stand up from a chair	☐	☐	☐	☐	☐
02	Put on footwear	☐	☐	☐	☐	☐
03	Get in and out of a car	☐	☐	☐	☐	☐
04	Walk for 10 minutes	☐	☐	☐	☐	☐

2. Single line

Please answer the questions based on what you feel is most appropriate for your ability. (tick the box)

01. When I need to stand up from a chair:
☐ Unable ☐ Able but very difficult ☐ Able but moderately difficult ☐ Able but a little difficult ☐ Able without Problems

02. When I need to put on footwear:
☐ Unable ☐ Able but very difficult ☐ Able but moderately difficult ☐ Able but a little difficult ☐ Able without Problems

3. Column

For the following questions please select the response that best describes how you view your level of function on average over the last month

01. When I need to stand up from a chair:
☐ Able without problems
☐ Able but a little difficult
☐ Able but moderately difficult
☐ Able but very difficult
☐ Unable

02. When I need to put on footwear:
☐ Able without problems
☐ Able but a little difficult
☐ Able but moderately difficult
☐ Able but very difficult
☐ Unable

Figure 3: Three questionnaire format styles

2.7 Development stage 4

2.7.1 Patient recruitment

Patients were recruited from Orthopaedic clinics at Broadgreen Hospital (Royal Liverpool and Broadgreen University Hospitals Trust) on the day that they were listed for joint arthroplasty surgery. As a routine these patients would be asked to complete PROMs required by the NJR. Recruited patients were asked to complete the novel PROM along side these PROMs. Patients were then given a stamped envelope with another copy of the novel PROM inside, and were instructed to complete and return it after 48 hrs. The following table (Table 5) summarises the patient demographics and the numbers used during each validation step.

Table 5: Summary Statistics for participants in Stage 4

Summary Statistic	**Value**
Number recruited	60
Mean age (years) (sd)	69 (11.05)
Male:Female ratio	30:30
BMI (Kg/M^2) (sd)	30.5 (4.7)
Validation Step	**Number**
Internal Consistency	60
Test-Retest	47
Correlation: OHS	27
Correlation: OKS	33
Correlation: WOMAC	60
Correlation: SF12	60
Floor and Ceiling	60

2.7.2 Validation

Patients recruited into this stage were given the newly developed PROM as well as the Oxford hip or knee score, WOMAC and SF-12 questionnaires to complete in clinic. Patients were then given an envelope containing another copy of the new PROM which was to be completed

again by the patient after 48 hours. The envelopes were already stamped and addressed to the primary research site. Data obtained from the questionnaires were used to test validity.

2.7.3 Internal consistency

Internal consistency of the PROM was measured statistically by Cronbach's Alpha. Cronbach's alpha is widely used to measure internal consistency among items within a questionnaire by investigating the interrelatedness of all the items within a presumed scale of subscale (Taber, 2018). It is measured by splitting the items of a scale into half and calculating the correlation between the two halves, so called split-half coefficient, the Cronbach's alpha is the factor of all split-half coefficients of that scale (Taber, 2018). Cronbach's alpha is presented as a number between 0 and 1, with alphas closer to 1 showing higher internal consistency.

Separate analysis was performed on the first 20 questions and on each subdomain. Calculation of Cronbach's Alpha was also performed on each subdomain with sequential deletion of each item to investigate if this would improve internal consistency. As the body map and VAS were standalone questions internal consistency was not performed.

2.7.4 Test-retest reliability

To assess reliability, interclass correlation coefficient (ICC) was perform between the PROM completed on the day of recruitment and the PROM that was returned to the researchers after 48 hours. ICC was calculated for each item individually, for the total scores of each subdomain and also for the total score of the first 20 questions. The two-way mixed effects model using absolute agreement was used to produce the ICC.

2.7.5 Criterion validity

To test for criterion validity of the new PROM the total scores derived from each subdomain were correlated against the Oxford hip or knee score, WOMAC and SF12. Where possible, convergent and divergent validity was assessed by comparing similar and unrelated subdomains respectively.

Descriptive statistics were performed individually for each of the PROMs to assess the pattern of distribution of the data set. As the total scores of the subdomains created a continuous scale, statistically they could be compared to the scales produced from the OHS, OKS and subscales of WOMAC. If the data was found to be parametric then Pearson's correlation coefficient was used for correlation analysis. For non-parametric data, Spearman's rank correlations were obtained. For the body map Spearman's rank correlation was used. Lastly, Pearson's correlation coefficient was used to compare the VAS question.

2.7.4.1 Oxford hip and knee score comparison

The recruited patients with either knee or hip OA were divided into two groups depending on the affected joint. Analysis of the two groups was undertaken separately. All subdomains from the new PROM were compared to the respective Oxford score. This was done to allow testing for convergence and divergence of themes. It was decided that Spearman's rank correlation coefficient would be used because of the data distribution.

2.7.4.2 WOMAC comparison

Only the pain and physical function subdomains of WOMAC were used for comparison. The pain subdomains from the new PROM and WOMAC were compared. For the physical function subdomain of WOMAC, correlation was made with lower limb, upper limb function and role limitations subdomains.

As the body map did not have a direct comparison in other PROMs, validating this question was done by treating it as a surrogate measure for pain. Thus the body map numerical score was correlated against the WOMAC pain subdomain.

2.7.4.3 SF12 comparison

The physical function composite (PCS) score was correlated against the lower limb function, upper limb function and role limitations subdomains of the new PROM. The mental health composite score was correlated against the general health subdomain as this assesses the effects of patients' mood and general health on their physical function.

As VAS score reflected the level of satisfaction of patients with their level of function and it was felt that the mental health composite score of the SF12 was best suited to assess criterion validity for this question.

2.7.6 Floor and ceiling effects

To test for floor and ceiling effects the total scores of each subscale and the total score of the first 20 questions were analysed for frequency of the minimum and maximum possible scores using the results from PROM completed by the patients on the day of recruitment. Floor and ceiling effects were not assessed for the body map or VAS questions.

Chapter 3: Results

3.1 Development stage 1

3.1.1 Qualitative data collection

The number of patients recruited to this stage was 50. The mean (±sd) age was 66 (±11.39), of which 27 (54%) were female. All patients were interviewed by the research team and a data collection sheet was used to record relevant information including the primary joint affected by OA, previous treatment(s), and other medical problems. An interview guide was used to prompt the researcher with open ended questions to ascertain aspects of daily function that were affected by the patient's condition. These ideas and concepts were transcribed onto the data collection sheets.

3.1.2 Qualitative data analysis

A total of 86 unique themes were identified. The frequency of each unique element was determined. Themes that were mentioned included physical activities patients found difficult, social activities and symptoms related to OA or other medical problems. The most frequent theme that was mentioned was 'walking' (82%). This was followed by 'using stairs' (48%) and 'getting in and out of cars' (38%). A list of the twenty most frequent themes identified are listed in Appendix O.

Each theme was then evaluated and converted into a single word or short phrase so that similar themes could be grouped together. For example, 'difficulty in getting in or out of a car' and 'difficulty getting on or off a bus' were rephrased as 'difficulty entering or exiting a vehicle'. The analysis and pooling of the items into common themes gave rise to five

subdomains: lower limb function; upper limb function; role limitation; pain; general health (Figure 4).

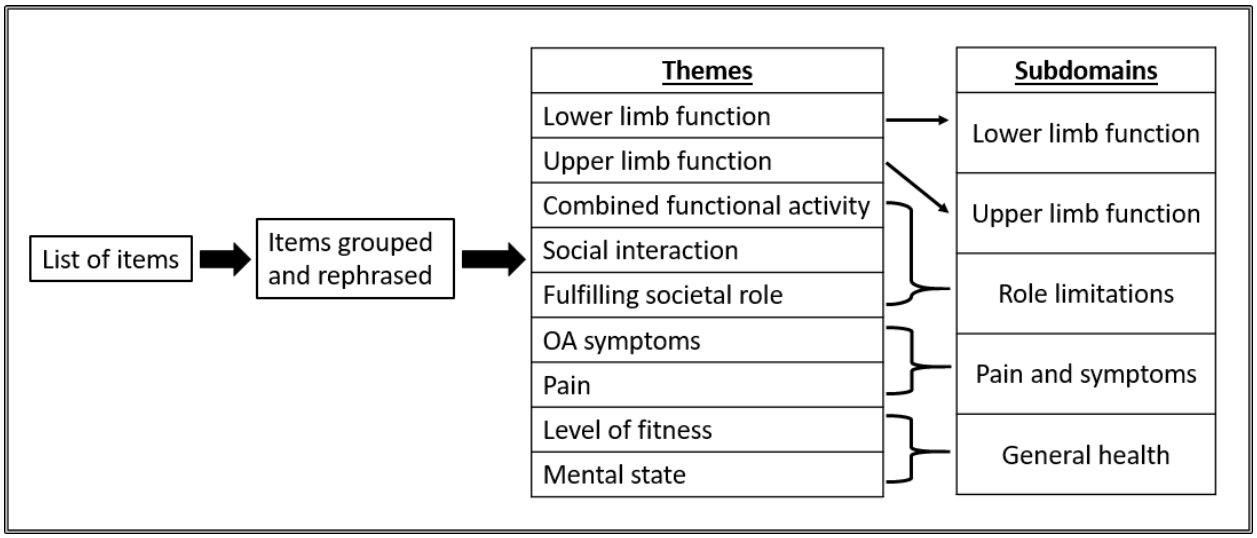


Figure 4: Grouping of themes into subdomains

3.2 Development stage 2

3.2.1 Expert participants

Sixteen expert participants were recruited for this stage of development. The panel consisted of orthopaedic consultants, clinical fellows, senior registrars and senior physiotherapists (Table 3).

Table 6: Breakdown of expert panel

Expert participant		Years of experience
Position	Number	Mean (SD)
Orthopaedic Consultant	6	21.83 (7.68)
Senior clinical fellow	3	9.33 (1.15)
Senior registrar	2	7.50 (3.54)
Physiotherapist	5	22.60 (8.62)
Total	16	17.94 (9.12)

3.2.2 Qualitative interviews with expert panel

Eight new themes were identified and were added to the existing pool of potential items that would be included in the final questionnaire (Table 4).

Table 7: List of new items added

New item description	**Subdomain**
Maximum walking distance Squatting Standing for a period of time Standing on tip toes Walking on uneven surfaces Gait abnormality Walking stability Walking up and down hills	Lower limb function
Stopping activity because of pain Worse pain versus usual pain	Pain and symptoms

3.2.3 Expert participant opinion

For the first round of questionnaires the response rate was 87.5%. Items which had a mean score greater than 5 where thought to be highly desirable for inclusion in the final questionnaire. If more than 70% of the participant responses were within the 95% confidence interval of the scores then it was decided that there was consensus for that item response. Appendix P contains completed statistics of this stage. Items with a mean score of less than 4 were thought to be undesirable and were removed from the pool. Only one item, 'standing on tip toes', was removed after the first round. A summary of the results was given to the expert panel prior to commencing the second round.

The response rate for the second round was 62.5%. All items that were desirable from the first round maintained a mean score above 5, and there was still consensus among the remaining participants for each of these items. Given that there was stability in the responses between the first and second rounds a further round was felt to be unnecessary. Following this process, the items that were identified as being desirable were added to the item bank to be selected for the final questionnaire. The number of items in the bank was 70.

3.3 Development stage 3

3.3.1 Questionnaire outline

The final version of the questionnaire included 22 questions: 20 questions derived from the item bank using Likert type responses, the body map and the VAS. The following table details how many questions were allocated to each subdomain. (Table 5)

Table 8: Number of questions included for each subdomain

Subdomain	**Number of questions**
Lower limb function	6
Upper limb function	5
Role limitation	4
Pain and symptoms	3
General health	2

3.3.2 Question design

3.3.2.1 Lower limb, upper limb and role limitation

The lower limb function, upper limb function and role limitation subdomains where grouped together as the same Likert scale could be used. This section of the questionnaire started with instructions for the patient asking them to select a single response that best describes their average level of function. The recall period was for one month.

The five-point scale was scored from 0 to 4. The scale consisted of options that would indicate no difficulty in performing an activity (score 4) at one end and unable to perform the activity (score 0) at the other, increments in between these extremes would increase in level of difficulty performing the activity. This was then followed by a prompt *'when I need to'* before listing included items from the three subdomains. The following figure demonstrates how the prompt and scale are laid out for the new PROM.

For the following items please select the response that best describes your **level of function** on average over the **last month** (*For each item please tick one box per row*)					
	Able without problems	**Able but a little difficult**	**Able but moderately difficult**	**Able but very difficult**	**Unable**
When I Need to:					

Figure 5: Instructions and scale used for first 15 questions in PROM

3.3.2.2 Pain and symptoms

The questions were split into two sections. The first section instructed patients to select which response best describes pain at rest. Here the five-point scale was used to indicate increasing levels of pain from no pain at rest (score 4) to constant pain (score 0). The second section asked how pain would affect the patient's overall function when trying to use either their arms or legs with a scale that indicated no limitation (score 4) to complete limitation (score 0). The next figure shows how the pain symptom questions are laid out in the PROM (Figure 6).

Please select the response that best describes your **level of pain while resting** (on average last month)						
		No pain	**Little pain**	**Moderate pain**	**Severe pain**	**Constant pain**
16	Level of pain	☐	☐	☐	☐	☐
Please select how **pain** from your joints limits your **overall function**						
		No Limitation	**Little limitation**	**Moderate limitation**	**A lot of limitation**	**Completely limited**
When I Need to:						
17	Use my legs	☐	☐	☐	☐	☐
18	Use my arms	☐	☐	☐	☐	☐

Figure 6: Pain subdomain questions

3.3.2.3General health

The general health subdomain used the same scale as in the second section of the pain subdomain however two individual questions were asked. Examples were included in the final version of the PROM to provide more clarity to the questions (Figure 7).

		No Limitation	Little limitation	Moderate limitation	A lot of limitation	Completely limited
19	How does your **general medical health (e.g.** asthma, **heart disease)** limit your overall function?	☐	☐	☐	☐	☐
20	How does your **mood (e.g. anxiety, depression)** limit your overall function?	☐	☐	☐	☐	☐

Figure 7: General health questions

3.3.2.4 Body map

The body map consisted of a figure of a human with boxes overlying the main joints of the body (Figure 8). Instructions were given to the patient to fill in the boxes with numbers (0-10) to indicate the amount of pain they experience from that particular joint.

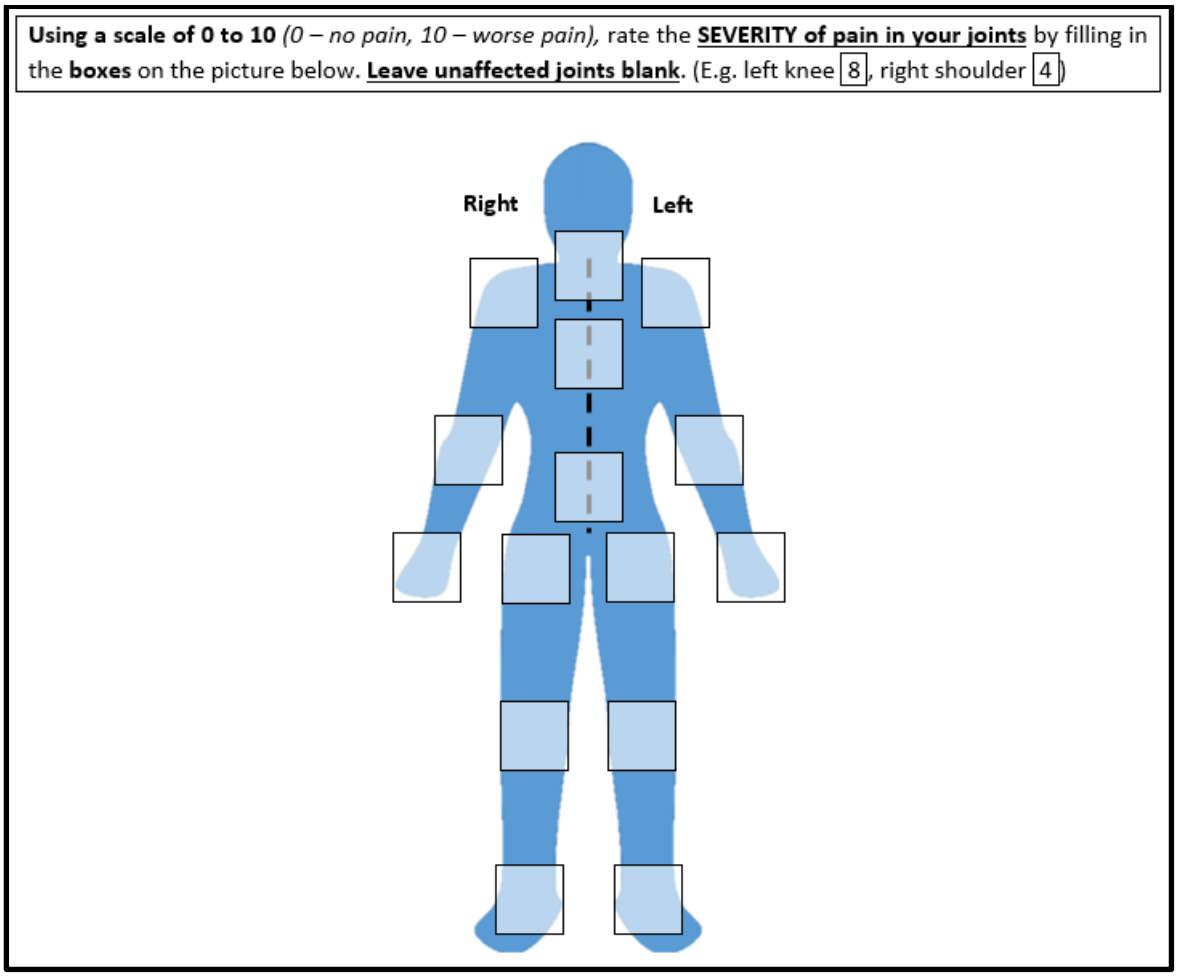


Figure 8: Body map question

3.3.2.5 VAS

This question instructed the patient to mark along the scale their level of satisfaction with their overall level of function.

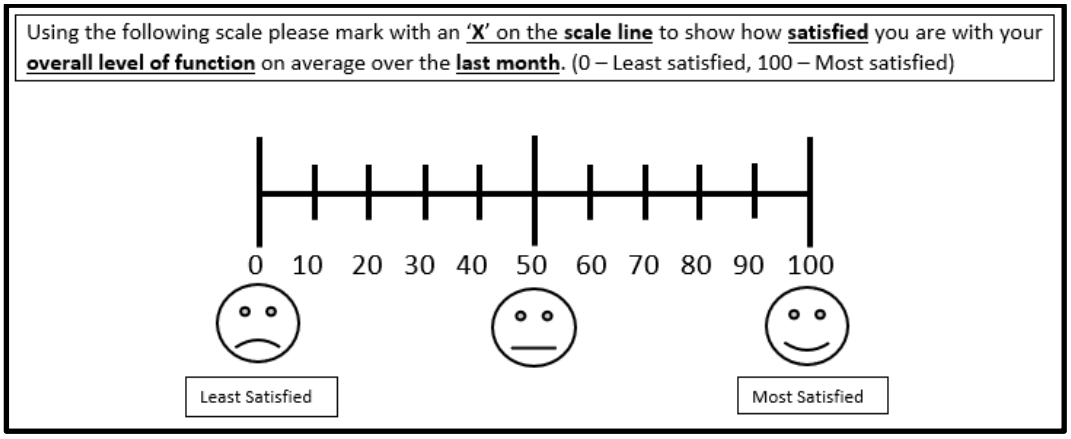


Figure 9: VAS question

3.3.3 Item inclusion

3.3.3.1 Lower limb function

The items included in this subdomain were activities which would gauge a patient's level of lower limb function. Because of the format of the questionnaire the activities only needed to be listed as per the following table. (Table 6)

Table 9: Included items in lower limb function subdomain

Activity to be performed by the patient
1. Stand up from a chair
2. Put on footwear
3. Get in and out of a car
4. Walk for 10 minutes
5. Go up a flight of stairs
6. Go down a flight of stairs

3.3.3.2 Upper limb function

Like the lower limb function subdomain, the activities listed would gauge the patient's upper limb function and are listed in the following table. (Table 7)

Table 10: Included items in the upper limb function subdomain

Activity to be performed by the patient
1. Carry things (e.g. shopping bag)
2. Do up buttons
3. Reach for something at shoulder height
4. Turn a key
5. Prepare a meal

3.3.3.3 Role limitations

This subdomain included activities that considers a person's interaction in society and activities that require whole body function to complete. Table 8 lists the activities that were included in the questionnaire.

Table 11: Included items in role limitations subdomain

Activity to be performed by the patient
1. Do my regular job or daily routine if retired
2. Perform leisure or sporting activities
3. Do housework
4. Go shopping on my own

3.3.3.4 Pain and symptoms

The pain and symptoms subdomain first elicit the patient's level of pain at rest. This is followed by two questions that asks how pain limits the patient's overall function when trying to use their legs and when they use their arms.

3.3.3.5 General health

The general health subdomain uses the two remaining questions to assess whether a patient's overall level of function is affected by other factors aside from their joint disease. The general health subdomains aim to elicit whether a patient's general medical health or their mood cause limitations to their overall level of function.

3.3.4 Qualitative pre-test

For pre-testing, twelve patients were recruited. The mean (±sd) age was 68.92 (±7.15) years, with seven male and five female patients. Nine of the patients were about to undergo a total knee replacement and the remaining three were due to have a total hip replacement.

Eleven (91.6%) of the patients commented that the PROM was easy to fill in and that assistance was not required. The one patient who needed help was assisted by her husband, and this was due to her having poor vision. Eight of the patients commented that the questions in the PROM were relevant. Only two questions were felt to be missing and these were 'getting in and out of a bath' and 'standing for long periods of time'. Other general comments regarding the clarity of the questionnaire included the use of examples for medical conditions and mood (general health subdomain) and rewording of some of the instruction to make them clearer.

With regards to general format of the questionnaire three format styles were presented to the patient. Patients were asked to comment on which format was most aesthetically pleasing. The results have been tabulated in the following table. (Table 9)

Table 12: Popularity of each format style

Format style	Number of votes
Table	7
Single line	2
Column	1
Undecided	2

The most popular style (table style) was selected as the format for the finalised questionnaire. Given that the majority of the patients felt that the questions included were all relevant none of the items were changed and only minor adjustments were made to for the sake of clarity. Appendix Q contains the final version of the PROM.

3.4 Development stage 4

3.4.1 Patient recruitment

The number of patients recruited for this stage was 60, mean (±sd) age was 69.03 (±11.05), 30 female and 30 males.

3.4.2 New PROM validation

Statistical analysis for internal consistency was completed using data from the new PROM completed on day of recruitment. Test-retest reliability was completed by using complete PROM on recruitment day and returned PROMS. Analysis for criterion validity was completed using data from questionnaires completed by patients on day of recruitment. Missing data has been acknowledged and addressed in the relevant sections where necessary.

3.4.3 Internal consistency

Results from the new PROM completed on day of recruitment were used (n=60). The Cronbach's alpha, item mean (±sd) and mean (±sd) total score (scale) for each subdomain are displayed in the following table (Table 10). A Cronbach's alpha value greater than '0.70' is desirable as it reflects good internal consistency.

Table 13: Summary of Cronbach's Alpha and means for first 20 questions in new PROM

New PROM	**Cronbach's Alpha**	**Item mean (sd)**	**Scale mean (sd)**
Question 1 – 20	0.93	2.23(0.84)	44.55(13.61)
Subdomain	**Cronbach's Alpha**	**Item mean (sd)**	**Scale mean (sd)**
Lower limb function	0.81	1.73(0.20)	10.35(4.25)
Upper limb function	0.87	3.08(0.65)	15.40(4.23)
Role limitation	0.81	1.61(0.47)	6.43(3.68)
Pain and symptoms	0.52	2.07(1.15)	6.22(2.01)
General health	0.74	3.08(0.22)	6.15(1.94)

Cronbach's Alpha for the first 20 questions suggests high internal consistency within the new PROM. Analysis of the subdomains also showed good internal consistency apart from the pain and symptoms subdomains. The following tables demonstrate Cronbach's Alpha for each subdomain, as well as the Cronbach's Alpha after sequential deletion of items from the related subdomain. (Table 11 - 14).

Table 14: Cronbach's Alpha for Lower limb function subdomain if item deleted

PROM question number	**Cronbach's Alpha if item deleted**
Question 1	0.76
Question 2	0.80
Question 3	0.76
Question 4	0.80
Question 5	0.77
Question 6	0.79
Cronbach's Alpha for Lower limb function subdomain = 0.81	

Table 15: Cronbach's Alpha for Upper limb function subdomain if item deleted

PROM question number	Cronbach's Alpha if item deleted
Question 7	**0.89**
Question 8	0.81
Question 9	0.82
Question 10	0.84
Question 11	0.84
Cronbach's Alpha for Lower limb function subdomain = 0.87	

Table 16: Cronbach's Alpha for Role limitation subdomain if item deleted

PROM question number	Cronbach's Alpha if item deleted
Question 12	0.77
Question 13	0.81
Question 14	0.71
Question 15	0.77
Cronbach's Alpha for Lower limb function subdomain = 0.81	

Table 17: Cronbach's Alpha for Pain and symptoms subdomain if item deleted

PROM question number	Cronbach's Alpha if item deleted
Question 16	0.55
Question 17	0.20
Question 18	0.53
Cronbach's Alpha for Lower limb function subdomain = 0.52	

Item deletion was not performed for the general health subdomain as there were only two items in this subdomain.

3.4.4 Test-retest reliability

Thirteen patients did not return the second questionnaire after 48hr giving a response rate of 78.33%. This missing data was handled statistically with listwise deletion (Roth, 1994) and therefore the sample size was reduced to 47. The following table lists the ICC for each item in the questionnaire. (Table 15)

Table 18: Interclass correlation coefficient for each question in new PROM

Question/Subdomain	ICC (N=47)	95% Confidence Interval
(Lower limb function)		
Question 1	0.73	0.52 – 0.85
Question 2	0.86	0.75 – 0.92
Question 3	0.84	0.72 – 0.91
Question 4	0.83	0.69 – 0.90
Question 5	0.77	0.60 – 0.87
Question 6	0.73	0.51 – 0.85
Subdomain scale	0.88	0.78 – 0.92
(Upper limb function)		
Question 7	0.80	0.64 – 0.89
Question 8	0.78	0.60 – 0.88
Question 9	0.85	0.73 – 0.92
Question 10	0.72	0.49 – 0.84
Question 11	0.82	0.68 – 0.90
Subdomain scale	0.88	0.79 – 0.94
(Role limitation)		
Question 12	0.74	0.54 – 0.86
Question 13	0.78	0.60 – 0.88
Question 14	0.82	0.67 – 0.90
Question 15	0.88	0.79 – 0.94
Subdomain scale	0.87	0.76 – 0.93
(Pain and symptoms)		
Question 16	0.80	0.64 – 0.89
Question 17	0.71	0.48 – 0.84
Question 18	0.61	0.31 – 0.78
Subdomain scale	0.80	0.64 – 0.89
(General health)		
Question 19	0.82	0.68 – 0.90
Question 20	0.88	0.78 – 0.93
Subdomain scale	0.88	0.78 – 0.93
Question 1 – 20 scale	0.92	0.86 – 0.96
Visual analogue scale	0.89	0.81 – 0.94

3.4.5 Criterion validity

3.4.5.1 New Prom data distribution

Plotting histograms of each scale subdomain obtained from the new PROM (n=60) suggested that the data obtained was not normally distributed (Figure 10). As a numerical score was assigned to each ordinal response in the first 20 questions a continuous integer scale can be obtained, however as the responses are based on ordinal data and the histograms displayed skewness non-parametric tests were used to calculate correlations.

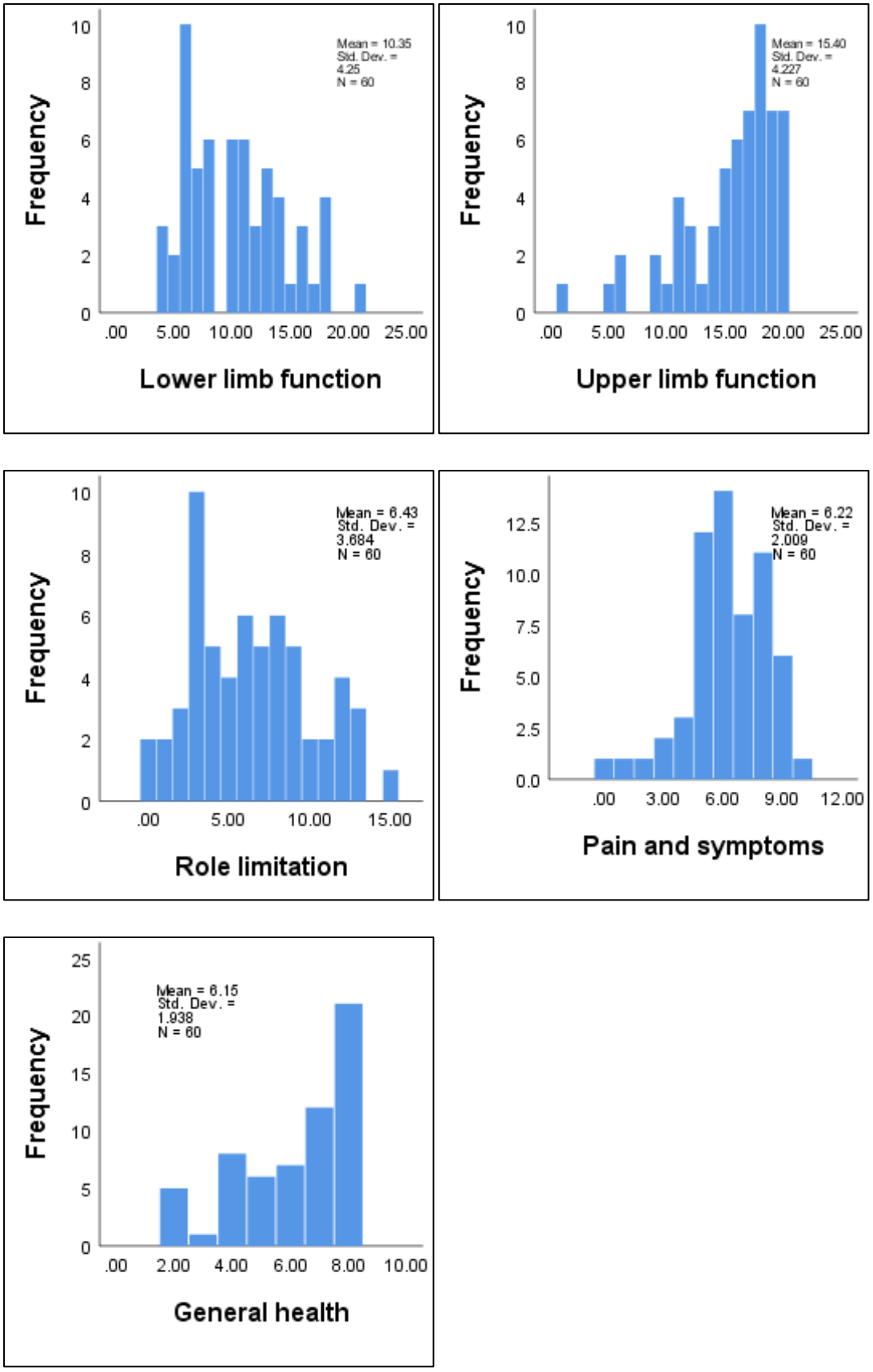


Figure 10: Histograms for the five subdomains of the new PROM

3.4.5.2 Oxford hip score comparison

Twenty-seven patients were assigned to this group. The following figures demonstrate scatter plots of the OHS against the five subdomains of the new PROM (Figure 11). Line of best fit was included along with 95% confidence intervals. The Spearman's rank correlation coefficients are listed in the following table including significance level. (Table 16)

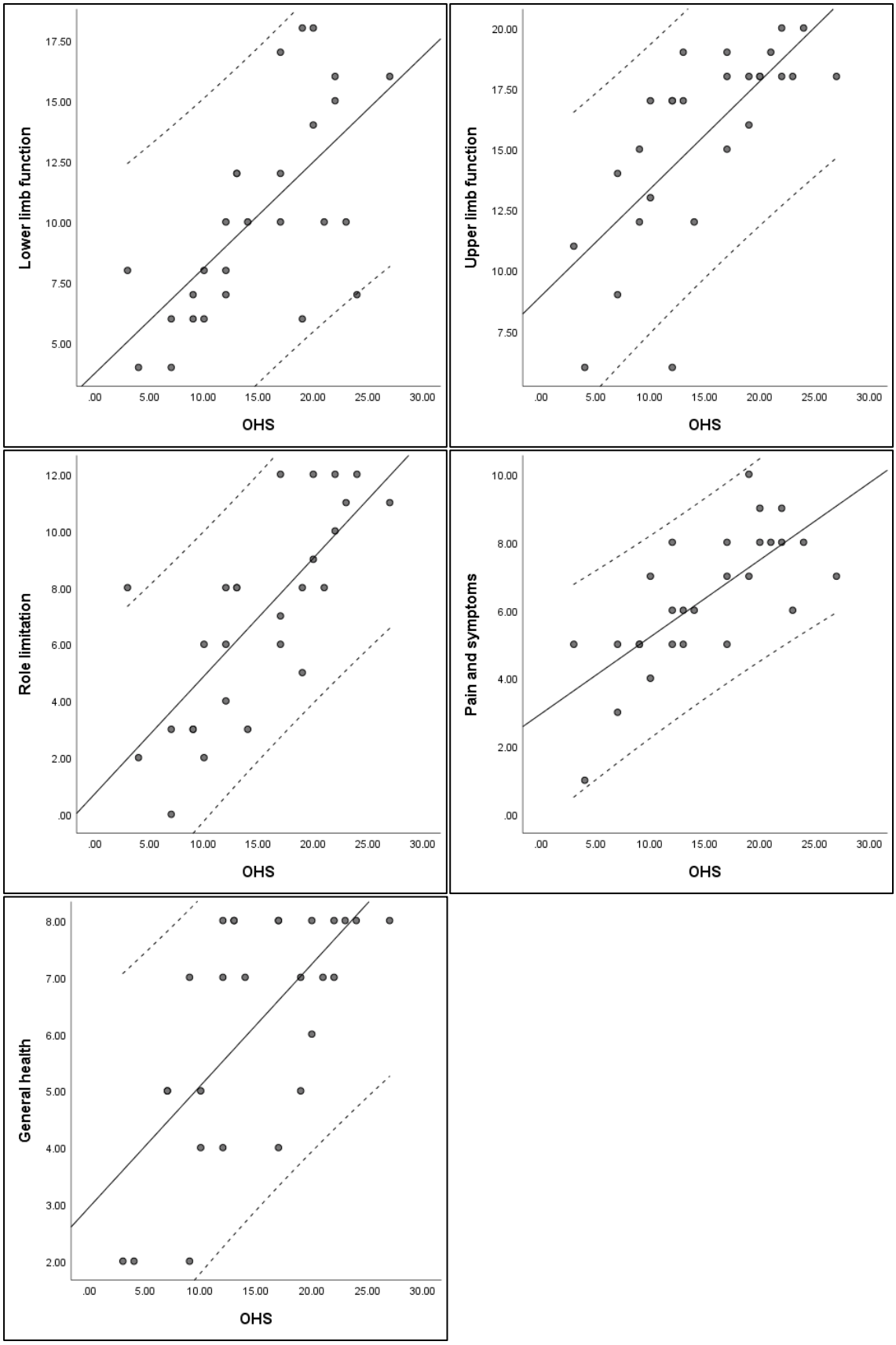


Figure 11: Scatter plots for OHS against PROM subdomains

Table 19: Summary of Spearman's Rank Correlation for OHS and new PROM

Subdomain	Correlation Coefficient	Significance Level
Lower limb function	0.65	<0.05
Upper limb function	0.79	<0.05
Role limitation	0.77	<0.05
Pain and symptoms	0.74	<0.05
General health	0.64	<0.05

Upper limb function and role limitation subdomains of the new PROM demonstrated strong correlation with the OHS. Lower limb function, pain and symptoms and general health subdomains showed moderate correlation with the OHS.

3.4.5.3 Oxford knee score comparison

Thirty-three patients were assigned to this group. Similar to the OHS, the following figures demonstrate scatter plots of the OHS against the five subdomains of the new PROM (Figure 12). Line of best fit was included along with 95% confidence intervals. The Spearman's rank correlation coefficients are list in the following table including significance level. (Table 17).

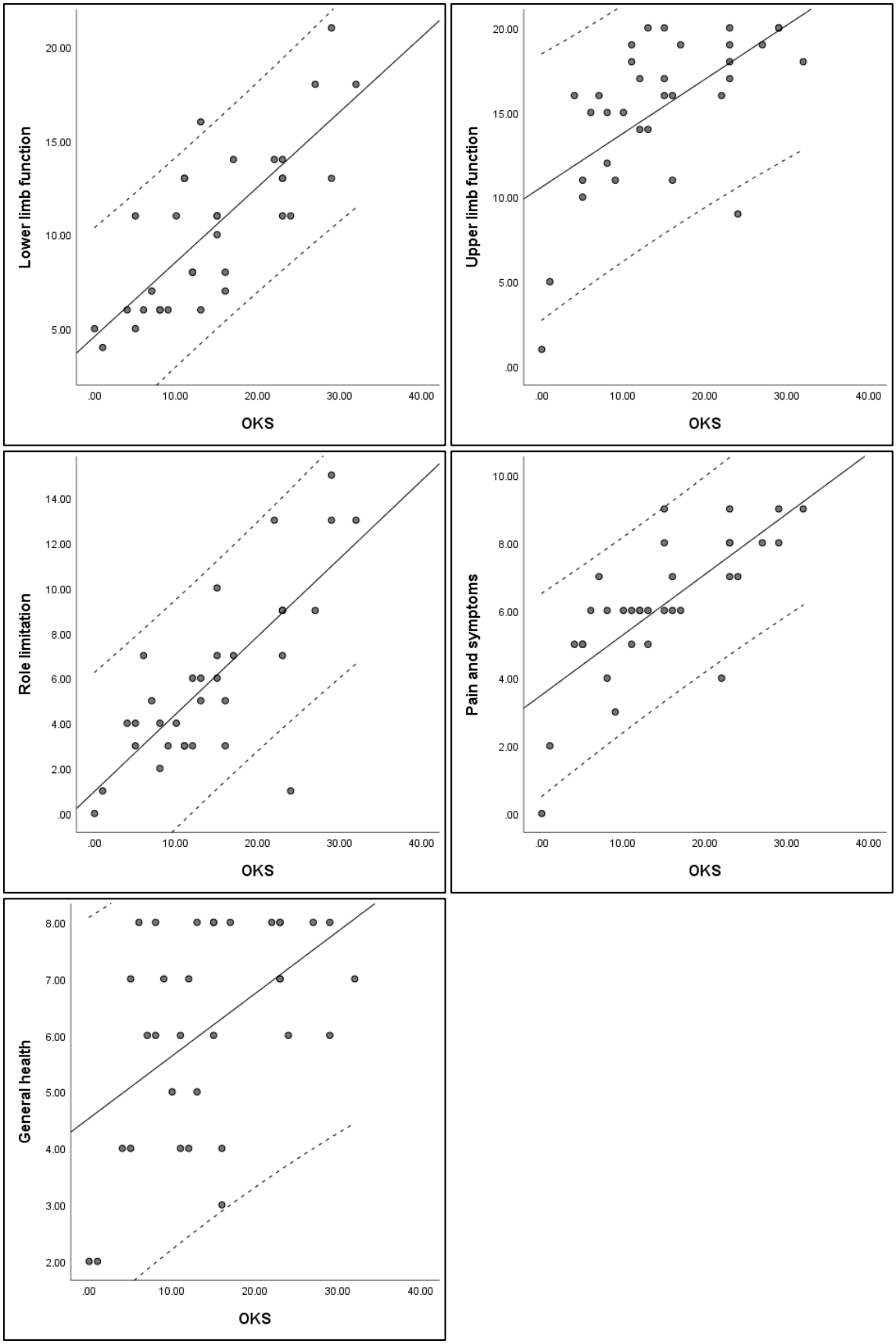


Figure 12: Scatter plots for OKS against PROM subdomains

Table 20: Summary of Spearman's Rank Correlation for OKS and new PROM

Subdomain	Correlation Coefficient	Significance Level
Lower limb function	0.77	<0.05
Upper limb function	0.61	<0.05
Role limitation	0.70	<0.05
Pain and symptoms	0.75	<0.05
General health	0.44	<0.05

Lower limb function and pain and symptoms subdomains of the new PROM demonstrated strong correlation with the OKS. Upper limb function and role limitation subdomains showed moderate correlation with the OKS. Only weak correlation was demonstrated from the general health subdomain which could suggest divergence and therefore confirm divergent validity.

3.4.5.4 WOMAC comparison

Pain subdomain

Sixty patients were included in analysis of the pain subdomains. The Spearman's rank correlation coefficient was -0.72 (significant at 0.01 level). This shows moderate correlation however the value is negative as the polarity of scoring is opposite compared to the new PROM (Figure 13). The other subdomains of the new PROM showed less correlation suggesting a divergent relationship. (Table 18).

When comparing the body map pain scores against the WOMAC pain score the Spearman's rank correlation was 0.48 (significant at 0.01 level).

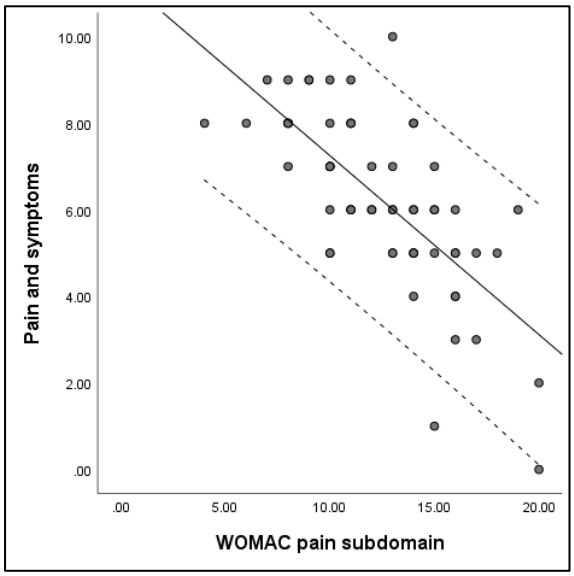


Figure 13: Scatter plot for WOMAC and new PROM pain subdomains

Table 21: Summary of Spearman's Rank Correlation for pain subdomain and new PROM

Subdomain	Correlation Coefficient	Significance Level
Lower limb function	-0.55	< 0.05
Upper limb function	-0.56	< 0.05
Role limitation	-0.54	< 0.05
Pain and symptoms	**-0.72**	< 0.05
General health	-0.40	< 0.05
Body map	0.48	< 0.05

Physical function subdomain

Three patients were excluded from the analysis (n=57) of the physical function subdomain due to incomplete data from WOMAC physical function subdomain. Moderate correlations were observed for all the subdomains of the new PROM against WOMAC physical function subdomain (Table 19). Scatter plots of this relationship are demonstrated in the following figure (Figure 14) .

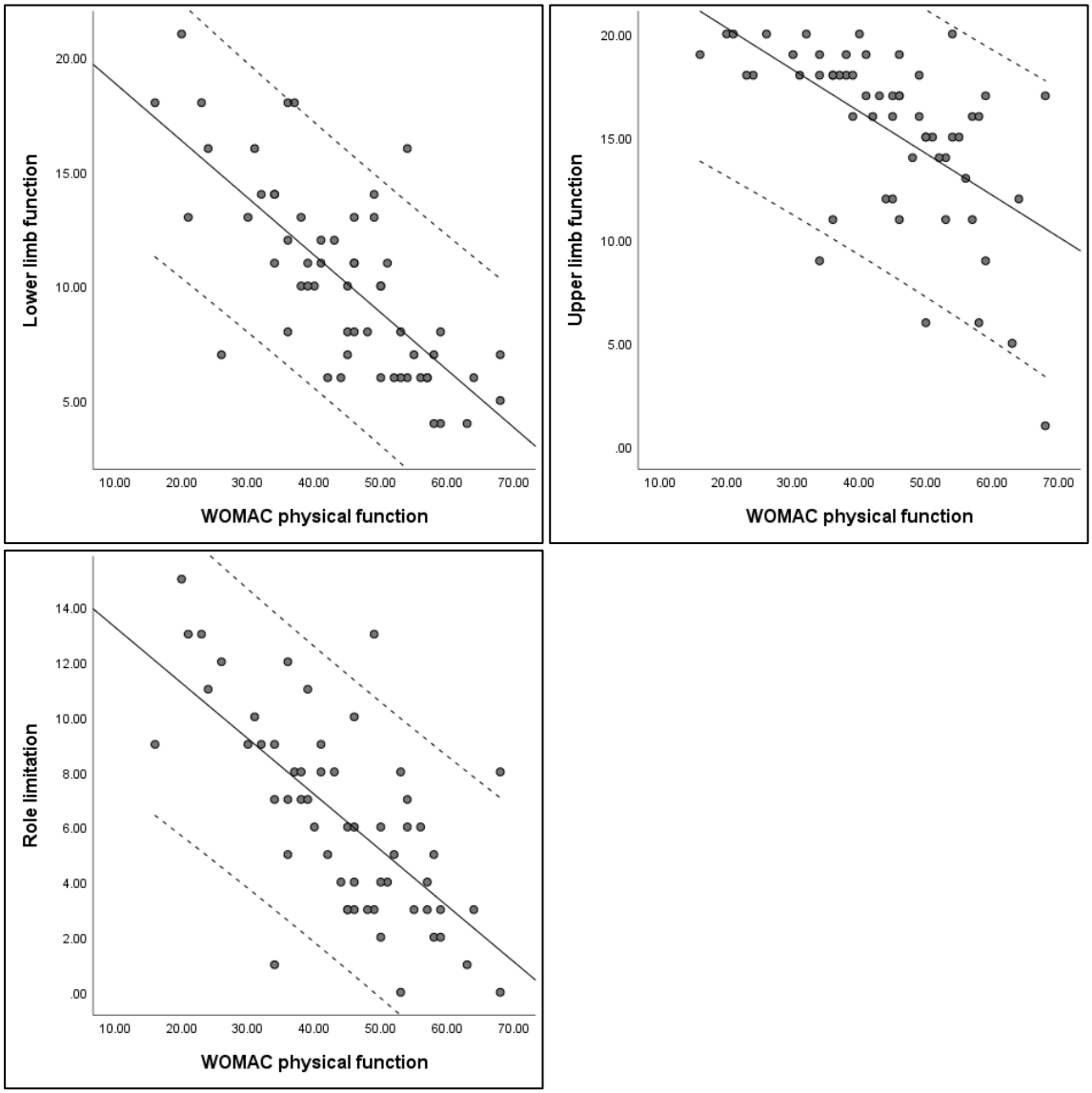


Figure 14: Scatter plots for WOMAC physical function against new PROM

Table 22: Summary of Spearman's Rank Correlation for physical function subdomain

Subdomain	Correlation Coefficient	Significance Level
Lower limb function	-0.73	< 0.05
Upper limb function	-0.65	< 0.05
Role limitation	-0.67	< 0.05
Pain and symptoms	-0.73	< 0.05
General health	-0.52	< 0.05

3.4.5.5 SF12 Comparison

Physical composite summary score

Sixty patients were included in the comparison analysis for the PCS of the SF12. Spearman's rank correlation coefficients did not show acceptable correlation between the PCS and the new PROM subdomains (Table 20). Scatter plots of this relationship with included line of best fit are demonstrated in Figure 15

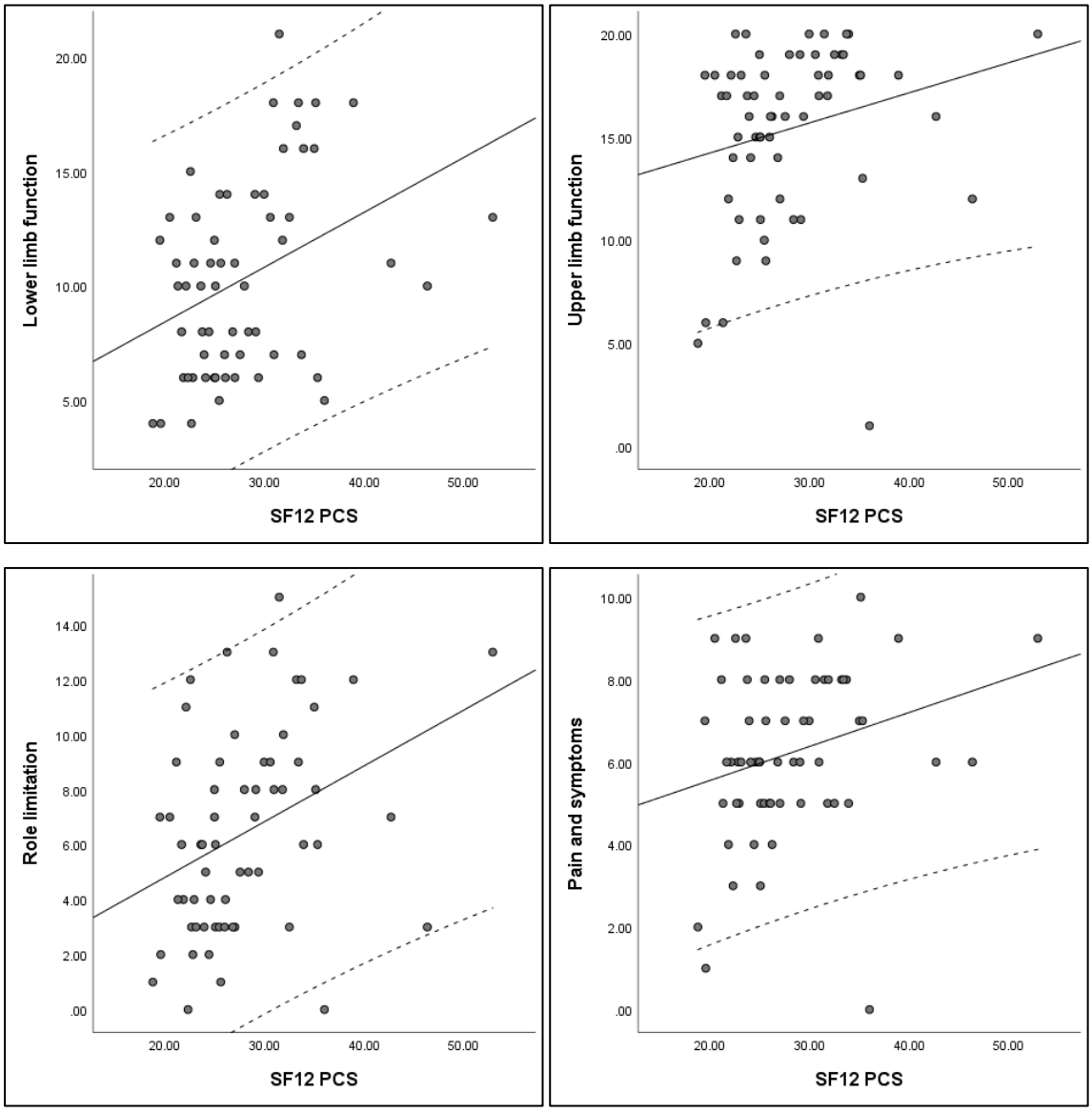


Figure 15: Scatter plots for SF12 PCS against new PROM

Table 23: Summary of Pearson's Correlation for SF12 PCS and new PROM

Subdomain	Correlation Coefficient	Significance Level
Lower limb function	0.38	0.003
Upper limb function	0.30	0.018
Role limitation	0.37	0.003
Pain and symptoms	0.27	0.038
General health	0.06	**0.670**

Mental Component Score

Sixty patients were included in the comparison analysis for the MCS of the SF12. The Spearman's rank correlation coefficients showed moderate correlation for the role limitation and general health subdomains, as well as the VAS. There was weak correlation for the lower limb function, upper limb function and pain and symptoms subdomains (Figure 16, Table 21).

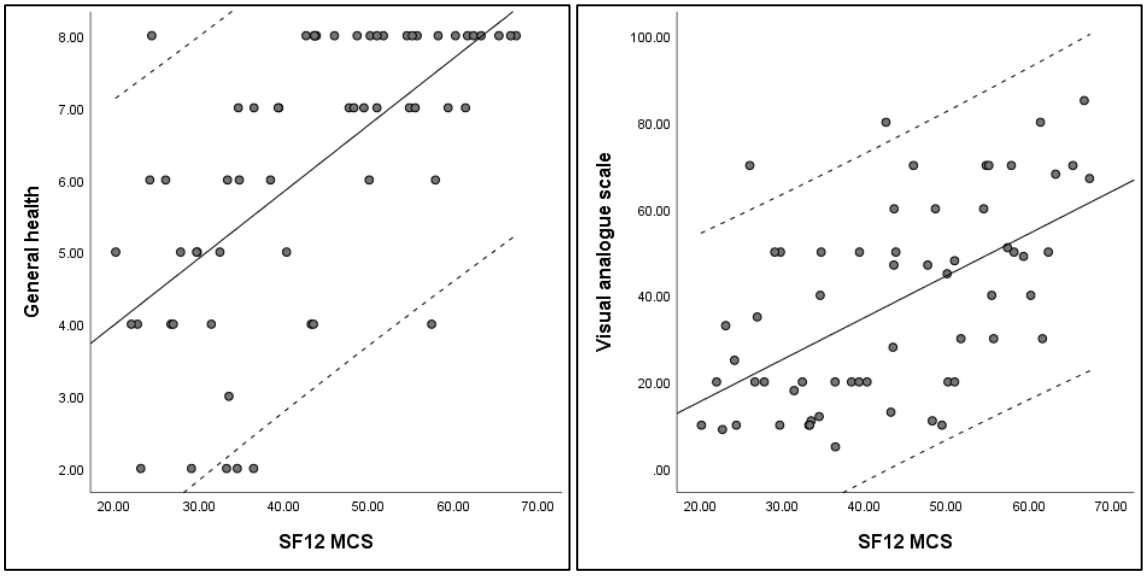


Figure 16: Scatter plots for SF12 MCS against new PROM

Table 24: Summary of Spearman's Correlation for SF12 MCS and new PROM

Subdomain	Correlation Coefficient	Significance Level
Lower limb function	0.45	< 0.05
Upper limb function	0.46	< 0.05
Role limitation	0.58	< 0.05
Pain and symptoms	0.47	< 0.05
General health	0.67	< 0.05
VAS	0.56	<0.05

3.4.6 Ceiling and floor effects

The sample size was 60. With the threshold frequency being set at 15%, nine patients would need to enter either the minimum or maximum score for the PROM to exhibit flooring or ceiling effects respectively. Only the general health subdomain demonstrated a ceiling effect with 21 patients entering the maximum score (35%). When considering the total scores of the first 20 questions no floor or ceiling effects were demonstrated. The follow table (Table 22) summarises these findings.

Table 25: Summary of frequencies for maximum and minimum scores

Subdomain	**Range**	**Frequencies**		**Effects**
		Minimum (%)	**Maximum (%)**	
Lower limb function	0-24	0 (0)	0 (0)	No effects
Upper limb function	0-20	0 (0)	7 (11.67)	No effects
Role limitation	0-16	2 (3.33)	0 (0)	No effects
Pain and symptoms	0-12	1 (1.67)	0 (0)	No effects
General health	0-8	0 (0)	21 (35)	Ceiling effect
Total scores	0-100	0 (0)	0 (0)	No effects

Chapter 4: Discussion

4.1 Discussion of research methodology

When considering the project in its entirety both qualitative and quantitative strategies have been utilised. This is in line with a mixed method of approach with regards to the study methodology. In view of the nature of designing a PROM neither method could have been used in isolation to arrive at the end point required. An intimate understanding of the effects of OA on a patient could only be explored by analysis of interviews with the recruited patients. And validating the new PROM could only be done by employing the previously mentions statistical methods.

In light of this, however, within each stage of the development process the separate research methodologies were not combined as such. And in summary stages 1-3 adopted a qualitative approach, whilst stage 4 was quantitative. The rest of the discussion will discussion will attempt to cover each stage in isolation before discussing the project objectives and limitations.

4.2 Discussion for stage 1

The impact of OA can vary between patients, the experiences in pain, disability and quality of life need to be put into the context of that individual (Nyvang, et al., 2016). In order to obtain this subjective perspective, it was important to conduct qualitative research by using semi-structured interviews and analyse the themes yielded from the information obtained.

This approach is similar to how the OHS and OKS were designed in that 20 patients were interviewed initially to obtain a list of themes to be incorporated in their questionnaires (Dawson, et al., 1996 and 1998). In this study the initial qualitative interview phase had 50 patients and it was felt that this increased the range of potential concepts that would be discovered. Given the range of patient age (43 – 87) and diversity in the recruited group (medical co-morbidities and extent of OA) a variety of activities were listed which demonstrates the diverse impact OA can have on individuals. For example, more rare activities that were mentioned included having a pedicure and reading books. Whereas the most common aspects affected by OA were due to reduced mobility such as walking.

When the list of themes obtained from this stage are compared with the questions of the OHS, OKS and WOMAC there are similarities between the different PROMs (Table 23). Although the context in which these themes were asked vary between PROMs the core themes are still being represented. This reinforces the fact that there are aspects common to persons suffering from OA and these have been revealed from this stage of the development process.

Table 26: List of most common theme and relative question found in other PROMs

New PROM	OHS	OKS	WOMAC
Theme	*(Question theme related to)*		
Walking	Q6	Q4	Q13
Using stairs (Up or Down)	Q7	Q12	Q8,Q9
Getting in and out of cars	Q3	Q3	Q14
Getting in and out of a bathtub			Q20
Doing housework	Q11	Q9	Q23,Q24
Standing up from sitting position	Q8	Q5	Q10
Pain	Q1, Q10,Q12	Q1, Q8	Q1-5
Doing shopping	Q5	Q11	Q15
Putting on shoes	Q4		Q16
Getting dressed			
Getting in and out of bed			Q16
Showering	Q2	Q2	
Socialising with friends and relatives			
Kneeling		Q7	
Feeling breathless			
Cold weather affecting pain			
Gardening			
Playing sports			
Carrying things			
Anxiety			
*Total number of questions in OHS and OKS = 12 **Total number of questions in WOMAC = 24			

When considering the list of themes obtained from the qualitative interviews it became apparent that they could be placed into certain categories depending on which aspect of the patient's life was affected by OA. To illustrate this, the three activities "playing sports", "walking the dog" and "hiking" could all be categorised as "recreational activities". This creation of categories and refinement of them eventually led to the five subdomains used in the new PROM.

4.3 Discussion for stage 2

With regards to perceptions of OA treatment goals there is disparity between patients and clinicians (Ramkumar, et al., 2015). Patients would give priority to elimination of symptoms and return to physical activity, whereas clinicians prioritise aspects of treatment related to treatment safety (Cordero-Ampuero, et al., 2012). Due to this discrepancy, it was felt that obtaining the point of view of the clinicians was integral in understanding holistically the aspects of treatment of OA in designing the new PROM.

A representative sample of experts from the different disciplines that manage OA in the hospital clinic setting was recruited. Based on qualitative interviews similar themes such as pain relief and regaining of function were highlighted to a certain degree. Regarding the new themes that were added, these were mainly focused on specific details pertaining to walking. From a patient's point of view, it may be enough for them to be able to walk from their front door to the bus stop, for example, and it is not so much of a concern about the speed at which they get there. As the PROM is mainly designed to assess these issues from the patient perspective, it was felt that questions should be more focused on the practical aspects of physical function rather than the technical.

From the iterative rounds for ranking the themes to be included in the questionnaire only "standing on tip toes" was removed for having a low ranking. This indicated that the remaining themes were viewed as being relevant when it came to assessing the level of function of a patient with OA. With a final count of 70 themes in the item bank after this stage, it was felt that the research team had comprehensively investigated aspects relating to physical function, pain and general health in patients suffering from OA.

4.4 Content validity

The COSMIN study design checklist was used to evaluate the methodological design in creation of the new PROM. The checklist is included in Appendix A. It was felt that this study adhered to the recommendations well according to their four-point marking scale. The design requirements and the score assigned based on this project's research methodology are summarised in the follow table (Table 24). It was felt that Content validity of the new PROM was confirmed based on the fact that the questions included in the PROM were relevant to a patient suffering from OA of the lower limb.

Table 27: List of design requirements from COSMIN checklist and score assigned

Content validity recommendations	**Score assigned**
Design requirements	
Use appropriate method to assess relevance, comprehensiveness and comprehensibility of PROM from perspective of patient	Very good
Use appropriate method to assess relevance and comprehensiveness of PROM from perspective of professionals	Very good
Include professionals from all relevant disciplines	Very good
Evaluate each item in an appropriate number of patients or professionals	Very good
Use skilled interviewers	Adequate
Base interviews on interview guide	Very good
Record and transcribe interviews	Adequate
Analyses	
Use appropriate approach to analyse data	Very good
Involve at least two researchers in the analysis	Very good

Only adequate scores were obtained for skilled interviewers and recording and transcribing of the interviews. This was due to the lack of experience and formal training of the researchers and the interviews were not transcribed verbatim. To mitigate this limitation a robust interview guide was created for the structured interviews.

4.5 Discussion for stage 3

4.5.1 Questionnaire length

The first consideration in formulating the questionnaire was the number of questions to be included. Although there is an association between questionnaire length and respondent burden one needs to consider the context of the questions included as well (Rolstad, et al., 2011). It was decided that 20 questions from the item bank would be ideal as this would fit on one page of A4 paper with the overleaf containing the body map and VAS questions. Although this would essentially make the new PROM "longer" than the OHS for example, the OHS is often combined with other questionnaires like the SF-12 therefore increasing the total length of all questionnaires that need to be completed by the patient.

In the context of collecting data for the NJR a patient would be given the OHS or OKS, SF-12, EQ-5D and WOMAC to complete in one sitting. This would amount to 6 or 7 pages of questions depending on layout. Iglesias and Togerson (2000) was able to conclude that increasing the number of pages from 5 to 7 reduces response rates for mailed questionnaires. If the new PROM were to be validated for this purpose, theoretically, the burden on the patient based on length of questionnaire can be reduced significantly.

4.5.2 Questionnaire design

The format and design of the first 20 questions was in such a way as to allow for it to be simple and intuitive for the patient to answer. The first 15 questions containing the lower limb, upper limb and role limitation domains followed the same format, allowing patients to provide perceptions on their level of function (Figure 17). For the pain subdomain the instructions were intended to be clear to allow patient to indicate level of pain at rest, and how pain limits activity (Figure 18). The general health subdomain questions were included in order to screen

whether or not a patient's functional limitation is affected by other factors apart from their symptoms due to OA, for example being out of breath limiting the ability to climb stairs (Figure 19).

For the following items please select the response that best describes your **level of function** on average over the **last month** (*For each item please tick one box per row*)						
		Able without problems	**Able but a little difficult**	**Able but moderately difficult**	**Able but very difficult**	**Unable**
When I Need to:						
01	Stand up from a chair	☐	☐	☐	☐	☐
02	Put on footwear	☐	☐	☐	☐	☐
03	Get in and out of a car	☐	☐	☐	☐	☐
04	Walk for 10 minutes	☐	☐	☐	☐	☐
05	Go up a flight of stairs	☐	☐	☐	☐	☐
06	Go down a flight of stairs	☐	☐	☐	☐	☐
07	Carry things (e.g. shopping bag)	☐	☐	☐	☐	☐
08	Do up buttons	☐	☐	☐	☐	☐
09	Reach out for something at shoulder height	☐	☐	☐	☐	☐
10	Turn a key	☐	☐	☐	☐	☐
11	Prepare a meal	☐	☐	☐	☐	☐
12	Do my regular job or daily routine if retired	☐	☐	☐	☐	☐
13	Perform leisure or sporting activities (e.g. Dancing, bowling, gardening)	☐	☐	☐	☐	☐
14	Do Housework	☐	☐	☐	☐	☐
15	Go shopping on my own	☐	☐	☐	☐	☐

Figure 17: Questions 1-15 of the new PROM using the same format

Please select the response that best describes your **level of pain while resting** (on average last month)						
		No pain	**Little pain**	**Moderate pain**	**Severe pain**	**Constant pain**
16	Level of pain	☐	☐	☐	☐	☐
Please select how **pain** from your joints limits your **overall function**						
		No Limitation	**Little limitation**	**Moderate limitation**	**A lot of limitation**	**Completely limited**
When I Need to:						
17	Use my legs	☐	☐	☐	☐	☐
18	Use my arms	☐	☐	☐	☐	☐

Figure 18: Questions 16-18 for assessing the pain subdomain

		No Limitation	Little limitation	Moderate limitation	A lot of limitation	Completely limited
19	How does your **general medical health (e.g. asthma, heart disease)** limit your overall function?	☐	☐	☐	☐	☐
20	How does your **mood (e.g. anxiety, depression)** limit your overall function?	☐	☐	☐	☐	☐

Figure 19: Questions 19 and 20 assessing general health and mood

The Likert scale was chosen as the rating scale for this PROM because if withs widespread use and familiarity (Bishop & Herron, 2015). And by having five closed ended responses the patient will be obliged to select the option that best fits their opinion. Also, each response of the Likert scale can be given a numerical value which can then be used for statistical analysis. Despite this however, Rensker et al (2018) showed that a VAS is the preferred rating scale among different types of PROMs.

4.5.3 Item inclusion

In the lower limb subdomain, the six items selected were intended to assess composite functions that would require the use of the whole limb to perform that activity. The items included all scored well during stage 2. Although “getting in and out a bath tub” scored high, it was decided that this question would be excluded as in terms of context not all patients would use a bath tub to bath and this item might be a potential left unanswered. “Getting in and out of bed” was also an item that received high scores however the composite movement of getting out of bed arguably includes the upper limb and back as well, therefore this item was reconsidered for the role limitation subdomain, but in view of the number of items already selected was ultimately left out.

In the upper limb subdomain, the items included aimed to assess five global functions of the upper limb which are required for daily activity. The item “carry things” aimed to assess overall arm strength. “Do up buttons” was included to assess patient’s fine motor movements. With the “reach for something at shoulder height” item the aim was to assess shoulder function as bringing the arm to at least shoulder height is important to allow patients to hand clothing, pick up objects from shelves and provide access to under the arm during bathing. The “Turn a key” and “Prepare a meal” items were intended to assess other different functional grips that the patient use during day to day activity. Although patients with OA of the upper limb were not included in the study population there are several PROMs for related to upper limb function (Smith, et al., 2012). As we have already established the importance of the kinetic chain, it was felt that dedicating a portion of the PROM to upper limb function was important to measure overall physical function (Karandikar & Vargas, 2011).

For the role limitation subdomain this included items which pertained to patient's daily activity in society and other complex daily activities. In agreement with Theis et al (2013), it was felt that social interaction was also an important aspect of daily function that could be affected by OA and that performing one's job or going for leisure activities would reflect this. With regards to "do housework" and "go shopping on my own", these items would help assess patients' level of independence as these are daily activities which would require overall good function to perform.

The pain and general health subdomains were subdomains with fewest items, three and two respectively. As we are already aware, pain is a common symptom in OA and can cause limitation in function therefore it was important to include questions that elicit this in the new PROM. As for the general health subdomain the aim was to ascertain whether medical health or mood were factors that affected a patient's level of function. These questions were included with examples to illustrate that the causes of the limitation in function may not solely be related to OA (World Health Organisation, 2020). Invariably a patient could have low mood due to the fact that they are severely limited in their activities, this was the reason why it was felt a question assessing a patient's level of satisfaction (the VAS) would be useful.

4.5.4 Body map

As the new PROM was designed to assess overall function, it could potentially be applied to a patient with OA in any joint. In terms of a PROM being a data collection tool, having an objective record of where the patient had pain built into the PROM has potential uses. It allowed the patient to conceptualise which joints caused the most problems and from the point of view of the clinician it provides an idea to the degree of severity of joints affected "at a glance" (Brummet, et al., 2016). The body map is also useful to correlate a patient's answers

to the first 20 questions to the joints affected by OA. For example, if the patient scored low on the lower limb subdomain only and then indicated a majority of the joints in the lower limb as being painful, this would confirm that pain from OA largely contributed to the loss of function.

4.5.5 VAS

The VAS was included to assess the patient's level of satisfaction with their level of function. Information obtained from this question would be useful to the clinician to assess whether an intervention was beneficial to the patient. In theory, a THR would eliminate patient's pain coming from the hip and if mobility improved then so would their level of satisfaction. How recent a recent study has shown that although patient satisfaction does improve follow THR, there was only modest corellation between satisfactions scores and both the OHS and EQ-5D (Maillot, et al., 2020).

A VAS from 0 – 100 was used instead of a Likert type response to allow the patient to have greater freedom to answer this question.

4.6 Discussion for stage 4

4.6.1 Internal consistency

Interpretation of Cronbach alpha's needs to be done with caution as without looking into the context of the items of the subdomain, incorrect assumptions can be drawn from these values (Gardner, 1995). Calculation of Cronbach's alpha of the first 20 questions revealed an alpha value much higher that any of the individual subdomains. This could signify that although there are five subdomains proposed during questionnaire design there may still be a degree of interrelatedness between the subdomains. Conversely this may be as a result of the fact that Cronbach's alpha is also sensitive to the number of items used to calculate it (Sijtsma, 2009). This phenomenon has also been reported in another development study of a questionnaire (Tuan, et al., 2005). Certainly, interpretation of Cronbach's alpha is more useful when considering individual subdomains (Taber, 2018).

The Cronbach's alpha for each of the other subdomains were acceptable (>0.70) apart from the pain and symptoms subdomain. When items were deleted from the subdomain and alpha calculated there was only one observed increase in Cronbach's alpha and that was for "carry things" question in the upper limb function subdomain. Arguably, the increase was marginal with an increase from 0.87 to 0.89, this could suggest that carrying things does not match well with the other items in measuring upper limb function.

With the pain and symptoms subdomain the Cronbach's alpha was 0.52. This is likely because the items included in this subdomain asked questions about pain in two difference circumstances; when patient is at rest and when they need to use their arms or legs. In this regard a myriad of combinations could be possible for each individual within the extent of their disease. For example, a patient may be severely limited while walking due to pain

however experience very little pain at rest. Alternatively, severe pain at rest might preclude patient from using the affected limb altogether. This may not necessary be a weakness in design as from the clinical point of view both facets of pain symptoms are important to elicit from the patient. The interpretation of Cronbach's alpha may suggest that the pain subdomain does have weakened internal consistency compared to the other subdomains in the PROM (Tavakol & Dennick, 2011).

Given that overall the Cronbach's alpha for the other subdomains were acceptable, when it comes to considering the total scores of the subdomain it can assumed that the subdomain measures a singular construct. When it comes to assessing correlations with other established PROMs, to confirm criterion validity, then using the total scores of the subdomains can be done with a degree of confidence.

It is worth noting that based on the COSMIN checklist it is recommended that >100 patients be used to calculate internal consistency to get a "very good" score for statistical analysis. The recruitment of 60 patients for this part of the study fell under the next best category, 50 – 99 patients, providing a score of "adequate".

4.6.2 Test-retest reliability

Thirteen patients did not return the second PROM which was given to them during recruitment. This meant that there was a response rate of 78.33%. Unreturned PROMs were treated as missing data and a decision needed to be made on how to deal with this. As data from the whole questionnaire was missing, techniques to replace these missing values would not be possible and these pairs were excluded from the analysis, thus reducing the sample size for test-retest reliability.

With the data that was available, ICC was calculated with the two-way mixed effects with absolute agreement model for ICC. This model was used at the patient was taken as an individual rater performing the same "test" (the new PROM) twice at different instances (Koo & Li, 2016). The ICC obtained for each item in the questionnaire showed good test-retest reliability with all items, except question number 18, having ICC value at least >0.70. Question number 18 asked a patient about how pain would limit their overall function when trying to use their arms and had an ICC coefficient score of 0.61 which still shows moderate reliability.

Overall, for a recall period of 48hrs, it was felt that the overall ICC scores for each item were good and therefore demonstrated that the PROM was reliable in terms of its test-retest reliability. However, due to the loss of data and lower number of patients based on COSMIN recommendations, statistical validation of reliability was scored as "doubtful" only.

4.6.3 Criterion validity

Exploratory analysis of the data from the new PROM suggested that the data obtained was not normally distributed. Furthermore, as the responses were based on an ordinal scale by design the data could not be treated as continuous data. The application of numerical scores to each response was done to allow for statistical assessment. Based on these reasons it was decided that non-parametric tests would be used.

As was already established in the introduction there is no gold standard to measure what the new PROM would like to measure and therefore correlations were made against the PROMs which are mainly used in orthopaedic practise. As data from the OHS, OKS, WOMAC and SF12 would be collected as a matter of routine during clinic visits the additional burden to the patient during this part of the research project was limited to the addition of the new PROM only.

Patient recruitment

There is limited evidence in the literature for justification on the ideal sample size, however a minimum of 100 patients is ideal of exploratory factor analysis (Anthoine, et al., 2014). During the recruitment of patients for this stage of the project only 60 patients were recruited. Initially it was overlooked that patients who attended clinic for hip OA would not complete the OKS, and the converse was also true. Because of this, comparisons with the OHS and OKS were further limited as approximately half the patients recruited attended for hip OA (n=27) and the rest attended for knee OA (n=33). This divide did not affect comparisons with WOMAC or the SF12 as all patients completed these PROMs regardless of joint affected by OA.

Currently, there is no recognised gold standard for measuring level of physical function or outcomes following joint arthroplasty. More recent PROMs that have been developed will frequently use the more established WOMAC and Oxford Scores for comparison, for example the Knee injury and Osteoarthritis Outcome Score (KOOS) (Roos & Toksvig-Larsen, 2003) and the Forgotten Joint Score (Behrend, et al., 2012). Typically, the psychometric criteria of each PROM would be evaluated and compared. Correlations between the new PROM and the established ones would be performed to measure criterion validity. This project has adopted the same approach to satisfy criterion validity.

OHS correlations

As a measure of physical function, the new PROM showed good correlations with the OHS. This would suggest that the new PROM was able to measure similar concepts that were being measured by the OHS (namely pain and physical function). Interestingly better correlation was observed with upper limb function compared to lower limb function.

OKS correlations

The correlations with the OKS were strongest for the lower limb function, role limitation and pain and symptoms subdomain of the PROM. As the OKS was designed to assess pain and function of patients with OA in the context of having a joint replacement, correlating well in these subdomains lends credit to the new PROM being able to similarly measure these concepts. The upper limb function subdomain did have moderate correlation as a measure of function, but it would be expected to be less compared to the lower limb subdomain. As for the general health subdomain this showed weak correlations which may suggest a more divergent quality.

WOMAC

The WOMAC was spilt into its pain and physical function subdomains for analysis. All the subdomains and the body map were compared and there was moderate correlation for the pain subdomains. When compared to the physical function subdomain there was moderate correlations for all of the subdomains of the new PROM. The correlation analysis with the WOMAC suggested that there were at least moderate correlations across the subdomains of the new PROM. Given that the WOMAC is a disease specific PROM having moderate correlation with the new PROM would suggest that it can be used to assess patients suffering from OA.

SF12

It was found that there was poor correlation for all subdomains when compared to the PCS. The correlations for the MCS were also relatively weak compared to the OHS, OKS and WOMAC. The reason for the poorer correlations is likely to be because the PCS and the MCS are composite scores derived from 8 different subdomains which were originally descried in

the SF36. Because of this the subdomains of the new PROM were likely to be measuring dissimilar concepts compared to the SF12.

4.6.4 Ceiling and floor effects

Ceiling and floor effects of the OHS were investigated by Lim et al (2015), and it was concluded that the OHS did not exhibit any ceiling or floor effects based on overall scores. On testing the overall data this study sample, no floor effects were observed. And the only ceiling effect observed was from the general health subdomain. Twenty-one patients (35%) reported full scores for this subdomain. When interpreting this subdomain, a maximum score would indicate that the patients were not limited in their physical function by other medical issues or their mood. When considering this as a "true" ceiling effect it needs to be viewed in the context of how the patient would answer the rest of the questionnaire. As the new PROM is designed to assess physical function of patients with OA, a certain degree of limitation may be due to other factors apart from symptoms from OA, but there will be a proportion of patients who are free from other co-morbidities that could cause functional limitations.

However, when the overall score of the first 20 questions of the new PROM are considered, there were no floor or ceiling effects. In this regard, when considering the new PROM in its entirety, could be suggested that the PROM is able to measure the full spectrum of physical function in patients suffering from OA (from completely able to completely disabled) as the lack of floor or ceiling effects would mean that there is enough range within the questionnaire for patients to assess their level of function.

4.7 Meeting project objectives

The objectives of this project were met with the creation of the new PROM. The initial population was defined as adult patients who suffer from OA of the lower limb. Patients were then recruited from this population for the different stages of questionnaire development. A representative group of patients were interviewed to obtain a list of themes which were investigated and included in an item bank.

An expert panel was successfully recruited for the second stage of questionnaire development. Opinions of the expert panel were recorded through semi-structured interviews. Questionnaires were used to allow the expert panel to rank items according to relevance. The items that were ranked satisfied content validity and would later be used to build the new PROM.

The PROM was created successfully from the item bank, the body map, and the VAS. The questionnaire was then assessed by another group of patients to qualitatively assess its design and content. Following this, minor revisions were done, and the final version of the new PROM was created.

Lastly a final group of patients were recruited to perform quantitative analysis to validate the questionnaire. Measurement properties that were validated include: internal consistency, test-retest reliability, criterion validity and floor and ceiling effects.

4.9 Project limitations

There are several limitations in this project.

The first limitation was that the number of patients recruited was less than 100. For any validation study it is recommended that there be at least 100 patients in order to increase statistical power. Due to this limitation, conclusions drawn from any quantitative analyses must be acknowledged as being made in light of this limitation. However, as this is a preliminary study to assess the measurement properties of a newly developed PROM, the results can still be taken to demonstrate the potential of it having robust psychometric criteria.

Another limitation is that the subdomains created, and the items included in each, were based of categories selected by pooling together common themes derived from the qualitative interviews with the recruited participants. Ultimately the selections were made at the discretion of the research team, albeit adhering to PROM development recommendations, and there will be an inherent weakness as the subjective opinions of the research team cannot completely be eliminated. If more patients were recruited, then exploratory factor analysis and confirmatory factor analysis could have been potential techniques used to identify subdomains and reduced the number of items.

A further limitation was the body map question. The method selected to assess criterion validity for this question was to use it as a substitute for a pain score and try to correlate it against other pain subdomains in other PROMs. The research team felt that this was the next best method to provide an attempt to validate it ensure that it assess what it was intended to. As the use of a body map is a novel concept for PROMs related to OA its potential use will have to be constrained until it can be fully validated.

4.10 Further work

There is scope for further work with the development of this new PROM. Due to the limited number of patients recruited, work from this project can be treated as a pilot study for further validation studies on a larger scale. Furthermore, a psychometric criterion which has not been investigated in this project was the sensitivity to detect change, and by extension, in the clinical setting the minimum clinical important difference (MCID). To do this, data before and after an intervention needs to be collected. In the context of OA, joint replacement surgery would be the ideal intervention to use as it has been proven to provide life changing benefit to the patient. Assessing sensitivity to change was beyond the scope of this project, but potential further work in this area is justified to completely validate the PROM and to detect the MCID.

Chapter 5: Conclusion

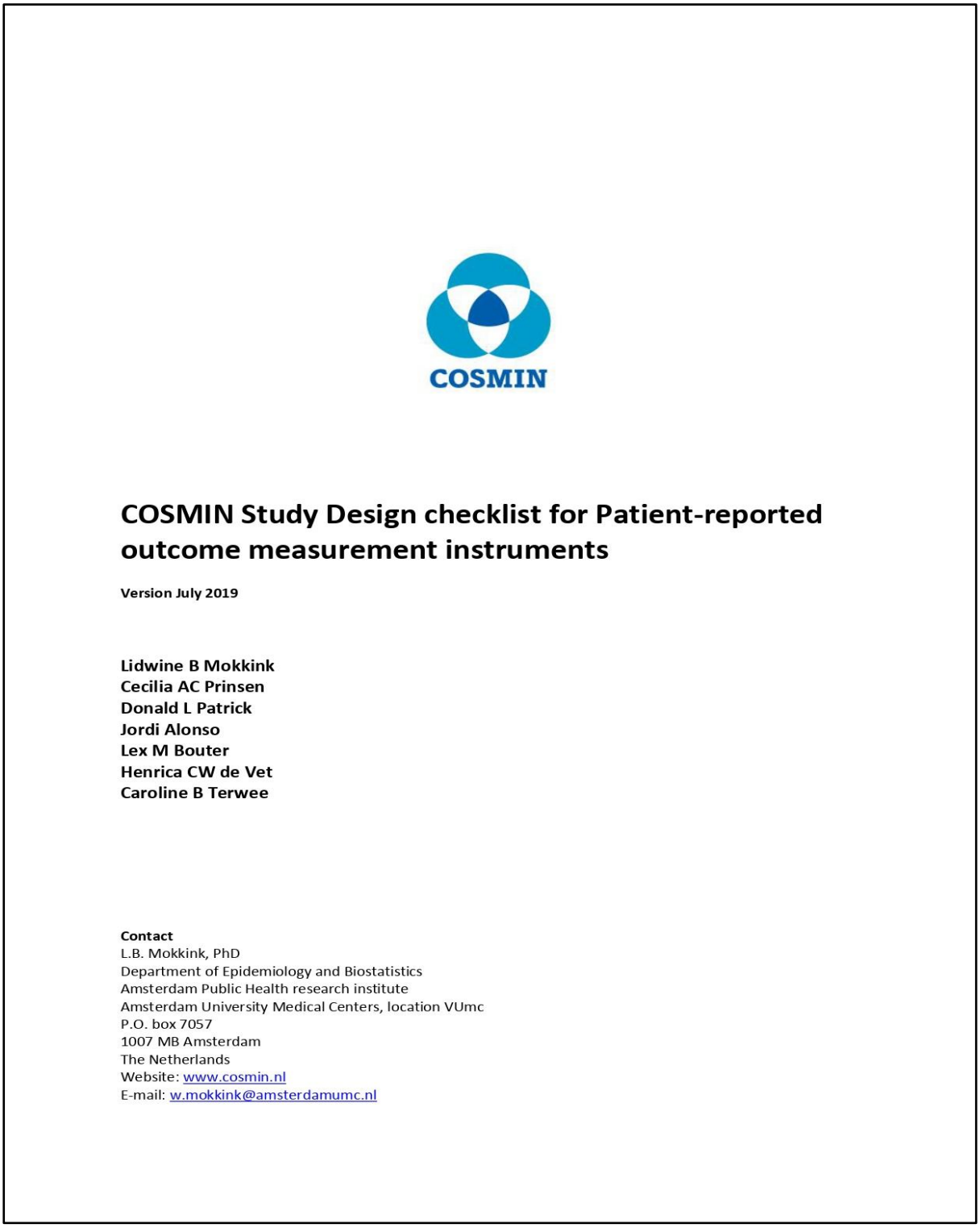
The completion of this project has led to the development of a new PROM for patients with OA that assesses the overall level of function. The population that the PROM should be used for has been clearly defined and evaluation of the themes measuring overall function have been undertaken in this same patient group. Through qualitative interviews and obtaining expert opinion the PROM has been developed from a list of items that satisfy content validity. The prototype of the questionnaire has been assessed qualitatively and has been found to be suitable by the target population.

Analysis of the new PROM has shown acceptable internal consistency across its subdomains and good test-retest reliability. On testing for criterion validity, it has demonstrated moderate correlations with the OHS and OKS, two of the most used joint specific PROMS, and the WOMAC, a well validated disease specific PROM.

Although this new PROM shows some promise the work outlined in this project only constitutes to preliminary study aimed to assess the viability of using such a PROM as a replacement or supplement to the more established and well-studied PROMs. Further work is required to fully validate the new PROM before it can be used in the general population. The PROM needs to be validated on a wider scale, sensitivity to detect change needs to be investigated and the MCID should also be explored.

Appendices

Appendix A: Cosmin study design checklist (Pages 1-5 of 32)



COSMIN

COSMIN Study Design checklist for Patient-reported outcome measurement instruments

Version July 2019

Lidwine B Mokkink
Cecilia AC Prinsen
Donald L Patrick
Jordi Alonso
Lex M Bouter
Henrica CW de Vet
Caroline B Terwee

Contact
L.B. Mokkink, PhD
Department of Epidemiology and Biostatistics
Amsterdam Public Health research institute
Amsterdam University Medical Centers, location VUmc
P.O. box 7057
1007 MB Amsterdam
The Netherlands
Website: www.cosmin.nl
E-mail: w.mokkink@amsterdamumc.nl

Table of Content

List of abbreviations

CTT:	classical test theory
IRT/Rasch:	Item Response Theory and Rasch analyses
NA:	not applicable
Original CC:	original COSMIN checklist [1]
PROM:	patient-reported outcome measure
RoB:	Risk of Bias; it refers to the COSMIN Risk of Bias checklist [2]

Instructions

The COSMIN Study Design checklist is recommended for designing studies to evaluate measurement properties of existing patient-reported outcome measures (PROMs). It can be used by researchers and clinicians or other professionals who are designing a study to evaluate measurement properties of an existing PROM, or by e.g. scientific committees and medical ethics committees who are appraising protocols of studies on measurement properties, or by reviewers for scientific journals that will publish study protocols of studies on measurement properties of PROMs.

The COSMIN Study Design checklist is based on the original version of the COSMIN checklist [1] [3], as well as on the recently developed COSMIN Risk of Bias checklist for PROMs [2]. Decisions on adaptations were made based on iterative discussions by the COSMIN steering committee, both at face-to-face meetings (LM, CP, HdV and CT) and by email discussions (entire COSMIN steering committee, i.e. all authors).

The COSMIN Study Design checklist consists of ten boxes. The first box, i.e. General recommendations for designing a study on a measurement properties, is relevant for all studies. It contains general standards which should be considered in the design of a study on any measurement property. The remaining boxes contain standards for specific studies on each of the nine measurement properties, i.e. content validity, structural validity, internal consistency, cross-cultural validity\measurement invariance, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness [2 4]. In addition, we provide standards for translating an existing PROM in the box Translation process.

In this checklist, each standard is also accompanied with a 4-point rating scale. This rating scale is based on the COSMIN Risk of Bias checklist [2]. The 4-point rating scale is added for illustrative purposes to better understand the consequences of choices made in the design of a study for the methodological quality of the study; it is not intended to be used to actually provide an overall rating (i.e. based on the worst-score counts principle) for your study design. The purpose of this checklist is only to check whether all important issues are considered when designing a study on measurement properties. Details on how to design and analyse these studies are described in the book Measurement in Medicine [5]. A clarification and explanation of most of the individual standards can be found in the user manuals (www.cosmin.nl) accompanying the COSMIN Risk of Bias checklist [6 7]. References, e.g. for sample size requirements, are also provided in the COSMIN user manuals [6 7].

Standards refer to potential risk of bias issues, reporting issues, or sample size issues. In this document for each standard a justification is added, referring to the number of the box (number between brackets refer to number of the specific standard in the specific box) from the COSMIN Risk of Bias checklist [2] (RoB), or the original COSMIN checklist [1] (CC); or by indicating that the standard is about sample size or it is a newly added standard.

General recommendation for the design of a study on measurement properties

The box General recommendations for designing a study on measurement properties is relevant for all studies on measurement properties. The aim of a study evaluating a measurement property of a PROM is to investigate (one or more aspects of) the quality of the PROM at issue. These studies require a clear research aim (i.e. referring to the measurement properties of interest), a clear description of the PROM and a clear description of the study population. The quality of a PROM should be determined in the target population in which the PROM will be used, because the results of studies on measurement properties depend on the sample included in the study.

General recommendations for the design of a study on measurement properties	**very good**	**adequate**	**doubtful**	**inadequate**	**justification**
Research aim					
1 Provide a clear research aim, including (1) the name and version of the PROM, (2) the target population, and (3) the measurement properties of interest	Research aim clearly described			Research aim not clearly described	New
PROM					
2 Provide a clear description of the construct to be measured	Construct clearly described			Construct not clearly described	RoB Box 1
3 Provide a clear description of the development process of the PROM, including a description of the target population for which the PROM was developed	Development process clearly described		Development process clearly described		RoB Box 1
4 The origin of the construct should be clear: provide a theory, conceptual framework (i.e. reflective or formative model) or disease model used or clear rationale to define the construct to be measured	Origin of the construct clear		Origin of the construct not clear		RoB Box 1

5 Provide a clear description of the structure of the PROM (i.e. the number of items and subscales included in the PROM, instructions given and response options) and its scoring algorithm	Structure and scoring algorithm clearly described			Structure and scoring algorithm not clearly described	RoB Box 1
6 Provide a clear description of existing evidence on the quality of the PROM	Existing evidence on the quality of the PROM clearly described		Existing evidence on the quality of the PROM not clearly described		New
7 Provide a clear description of the context of use*	Context of use clearly described		Context of use not clearly described		RoB Box 1
Target population					
8 Provide a clear description of in- and exclusion criteria to select patients, e.g. in terms of disease condition and characteristics like age, gender, language or country, and setting (e.g. general population, primary care or hospital/rehabilitation care)	In- and exclusion criteria for patients clearly described			In- and exclusion criteria for patients not clearly described	Characteristic of study population [6]
9 Provide a clear description of the method used to select the patients for the study (e.g. convenience, consecutive, or random)	Method for patient selection clearly described		Method of patient selection not clearly described		New
10 Describe whether the selected sample is representing the target population in which the PROM will be used in terms of age, gender, important disease characteristics (e.g. severity, status, duration)	Study sample representing the target population clearly described	Assumable that the study sample is representing the target population, but not clearly described	Unclear whether the study sample is representing the target population	Study will not be performed in a sample representing the target population	RoB Box 1

* The context of use refers to the intended application of the PROM (e.g. for research or clinical practice), to a specific setting for which the PROM was developed (e.g. for use in a hospital or at home) or to a specific administration mode (e.g. paper or computer-administered). If the PROM was developed for use across multiple contexts, this should be described.

Appendix B: Oxford hip score

Problems with your hip

During the past 4 weeks..

✓tick one box for every question.

1. ***During the past 4 weeks***.....

How would you describe the pain you usually had from your hip?

None	Very mild	Mild	Moderate	Severe
❑	❑	❑	❑	❑

2. ***During the past 4 weeks***.....

Have you had any trouble with washing and drying yourself (all over) because of your hip?

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
❑	❑	❑	❑	❑

3. ***During the past 4 weeks***.....

Have you had any trouble getting in and out of a car or using public transport because of your hip? *(whichever you tend to use)*

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
❑	❑	❑	❑	❑

4. ***During the past 4 weeks***.....

Have you been able to put on a pair of socks, stockings or tights?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
❑	❑	❑	❑	❑

5. ***During the past 4 weeks***.....

Could you do the household shopping on your own?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
❑	❑	❑	❑	❑

6. ***During the past 4 weeks***.....

For how long have you been able to walk before pain from your hip becomes severe? (*with or without a stick)*

No pain/ More than 30 minutes	16 to 30 minutes	5 to 15 minutes	Around the house only	Not at all -pain severe on walking
❑	❑	❑	❑	❑

P.T.O./

During the past 4 weeks...

✓tick one box for every question

7 ***During the past 4 weeks*.....**

Have you been able to climb a flight of stairs?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
❑	❑	❑	❑	❑

8 ***During the past 4 weeks*.....**

After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your hip?

Not at all painful	Slightly painful	Moderately painful	Very painful	Unbearable
❑	❑	❑	❑	❑

9 ***During the past 4 weeks*.....**

Have you been limping when walking, because of your hip?

Rarely/ never	Sometimes, or just at first	Often, not just at first	Most of the time	All of the time
❑	❑	❑	❑	❑

10 ***During the past 4 weeks*.....**

Have you had any sudden, severe pain - 'shooting', 'stabbing' or 'spasms' - from the affected hip?

No days	Only 1 or 2 days	Some days	Most days	Every day
❑	❑	❑	❑	❑

11 ***During the past 4 weeks*.....**

How much has pain from your hip interfered with your usual work *(including housework)*?

Not at all	A little bit	Moderately	Greatly	Totally
❑	❑	❑	❑	❑

12 ***During the past 4 weeks*.....**

Have you been troubled by pain from your hip in bed at night?

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
❑	❑	❑	❑	❑

Appendix C: Oxford knee score

PROBLEMS WITH YOUR KNEE

During the past 4 weeks..

✓tick one box for every question

1 *During the past 4 weeks.....*

How would you describe the pain you usually have from your knee?

None	Very mild	Mild	Moderate	Severe
❑	❑	❑	❑	❑

2 *During the past 4 weeks.....*

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	Impossible to do
❑	❑	❑	❑	❑

3 *During the past 4 weeks.....*

Have you had any trouble getting in and out of a car or using public transport because of your knee? (whichever you would tend to use)

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
❑	❑	❑	❑	❑

4 *During the past 4 weeks.....*

For how long have you been able to walk before pain from your knee becomes **severe**? (*with or without a stick)*

No pain/ More than 30 minutes	16 to 30 minutes	5 to 15 minutes	Around the house only	Not at all - pain severe when walking
❑	❑	❑	❑	❑

5 *During the past 4 weeks.....*

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 *During the past 4 weeks.....*

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
❑	❑	❑	❑	❑

/P.T.O

During the past 4 weeks...

✓tick one box for every question

7 *During the past 4 weeks.....*

Could you kneel down and get up again afterwards?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
❑	❑	❑	❑	❑

8 *During the past 4 weeks.....*

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 *During the past 4 weeks.....*

How much has pain from your knee interfered with your usual work *(including housework)*?

Not at all	A little bit	Moderately	Greatly	Totally
❑	❑	❑	❑	❑

10 *During the past 4 weeks.....*

Have you felt that your knee might suddenly 'give way' 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 *During the past 4 weeks.....*

Could you do the household shopping on your own?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
❑	❑	❑	❑	❑

12 *During the past 4 weeks.....*

Could you walk down one flight of stairs?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
❑	❑	❑	❑	❑

Appendix D: WOMAC

Name:_____ Date:_____

WOMAC OSTEOARTHRITIS INDEX

1. The following questions concern the **amount of pain** you are currently experiencing in your **Hip or knees**. For each situation, please enter the amount of pain you have experienced in the past **48 hours**.

	None	mild	Moderate	severe	extreme
A. Walking on a flat surface	☐	☐	☐	☐	☐
B. Going up or down stairs	☐	☐	☐	☐	☐
C. At night while in bed	☐	☐	☐	☐	☐
D. Sitting or lying	☐	☐	☐	☐	☐
E. Standing upright	☐	☐	☐	☐	☐

Stiffness	None	mild	Moderate	severe	extreme
2. How severe is your stiffness after first awakening in the morning?	☐	☐	☐	☐	☐
	None	**mild**	**Moderate**	**severe**	**extreme**
3. How severe is your stiffness after sitting, lying, or resting later in the day?	☐	☐	☐	☐	☐

4. The following 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 have experienced in the last **48 hours**, in your **Hips or knees**.

What degree of difficulty do you have with:

	None	mild	Moderate	severe	extreme
A Descending (going down) stairs	☐	☐	☐	☐	☐
B Ascending (going up) stairs	☐	☐	☐	☐	☐
C Rising from sitting	☐	☐	☐	☐	☐
D Standing	☐	☐	☐	☐	☐
E Bending to floor	☐	☐	☐	☐	☐
F Walking on a flat surface	☐	☐	☐	☐	☐
G Getting in/out of car	☐	☐	☐	☐	☐
H Going shopping	☐	☐	☐	☐	☐
I Putting on socks/stockings	☐	☐	☐	☐	☐
J Rising from bed	☐	☐	☐	☐	☐
K Taking off socks/stockings	☐	☐	☐	☐	☐
L Lying in bed	☐	☐	☐	☐	☐
M Getting in/out of bath	☐	☐	☐	☐	☐
N Sitting	☐	☐	☐	☐	☐
O Getting on/off toilet	☐	☐	☐	☐	☐
P Heavy domestic duties (mowing the lawn, lifting heavy grocery bags)	☐	☐	☐	☐	☐
Q Light domestic duties (such as tidying a room, dusting, cooking)	☐	☐	☐	☐	☐

Patient Identifier:_____

Appendix E: SF-12

SF-12 Health Survey

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities. **Answer each question by choosing just one answer**. If you are unsure how to answer a question, please give the best answer you can.

1. **In general, would you say your health is**:

☐₁ Excellent ☐₂ Very good ☐₃ Good ☐₄ Fair ☐₅ Poor

The following questions are about activities you might do during a typical day. Does <u>your health now limit you</u> in these activities? If so, how much?

	YES, limited a lot	**YES, limited a little**	**NO, not limited at all**
2. **Moderate activities** such as moving a table, pushing a vacuum cleaner, bowling, or playing golf.	☐₁	☐₂	☐₃
3. Climbing **several** flights of stairs.	☐₁	☐₂	☐₃

During the <u>past 4 weeks</u>, have you had any of the following problems with your work or other regular daily activities <u>as a result of your physical health</u>?

	YES	**NO**
4. **Accomplished less** than you would like.	☐₁	☐₂
5. Were limited in the **kind** of work or other activities.	☐₁	☐₂

During the <u>past 4 weeks</u>, have you had any of the following problems with your work or other regular daily activities <u>as a result of any emotional problems</u> (such as feeling depressed or anxious)?

	YES	**NO**
6. **Accomplished less** than you would like.	☐₁	☐₂
7. Did work or activities **less carefully than usual**.	☐₁	☐₂

8. **During the <u>past 4 weeks</u>, how much <u>did pain interfere</u> with your normal work (including work outside the home and housework)?**

☐₁ Not at all ☐₂ A little bit ☐₃ Moderately ☐₄ Quite a bit ☐₅ Extremely

These questions are about how you have been feeling during the <u>past 4 weeks</u>.
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 <u>past 4 weeks</u>…

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
9. Have you felt calm & peaceful?	☐₁	☐₂	☐₃	☐₄	☐₅	☐₆
10. Did you have a lot of energy?	☐₁	☐₂	☐₃	☐₄	☐₅	☐₆
11. Have you felt down-hearted and blue?	☐₁	☐₂	☐₃	☐₄	☐₅	☐₆

12. **During the <u>past 4 weeks</u>, how much of the time has your <u>physical health or emotional problems</u> interfered with your social activities (like visiting friends, relatives, etc.)?**

☐₁ All of the time ☐₂ Most of the time ☐₃ Some of the time ☐₄ A little of the time ☐₅ None of the time

Patient name: Date: PCS: MCS:

Visit type (circle one)
Preop 6 week 3 month 6 month 12 month 24 month Other:_____

Appendix F: Patient information sheet

UNIVERSITY OF LIVERPOOL

IRAS: 207639
Study: DevPROM
Version 2.1: 15/7/16

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

PATIENT PARTICIPANTS INFORMATION SHEET

Study Title: Development and Validation of a Patient Reported Outcome Measure Assessing Global Physical Function in Patients Suffering from Lower Limb Osteoarthritis

Contact information:

Mr Joseph Alsousou (Principal investigator)
T: 0151 794 8971
E: josephalsousou@doctors.org.uk

Ms Amanda Wood (Clinical Research Nurse)
T: 0151 794 8993
E: Amanda.wood@liverpool.ac.uk

Address:
Mr Alsousou or Amanda Wood,
North West Cancer Research Centre,
200 London Road, Liverpool, L3 9TA.

We would like to invite you to take part in a research study. Before you decide whether to take part, you need to understand why the research is being done and what it will involve. Please take the time to read the following information carefully and discuss it with others if you wish. This should take about 5 minutes of your time. If you have any questions do not hesitate to contact anyone listed at the end of this sheet.

Study Title

Development and validation of a patient reported outcome measure assessing lower limb joint function.

What is the purpose of this study?

Questionnaires or outcome measures are tools used to assess quality of life. They can be made up of different aspects such as pain, sleep function and mobility. These questionnaires are used to try and understand a patients' level of health at a certain point in time. They may also be used at different times to assess any potential changes in health due to treatment or progression of disease. Currently, several questionnaires are being used to assess mobility in patients with osteoarthritis of the hip, knee, foot or ankle. You may have already filled in similar questionnaires before. Although these are the best available, they do come with problems, for example some people may find the questions irrelevant, too difficult to understand or may take too much of their time! This study aims to develop a new questionnaire that resolves some of these problems. We are inviting patients with osteoarthritis of the hip, knee, ankle or foot. The study will be performed at the Royal and Broadgreen University Hospital Trust. This study will take part in many stages but you are being invited to one of them. Please continue reading to find out more details.

Why have I been chosen?

You have been selected because you have osteoarthritis of the hip, knee, foot or ankle joints.

What would taking part involve?

We would like to take the opportunity to inform you about the study. The study is divided into several stages that aim to create a new questionnaire. You are being invited to one of these stages. The aim of this stage is to develop questions that may potentially be used in the questionnaire. If you agree to take part in this study you will be given a consent sheet to sign and return. You will be asked to attend the orthopaedic clinic at Broadgreen Hospital, where a suitable time will be organized with you. Some personal information will be noted which will include your age, gender, weight, height, duration of symptoms, physical activity levels, medical history, current and past treatments of your osteoarthritis. We will also need to confirm your diagnosis of osteoarthritis by looking at your clinical notes. This research will involve you having an up to 30 minute interview with one member of the research team. They will ask you to comment on any problems that you face as a result of your osteoarthritis. For example walking, sitting down and getting in and out of a car etc. The researcher will also ask you to comment on anything that you feel affects your ability to move. For example pain, stiff joints, fear of an injury etc. The information you provide will be noted down on paper and will help us to create the new questions. This information will be securely stored and anonymized so that your personal details will NOT have any identifying information on it. After the interview has ended, you will no longer be part of the study. However, if you are interested in participating in other parts of this research please inform the research team.

What are the advantages and disadvantages of taking part?

There is no direct benefit to participating in this research. Your data will remain anonymized for the entirety of the study. We do not anticipate any risks or harm towards you during our study.

What if something goes wrong?

If you have a concern about any aspect of this study, you should speak to the researchers who will do their best to answer your questions. The principal investigator Joseph Alsousou can also be contacted on 0151 794 8971, clinical research nurse Amanda Wood on 0151 794 8993

What will happen if I don't want to carry on with the study?

Your decision to take part in this study is entirely voluntary; you can change your mind at any time. Any decision to take part, reject or withdraw from the study will NOT affect the NHS care you receive. Should you wish to withdraw from the study, a member of the research team can be approached. Your intentions regarding any information collected will be ascertained for example destroying the information.

Will my information be kept confidential?

All information that is collected will be kept strictly confidential. Information collected from you will be anonymised and stored in secure storage. Identifiable information will be stored separately to the information you provide as part of the study. Storage is located at the University of Liverpool.

What will happen to my information after the study?

At the end of the study we will, with your consent, submit an anonymised data set containing your information to the University of Liverpool's Data Archive. This data will be held indefinitely and will assist other researchers

to undertake future studies. Any information will be fully anonymised and can't be traced back to you.

What will happen to the results of this study?

The results will be written up as part of a masters thesis and will be published in an academic journal. A summary of the results, the thesis or any publications will be offered to each participant. If you are interested in receiving any results please inform the research team. Postage costs will be paid by the research team. It is important to note that you will not be identifiable in any publication.

Who has reviewed the study?

This study has been ethically approved by the Health Research Authority. It is audited and sponsored by the University of Liverpool's Joint Research Office.

Further information and contact details?

If you would like to discuss this study with anyone please do not hesitate to contact anyone below. A copy of the contact details are also on the first page of this information sheet. If you would like to find out more about research in healthcare please visit http://www.hra.nhs.uk. If you have any complaints, please do not hesitate to contact the people below.

Mr Joseph Alsousou (Principal Investigator)
T: 0151 794 8971
E: josephalsousou@doctors.org.uk

Ms Amanda Wood (Clinical Research Nurse)
T: 0151 794 8993
E: Amanda.wood@liverpool.ac.uk

Appendix G: Consent form for patient participation

IRAS: 207639
Study: DevPROM
Version 2: 15/7/16

Centre Number:
Study Number:
Participant Identification Number for this trial:

Study Title: Development and Validation of a Patient Reported Outcome Measure Assessing Global Physical Function in Patients Suffering from Lower Limb Osteoarthritis

Consent Form 1 – Patient Participants.

Name of Researcher:

Please initial box

1. I confirm that I have read the information sheet dated..... (version.....) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by members of the research team and regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
4. I agree to be audio recorded during the interview. ☐
5. I agree for my information to be submitted to the University of Liverpool's Data Archive and held indefinitely for future research. ☐
6. I agree to take part in this study ☐

Name of Participant	Date	Signature

When completed: 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

Appendix H: Consent form for expert participation

UNIVERSITY OF LIVERPOOL

IRAS: 207639
Study: DevPROM
Version 2: 15/07/2016

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

Centre Number:

Study Number:

Participant Identification Number for this trial:

Study Title: Development and Validation of a Patient Reported Outcome Measure Assessing Global Physical Function in Patients Suffering from Lower Limb Osteoarthritis

Consent Form 2 – Expert Participants

Name of Researcher: Trupesh Patel

Please initial box

1. I confirm that I have read the information sheet dated..... (version.....) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐

3. I agree for my information to be submitted to the University of Liverpool's Data Archive and held indefinitely for future research. ☐

4. I agree to take part in this study ☐

Name of Participant	Date	Signature
Name of Person	Date	Signature

When completed: 1 for participant; 1 for researcher site file;

Appendix I: Ethical approval letter

NHS
Health Research Authority

Professor Simon P Frostick
Department of Molecular and Clinical Cancer Medicine
Royal Liverpool University Hospital
Liverpool
L69 3GA

Email: hra.approval@nhs.net

28 July 2016

Dear Professor Frostick

Letter of HRA Approval

Study title:	**Development and Validation of a Patient Reported Outcome Measure Assessing Lower Limb Joint Function**
IRAS project ID:	**207639**
Protocol number:	**UoL001228**
REC reference:	**16/NW/0534**
Sponsor	**University of Liverpool**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.

It is critical that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details and further information about working with the research management function for each organisation can be accessed from www.hra.nhs.uk/hra-approval.

Appendices

The HRA Approval letter contains the following appendices:

- A – List of documents reviewed during HRA assessment
- B – Summary of HRA assessment

After HRA Approval

The document *"After Ethical Review – guidance for sponsors and investigators",* issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

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

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

In addition to the guidance in the above, please note the following:

- HRA Approval applies for the duration of your REC favourable opinion, unless otherwise notified in writing by the HRA.
- Substantial amendments should be submitted directly to the Research Ethics Committee, as detailed in the *After Ethical Review* document. Non-substantial amendments should be submitted for review by the HRA using the form provided on the HRA website, and emailed to hra.amendments@nhs.net.
- The HRA will categorise amendments (substantial and non-substantial) and issue confirmation of continued HRA Approval. Further details can be found on the HRA website.

Scope

HRA Approval provides an approval for research involving patients or staff in NHS organisations in England.

If your study involves NHS organisations in other countries in the UK, please contact the relevant national coordinating functions for support and advice. Further information can be found at http://www.hra.nhs.uk/resources/applying-for-reviews/nhs-hsc-rd-review/.

If there are participating non-NHS organisations, local agreement should be obtained in accordance with the procedures of the local participating non-NHS organisation.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application

procedure. If you wish to make your views known please email the HRA at hra.approval@nhs.net. Additionally, one of our staff would be happy to call and discuss your experience of HRA Approval.

HRA Training

We are pleased to welcome researchers and research management staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

Your IRAS project ID is **207639**. Please quote this on all correspondence.

Yours sincerely

Dr Claire Cole

Senior Assessor

Email: hra.approval@nhs.net

Copy to: *Mr Alex Astor, University of Liverpool (Sponsor Contact)*
Ms Heather Rogers, Royal Liverpool and Broadgreen University Hospital (Lead NHS R&D Contact)
Mr Trupesh Patel (Student)

IRAS project ID	207639

Appendix A - List of Documents

The final document set assessed and approved by HRA Approval is listed below.

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Staff Recruitment Poster]	1	29 May 2016
Covering letter on headed paper [Cover Letter]	1	23 June 2016
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Insurance letter]	1	05 August 2015
IRAS Application Form [IRAS_Form_24062016]		24 June 2016
Letter from sponsor [Sponsorship Approval]	1	17 June 2016
Non-validated questionnaire [Baseline Assessment - Patient Participants]	1	29 May 2016
Non-validated questionnaire [Baseline Assessment Expert Participants]	1	29 May 2016
Other [Unfavourable Opinion NW PRESTON]	1	27 April 2016
Other [Email clarification re. quantitative aspects of the study]		29 June 2016
Other [Schedule of Events]	1.0	25 July 2016
Other [Statement of Activities]	1.0	25 July 2016
Participant consent form [Consent Form 1 - Patient Participants]	2	15 July 2016
Participant consent form [Consent Form 2]	3	27 July 2016
Participant information sheet (PIS) [Patient information sheet 1 - item generation]	2	15 July 2016
Participant information sheet (PIS) [Patient information sheet 2 - qualitative pilot]	2	15 July 2016
Participant information sheet (PIS) [Patient information sheet 3 - responsiveness]	2	15 July 2016
Participant information sheet (PIS) [Patient information sheet 4 - test retest]	2	15 July 2016
Participant information sheet (PIS) [Experts information sheet - delphi analysis]	2	15 July 2016
Research protocol or project proposal [Protocol]	2	29 May 2016
Summary CV for Chief Investigator (CI) [CV for Chief Investigator]	1	01 October 2015
Summary CV for student [CV Trupesh Patel]	1	01 June 2016
Summary CV for supervisor (student research) [CV Joseph Alsousou]	1	01 June 2016
Validated questionnaire [Oxford Hip Score]	1	29 April 2016
Validated questionnaire [Oxford Knee Score]	1	29 May 2016
Validated questionnaire [SF-36]	1	29 May 2016
Validated questionnaire [WOMAC]	1	29 May 2016

Appendix B - Summary of HRA Assessment

This appendix provides assurance to you, the sponsor and the NHS in England that the study, as reviewed for HRA Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England to assist in assessing and arranging capacity and capability.

For information on how the sponsor should be working with participating NHS organisations in England, please refer to the, *participating NHS organisations, capacity and capability* and *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* sections in this appendix.

The following person is the sponsor contact for the purpose of addressing participating organisation questions relating to the study:

Mr Alex Astor (01517948739; sponsor@liverpool.ac.uk)
Mr Trupesh Patel (07794763419; t.patel@liv.ac.uk)

HRA assessment criteria

Section	HRA Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	Researcher confirmed the study is not applying for portfolio adoption.
2.1	Participant information/consent documents and consent process	Yes	Minor changes have been made to Consent Form 2 – Expert participants after REC favourable opinion to bring it in line with HRA standards.
3.1	Protocol assessment	Yes	No comments
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	The statement of activities will act as the agreement between the sponsor and the site.
4.2	Insurance/indemnity arrangements assessed	Yes	Where applicable, independent contractors (e.g. General Practitioners) should ensure that the professional indemnity provided by their medical defence organisation covers the activities expected of them for this

Section	HRA Assessment Criteria	Compliant with Standards	Comments
			research study
4.3	Financial arrangements assessed	Yes	As detailed in the Statement of Activities no funding will be provided to sites.
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	No comments
5.3	Compliance with any applicable laws or regulations	Yes	No comments
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	No comments
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	No comments
6.3	Devices – MHRA notice of no objection received	Not Applicable	No comments
6.4	Other regulatory approvals and authorisations received	Not Applicable	No comments

Participating NHS Organisations in England

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.
There is one site in this study. All study activities as detailed in the Schedule of Events will occur at site. The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the research management function at the participating organisation. For NIHR CRN Portfolio studies, the Local LCRN contact should also be copied into this correspondence. For further guidance on working with participating NHS organisations please see the HRA website.

If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England which are not provided in IRAS or on the HRA website, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net. The HRA will work with these organisations to achieve a consistent approach to information provision.

Confirmation of Capacity and Capability

This describes whether formal confirmation of capacity and capability is expected from participating NHS organisations in England.

Participating NHS organisations in England **will be expected to formally confirm their capacity and capability to host this research**.

- Following issue of this letter, participating NHS organisations in England may now confirm to the sponsor their capacity and capability to host this research, when ready to do so. How capacity and capacity will be confirmed is detailed in the *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* section of this appendix.
- The Assessing, Arranging, and Confirming document on the HRA website provides further information for the sponsor and NHS organisations on assessing, arranging and confirming capacity and capability.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and the minimum expectations for education, training and experience that PIs should meet (where applicable).

PI will be required at site.

Sponsor expects those working on the study to have undergone GCP training.

GCP training is not a generic training expectation, in line with the HRA statement on training expectations.

HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken

The student will have access to site as part of their degree training therefore no letters of access will be required. All others members of the research team will be local staff who have a contractual relationship with the organisation. Therefore no honorary research contracts or letters of access are expected for this study.

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in

England to aid study set-up.
The applicant has indicated that they do not intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix J: GCP Certificate

NHS
National Institute for Health Research
Clinical Research Network

CERTIFICATE of ACHIEVEMENT

This is to certify that

Mohammad Haji Yakob

has completed the course

Introduction to Good Clinical Practice eLearning (Secondary Care)

May 30, 2017

Modules completed:

Introduction to Research in the NHS
Good Clinical Practice and Standards in Research
Study Set Up and Responsibilities
The Process of Informed Consent
Data Collection and Documentation
Safety Reporting

This course is worth 4 CPD credits

CPD
The CPD Certification Service

National Institute for Health Research
Clinical Research Network

Appendix K: Interview Guide – Stage 1

Development of PROM interview guide – First Stage

1. Introduce self and explain purpose of interview.
2. Establish understanding from patient and obtain verbal consent.
3. Obtain written consent (complete study consent form).
4. Collect basic information using data collection sheet.
5. Semi-structured interview
    a. Ask about pain symptoms
        i. Nature of pain
        ii. Location
        iii. Frequency and duration
    b. Ask about previous and current treatments for osteoarthritis
    c. Ask about the impact of osteoarthritis on patient's life
    d. Prompts to be used to direct interview
        i. Walking
        ii. Ability to use stairs
        iii. Change in body position
        iv. Maintaining certain body positions
        v. Use of upper limbs – fine and gross movements
        vi. Daily activities
        vii. Bathing oneself
        viii. Dressing oneself
        ix. Preparing meals
        x. Eating meals
        xi. Housework
        xii. Shopping
        xiii. Using public and own transportation
        xiv. Social interactions – including public/private/intimate
        xv. Engaging in social activities
        xvi. Aspects around job/employment
        xvii. Recreational activities – indoor and outdoor
    e. Establish if there are other factors limiting patient's overall functional ability.
    f. Ask about patient's perception of their general health and functional ability.
6. Thank the patient.

Appendix L: Qualitative information gathering sheet

IRAS: 207639
Study: DevPROM
Version 2: 15/7/16

UNIVERSITY OF LIVERPOOL

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

Date Signature

Baseline Assessment – Patient Participants

Participant number	
Age	
Sex	
Weight	
Height	
OA diagnosis?	
Duration of Symptoms	
Physical Activity Level (hours)	
Past Medical History	
Past treatment for OA.	

When completed: 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

IRAS: 207639
Study: DevPROM
Version 2: 15/7/16

Data collection sheet.

Participant number		Page

When completed: 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

Appendix M: Questionnaire for expert panel

UNIVERSITY OF LIVERPOOL

Development of PROMs Study

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

Expert participant questionnaire 1

Participant Name:

Lower Limb Function

For each of the following activities please provide one response to indicate how strongly you feel about its importance in assessing lower limb physical function. (click the box)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree	N/A
Getting In and out of bed	☐	☐	☐	☐	☐	☐	☐
Getting In and out of bath	☐	☐	☐	☐	☐	☐	☐
Getting In and out of shower	☐	☐	☐	☐	☐	☐	☐
Getting On and off toilet seat	☐	☐	☐	☐	☐	☐	☐
Dressing: Pulling up Trousers	☐	☐	☐	☐	☐	☐	☐
Dressing: Putting on Socks	☐	☐	☐	☐	☐	☐	☐
Dressing: Putting on Shoes	☐	☐	☐	☐	☐	☐	☐
Climbing up one flight of stairs	☐	☐	☐	☐	☐	☐	☐
Going down one flight of stairs	☐	☐	☐	☐	☐	☐	☐
Going up and down one flight of stairs	☐	☐	☐	☐	☐	☐	☐
Get in and out of car	☐	☐	☐	☐	☐	☐	☐
Get on and off bus	☐	☐	☐	☐	☐	☐	☐
Walking on flat surface	☐	☐	☐	☐	☐	☐	☐
Stand up from sitting position	☐	☐	☐	☐	☐	☐	☐
Kneeling down and getting back up	☐	☐	☐	☐	☐	☐	☐
Maximum walking distance	☐	☐	☐	☐	☐	☐	☐
Walking on uneven surface	☐	☐	☐	☐	☐	☐	☐
Abnormal gait	☐	☐	☐	☐	☐	☐	☐
Walking with aids	☐	☐	☐	☐	☐	☐	☐
Walking up and down hills	☐	☐	☐	☐	☐	☐	☐
Squatting	☐	☐	☐	☐	☐	☐	☐
Standing on tip toes	☐	☐	☐	☐	☐	☐	☐
Standing for a period of time	☐	☐	☐	☐	☐	☐	☐

Comments:

Upper Limb Function

For each of the following activities please provide one response to indicate how strongly you feel about its importance in assessing upper limb physical function. (click the box)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree	N/A
Brushing teeth	☐	☐	☐	☐	☐	☐	☐
Dressing: Doing up Buttons	☐	☐	☐	☐	☐	☐	☐
Grooming (comb hair) & Make-up	☐	☐	☐	☐	☐	☐	☐
Prepare a meal	☐	☐	☐	☐	☐	☐	☐
Carry a tray	☐	☐	☐	☐	☐	☐	☐
Turning a key	☐	☐	☐	☐	☐	☐	☐
Carry shopping bags	☐	☐	☐	☐	☐	☐	☐
Opening a tight jar	☐	☐	☐	☐	☐	☐	☐
Reaching for object above head	☐	☐	☐	☐	☐	☐	☐
Doing Dishes (manually or dishwasher)	☐	☐	☐	☐	☐	☐	☐
Opening a door	☐	☐	☐	☐	☐	☐	☐
Reach behind back	☐	☐	☐	☐	☐	☐	☐
Using keyboard / computer	☐	☐	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐	☐	☐
Picking up small things (e.g. coins, keys etc)	☐	☐	☐	☐	☐	☐	☐
Lifting things	☐	☐	☐	☐	☐	☐	☐
Washing self all over	☐	☐	☐	☐	☐	☐	☐
Using knife and fork	☐	☐	☐	☐	☐	☐	☐
Hang clothes	☐	☐	☐	☐	☐	☐	☐

Comments:

Usual Activities

For each of the following activities please provide one response to indicate how strongly you feel about its importance in assessing a patient usual activities. (click the box)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree	N/A
Cleaning around the house	☐	☐	☐	☐	☐	☐	☐
Light recreational activities (e.g. knitting, reading)	☐	☐	☐	☐	☐	☐	☐
Heavy recreational activities (e.g. gardening)	☐	☐	☐	☐	☐	☐	☐
Ability to do Usual Work or daily job	☐	☐	☐	☐	☐	☐	☐
Socialising with friends and family	☐	☐	☐	☐	☐	☐	☐
Light Sporting activities or exercise (e.g walking, golf)	☐	☐	☐	☐	☐	☐	☐
Heavy Sporting activities or exercise (e.g Tennis)	☐	☐	☐	☐	☐	☐	☐
Engaging in sexual activities	☐	☐	☐	☐	☐	☐	☐

Comments:

General Health

For each of the following activities please provide one response to indicate how strongly you feel about its importance in assessing a patient's general health. (click the box)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree	N/A
Level of General physical health	☐	☐	☐	☐	☐	☐	☐
General health affecting lower limb function	☐	☐	☐	☐	☐	☐	☐
General health affection upper limb function	☐	☐	☐	☐	☐	☐	☐
Level of Emotional wellbeing	☐	☐	☐	☐	☐	☐	☐
Level of Anxiety	☐	☐	☐	☐	☐	☐	☐
Level of Depression	☐	☐	☐	☐	☐	☐	☐
Feeling self-conscious	☐	☐	☐	☐	☐	☐	☐
Pre-occupied by symptoms	☐	☐	☐	☐	☐	☐	☐

Comments:

Pain / Symptoms

For each of the following activities please provide one response to indicate how strongly you feel about its importance in assessing patient arthritic pain and symptoms. (click the box)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree	N/A
Pain Severity	☐	☐	☐	☐	☐	☐	☐
Pain Frequency	☐	☐	☐	☐	☐	☐	☐
Pain Limiting Lower limb function	☐	☐	☐	☐	☐	☐	☐
Pain Limiting Upper limb function	☐	☐	☐	☐	☐	☐	☐
Awareness of pain	☐	☐	☐	☐	☐	☐	☐
Pain symptoms disturbing sleep	☐	☐	☐	☐	☐	☐	☐
Pain during activity	☐	☐	☐	☐	☐	☐	☐
Pain at rest	☐	☐	☐	☐	☐	☐	☐
Stopping activity because of pain	☐	☐	☐	☐	☐	☐	☐
Pre-occupied by pain	☐	☐	☐	☐	☐	☐	☐
Worse pain vs Usual pain	☐	☐	☐	☐	☐	☐	☐
Joint stiffness	☐	☐	☐	☐	☐	☐	☐

Comments:

Appendix N: Interview Guide – Stage 3

Qualitative pilot interview guide – Third Stage

1. Introduce self and explain purpose of interview.
2. Establish understanding from patient and obtain verbal consent.
3. Obtain written consent (complete study consent form).
4. Collect basic information using data collection sheet.
5. Ask patient to complete the prototype PROM
6. Once completed conduct interview
    a. Questions to ask:
        i. How long did it take to complete the PROM?
        ii. Was easy to understand and complete?
        iii. Were the instructions clear?
        iv. Were the questions relevant?
        v. Were there any questions that should have been included?
        vi. Were there any questions that were unnecessary?
7. Present the patient with other formatting styles of the PROM and ask them to select their preferred style.
8. Thank the patient

Appendix O: Most frequent themes identified from stage 1

Item description	Frequency from qualitative analysis, N(%)
Walking	41(82)
Using stairs (Up or Down)	24(48)
Getting in and out of cars	19(38)
Getting in and out of a bathtub	16(32)
Doing housework	16(32)
Standing up from sitting position	14(28)
Pain	14(28)
Doing shopping	12(24)
Putting on shoes	11(22)
Getting dressed	9(18)
Getting in and out of bed	8(16)
Showering	7(14)
Socialising with friends and relatives	7(14)
Kneeling	7(14)
Feeling breathless	6(12)
Cold weather affecting pain	6(12)
Gardening	6(12)
Playing sports	5(10)
Carrying things	5(10)
Anxiety	5(10)

Appendix P: Summary results of expert panel questionnaire

Item	Domain	n	Median	Mean	sds	95	%	CI	Concensus?
Getting In and out of bed (5.14)	Lower limb function	10	5.00	5.20	0.63	4.81	-	5.59	Yes
Getting In and out of bath (4.57)	Lower limb function	10	5.00	5.00	0.67	4.59	-	5.41	Yes
Getting In and out of shower (5.07)	Lower limb function	10	5.00	5.30	0.67	4.88	-	5.72	Yes
Dressing: Pulling up Trousers (4.93)	Lower limb function	10	5.00	4.50	1.08	3.83	-	5.17	No
Dressing: Putting on footwear (new)	Lower limb function	10	6.00	5.70	0.48	5.40	-	6.00	Yes
Climbing up one flight of stairs (5.50)	Lower limb function	10	5.00	5.20	0.92	4.63	-	5.77	Yes
Going down one flight of stairs (5.43)	Lower limb function	10	5.00	5.20	0.92	4.63	-	5.77	Yes
Going up and down one flight of stairs (4.93)	Lower limb function	10	5.00	4.40	1.43	3.51	-	5.29	Yes
Get in and out of car (5.29)	Lower limb function	10	5.00	5.10	0.57	4.75	-	5.45	Yes
Get on and off bus (4.57)	Lower limb function	10	5.00	4.90	1.10	4.22	-	5.58	Yes
Walking on flat surface (5.57)	Lower limb function	10	6.00	5.60	0.70	5.17	-	6.03	Yes
Stand up from sitting position (5.86)	Lower limb function	9	6.00	5.89	0.33	5.67	-	6.11	Yes
Kneeling down and getting back up (4.86)	Lower limb function	10	4.50	4.50	0.85	3.97	-	5.03	Yes
Maximum walking distance (5.21)	Lower limb function	10	6.00	5.60	0.52	5.28	-	5.92	Yes
Walking on uneven surface (5.07)	Lower limb function	10	5.00	5.00	0.47	4.71	-	5.29	Yes
Walking with aids (5.50)	Lower limb function	10	5.00	5.20	0.92	4.63	-	5.77	Yes
Walking up and down hills (4.57)	Lower limb function	10	4.00	4.20	0.79	3.71	-	4.69	Yes
Squatting (4.71)	Lower limb function	10	4.00	4.10	0.99	3.48	-	4.72	No
Standing for a period of time (5.43)	Lower limb function	10	5.00	5.40	0.52	5.08	-	5.72	Yes
Brushing teeth (5.09)	Upper limb function	9	6.00	5.56	0.53	5.21	-	5.90	Yes
Dressing: Doing up Buttons (5.75)	Upper limb function	9	6.00	5.67	0.50	5.34	-	5.99	Yes
Grooming (comb hair) & Make-up (5.50)	Upper limb function	9	6.00	5.56	0.53	5.21	-	5.90	Yes
Prepare a meal (5.75)	Upper limb function	9	6.00	5.56	0.53	5.21	-	5.90	Yes
Carry a tray (5.25)	Upper limb function	9	5.00	5.33	0.50	5.01	-	5.66	Yes
Turning a key (5.25)	Upper limb function	9	6.00	5.56	0.53	5.21	-	5.90	Yes
Carry shopping bags (5.42)	Upper limb function	9	5.00	5.22	0.83	4.68	-	5.77	Yes
Opening a tight jar (5.17)	Upper limb function	9	5.00	5.00	0.87	4.43	-	5.57	Yes
Reaching for object above head (5.75)	Upper limb function	9	6.00	5.44	0.88	4.87	-	6.02	Yes
Doing Dishes (manually or dishwasher) (5.17)	Upper limb function	9	5.00	5.33	0.71	4.87	-	5.80	Yes
Opening a door (5.50)	Upper limb function	9	6.00	5.56	0.73	5.08	-	6.03	Yes
Reach behind back (5.58)	Upper limb function	9	5.00	5.22	0.44	4.93	-	5.51	Yes
Using keyboard / computer (4.92)	Upper limb function	8	5.50	5.50	0.53	5.13	-	5.87	Yes
Writing (5.58)	Upper limb function	9	6.00	5.33	1.00	4.68	-	5.99	Yes
Picking up small things (e.g. coins, keys etc) (5.67)	Upper limb function	9	5.00	5.44	0.53	5.10	-	5.79	Yes
Lifting things (5.17)	Upper limb function	9	5.00	5.22	0.67	4.79	-	5.66	Yes
Washing self all over (5.67)	Upper limb function	9	6.00	5.67	0.50	5.34	-	5.99	Yes
Using knife and fork (5.75)	Upper limb function	9	6.00	5.67	0.50	5.34	-	5.99	Yes
Hang clothes (5.33)	Upper limb function	9	5.00	5.11	0.60	4.72	-	5.50	Yes
Cleaning around the house (5.07)	Role restriction	10	5.00	5.40	0.52	5.08	-	5.72	Yes
Recreational activities (new)	Role restriction	10	5.00	5.20	0.63	4.81	-	5.59	Yes
Ability to do work or daily job (5.57)	Role restriction	10	5.50	5.50	0.53	5.17	-	5.83	Yes
Socialising with friends and family (4.86)	Role restriction	10	5.00	5.00	0.67	4.59	-	5.41	Yes
Sporting activities or exercise (new)	Role restriction	10	5.00	5.30	0.48	5.00	-	5.60	Yes
Engaging in sexual activities (4.64)	Role restriction	10	5.00	4.90	0.57	4.55	-	5.25	Yes

Level of General physical health (5.64)	General health	10	6.00	5.80	0.42	5.54	-	6.06	Yes
General health affecting lower limb function (5.36)	General health	10	6.00	5.70	0.48	5.40	-	6.00	Yes
General health affection upper limb function (5.36)	General health	10	6.00	5.70	0.48	5.40	-	6.00	Yes
Level of Emotional wellbeing (5.57)	General health	10	5.50	5.50	0.53	5.17	-	5.83	Yes
Level of Anxiety (5.64)	General health	10	6.00	5.70	0.48	5.40	-	6.00	Yes
Level of Depression (5.57)	General health	10	6.00	5.50	0.71	5.06	-	5.94	Yes
Feeling self-conscious (5.07)	General health	10	5.00	5.20	0.63	4.81	-	5.59	Yes
Pre-occupied by symptoms (5.21)	General health	10	5.00	5.40	0.52	5.08	-	5.72	Yes
Pain Severity (5.93)	Pain and symptoms	10	6.00	5.90	0.32	5.70	-	6.10	Yes
Pain Frequency (5.85)	Pain and symptoms	10	6.00	5.90	0.32	5.70	-	6.10	Yes
Pain Limiting Lower limb function (5.79)	Pain and symptoms	10	6.00	5.90	0.32	5.70	-	6.10	Yes
Pain Limiting Upper limb function (5.79)	Pain and symptoms	10	6.00	5.90	0.32	5.70	-	6.10	Yes
Awareness of pain (5.15)	Pain and symptoms	10	5.00	5.40	0.52	5.08	-	5.72	Yes
Pain symptoms disturbing sleep (5.79)	Pain and symptoms	10	6.00	5.60	0.84	5.08	-	6.12	Yes
Pain during activity (5.71)	Pain and symptoms	10	6.00	5.70	0.67	5.28	-	6.12	Yes
Pain at rest (5.93)	Pain and symptoms	10	6.00	5.90	0.32	5.70	-	6.10	Yes
Stopping activity because of pain (5.93)	Pain and symptoms	10	6.00	5.70	0.48	5.40	-	6.00	Yes
Pre-occupied by pain (5.07)	Pain and symptoms	10	5.00	5.40	0.52	5.08	-	5.72	Yes
Worse pain vs Usual pain (5.00)	Pain and symptoms	10	5.00	5.20	0.79	4.71	-	5.69	Yes
Joint stiffness (5.50)	Pain and symptoms	10	5.50	5.50	0.53	5.17	-	5.83	Yes

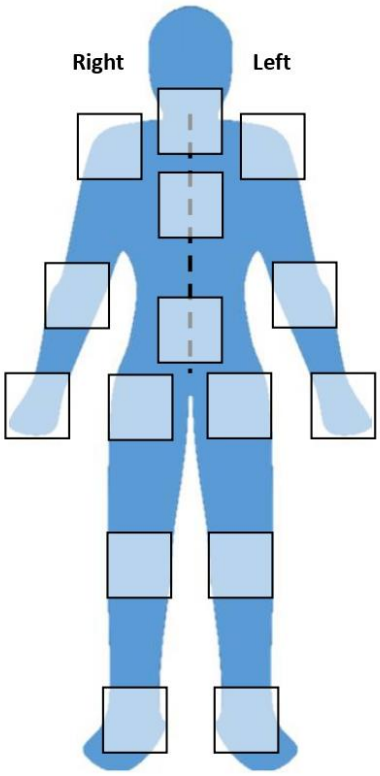
Appendix Q: Final version of new PROM

New PROM

Patient Name:		Date of Birth:		Today's Date:		
For the following items please select the response that best describes your **level of function** on average over the **last month** (*For each item please tick one box per row*)						
		Able without problems	**Able but a little difficult**	**Able but moderately difficult**	**Able but very difficult**	**Unable**
When I Need to:						
01	Stand up from a chair	☐	☐	☐	☐	☐
02	Put on footwear	☐	☐	☐	☐	☐
03	Get in and out of a car	☐	☐	☐	☐	☐
04	Walk for 10 minutes	☐	☐	☐	☐	☐
05	Go up a flight of stairs	☐	☐	☐	☐	☐
06	Go down a flight of stairs	☐	☐	☐	☐	☐
07	Carry things (e.g. shopping bag)	☐	☐	☐	☐	☐
08	Do up buttons	☐	☐	☐	☐	☐
09	Reach out for something at shoulder height	☐	☐	☐	☐	☐
10	Turn a key	☐	☐	☐	☐	☐
11	Prepare a meal	☐	☐	☐	☐	☐
12	Do my regular job or daily routine if retired	☐	☐	☐	☐	☐
13	Perform leisure or sporting activities (e.g. Dancing, bowling, gardening)	☐	☐	☐	☐	☐
14	Do Housework	☐	☐	☐	☐	☐
15	Go shopping on my own	☐	☐	☐	☐	☐
Please select the response that best describes your **level of pain while resting** (on average last month)						
		No pain	**Little pain**	**Moderate pain**	**Severe pain**	**Constant pain**
16	Level of pain	☐	☐	☐	☐	☐
Please select how **pain** from your joints limits your **overall function**						
		No Limitation	**Little limitation**	**Moderate limitation**	**A lot of limitation**	**Completely limited**
When I Need to:						
17	Use my legs	☐	☐	☐	☐	☐
18	Use my arms	☐	☐	☐	☐	☐
		No Limitation	**Little limitation**	**Moderate limitation**	**A lot of limitation**	**Completely limited**
19	How does your **general medical health (e.g. asthma, heart disease)** limit your overall function?	☐	☐	☐	☐	☐
20	How does your **mood (e.g. anxiety, depression)** limit your overall function?	☐	☐	☐	☐	☐

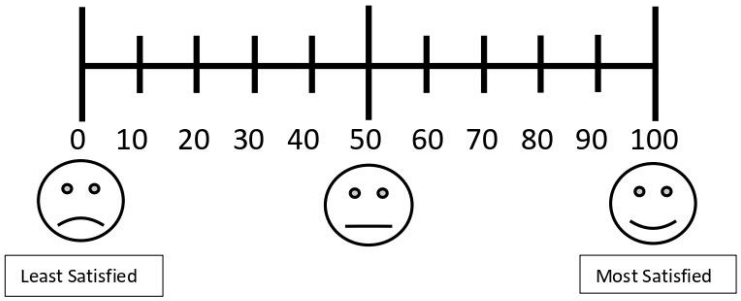
Question 21

Using a scale of 0 to 10 *(0 – no pain, 10 – worse pain),* rate the **SEVERITY of pain in your joints** by filling in the **boxes** on the picture below. **Leave unaffected joints blank**. (E.g. left knee [8], right shoulder [4])



Question 22

Using the following scale please mark with an '**X**' on the **scale line** to show how **satisfied** you are with your **overall level of function** on average over the **last month**. (0 – Least satisfied, 100 – Most satisfied)



References

Altman, R. et al., 1986. Development of creiteria for the classification and reporting osteoarthritis. Classification of oesteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum,* Volume 29, pp. 1039-49.

Anthoine, E. et al., 2014. Sample size used to validate a scale: a review of publications on newly-developed pateint reported outcome measures. *Health quality of life outcomes,* Volume 12, p. 176.

Arthritis research UK, 2019. *Versusarthritis.org.* [Online]
Available at: https://www.versusarthritis.org/
[Accessed December 2019].

Behrend, H., Giesinger, K., Giesinger, J. M. & Kuster, M. S., 2012. The "Forgotten Joint" as the Ultimate Goal in Joint Arthroplasty - Validation of a New Patient-Reported Outcome MEasure. *The Journal of Arthroplasty,* 27(3), pp. 430-436.

Bellamy, N. et al., 1988. Validation study of WOMAC: health status instrument for emasruing cinically important pateint relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol,* Volume 15, pp. 1833-40.

Bellamy, N. et al., 1988. Validation study of WOMAC: a health status instrumen for measuring clincial important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. *J Orthop Rheumatol,* Volume 1, pp. 95-108.

Bishop, P. A. & Herron, R. L., 2015. Use and Misuse of the Likert Item Responses and Other Ordinal Measures. *International Journal of Exercise Science,* 8(3), pp. 297-302.

Bolarinwa, O., 2015. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Niger Postgrad Med J,* Volume 22, pp. 195-201.

Brazier, J. et al., 1992. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ,* Volume 305, pp. 160-164.

Brummet, C. M. et al., 2016. Preliminary Validation of the Michigan Body Map (MBM). *Pain,* Volume 157(6), pp. 1205-12.

Burkhart, S., Morgan, C. & Kibler, W., 2003. The disabled throwing shoulder: Spectrum of pathology Part 1111 The SICK scapula, scapular dyskinesis, the kinetic chain and rehabilitation.. *Arthroscopy,* Volume 19, pp. 641-61.

Butler, P. B. & Major, R. E., 2003. The missing link? Therapy issues of open and closed chains.. *Physiotherapy,* Volume 89, pp. 465-70.

Cho, E., 2016. Making Reliability Reliable. *Organizational Research Methods,* 19(4), pp. 651-682.

Clarke, S., Lock, V. & Duffy, J., 2005. Intra-articular hylan G-F 20 (Synvisc) in the management of patellofemoral osteoarthritis of the knee (POAK). *Knee,* Volume 12, pp. 57-62.

Clynes, M. A. et al., 2019. Impact of osteoarthritis on activities of daily living: does joint site matter?. *Aging Clinical and Experimental Research,* Volume 31, pp. 1049-1056.

Cordero-Ampuero, J. et al., 2012. Evaluation of patients' and physicians' expectations and attributes of osteoarthritis treatment using Kano methodology. *Qual Life res,* Volume 21, pp. 1391-1404.

Creswell, J. W., 2009. *Research Design.* 3rd ed. s.l.:SAGE.

Dawson, J. et al., 2010. The routine use of patient reported outcome measures in healthcare settings. *The BMJ,* Volume 340:c186.

Dawson, J., Fitzpatrick, R., Carr, A. & Murray, D., 1996. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg [Br],* Volume 78, pp. 185-90.

Dawson, J., Fitzpatrick, R., Murray, D. & Carr, A., 1998. Questionnaire on the preceptions of patients about total knee replacement. *J Bone Joint Surg [Br],* Volume 80, pp. 63-9.

Dillman, C., Murray, T. & Hintermeister, R., 1994. Biomechanical differences of open and closed chain exercises with respect of proprioception, neuromuscular contraol and sunsequesnt functional stability of the joint.. *J Sport Rehad,* Volume 3, pp. 228-38.

Dunbar, M., Robertsson, O., Ryd, L. & Lidgren, L., 2000. Translation and validaton of the Oxford-12 item knee score for use in Sweden. *Act Orthopaedica,* 71(3), pp. 268-274.

Felson, D. et al., 1987. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum,* Volume 30, pp. 914-918.

Fitzpatrick, R., Davey, C., Buxton, M. & Jones, D., 1998. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess,* Volume 2, pp. 1-74.

Fransen, M. et al., 2015. Exercise for osteoarthritis of the knee:a Cochrane systematic review. *Br J Sports Med,* Volume 49, pp. 1554-57.

Gagnier, J. et al., 2017. A systematic review of measurement properties of patient-reported outcome measures used in patients undergoing total knee arthroplasty. *Journal of Arthroplasty,* Volume 32, pp. 1688-97.

Gardner, P. L., 1995. Measuring attitudes to science: unidimensionality and internal consistency revisited.. *Research in Science Education,* 25(3), p. 283–289.

Global burden of disease study, 2016. GLobal, regional and national incidence, prevalence and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the GLobal BUrden of Diesase Study 2015. *Lancet,* Volume 388, pp. 1545-602.

Grossoehme, D. H., 2014. Research Methodology Overview of Qualitative Research. *J Health Care Chaplain,* 20(3), pp. 109-22.

Gupta, S. et al., 2005. The economic burden of disabling hip and knee osteoarthrits (OA) from the perspective of individuals living with this condition. *Rheumatology,* Volume 44, pp. 1531-37.

Hadler, N., 1992. Knee pain is the malady - not osteoarthritis. *Ann Intern Med,* Volume 116, pp. 598-9.

Hannan, M., Felson, D. & Pincus, T., 2000. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *J Rheumatol,* Volume 27, pp. 1513-7.

Harris, K. et al., 2016. Systematic review of measurement properties of patient-reported outcome measures used in patients undergoing hip and knee arthroplasty. *Patient Related Outcome Measures,* Volume 7, pp. 101-108.

Herberts, P., Ahnfelt, L., Malchau, H. & Andersson, B., 1989. Multicenter clinical trials and their value in assessing total joint arthroplasty.. *Clinical Orthopaedics and Related Research,* Issue No. 249, pp. 48-55.

Hsia, A., AJ, E. & FD, T., 2018. Osteophytes and fracture calluses share developental milestones and are diminished by unloading. *J Orthop Res,* Volume 36, pp. 699-710.

Hunter, D. J. & Bierma-Zeinstra, S., 2019. Osteoarthritis. *Lancet,* Volume 393, pp. 1745-1759.

Hunter, D., Peitro-Alhambra, D. & Arden, N., 2014. *Osteoarthritis: the facts.* 2nd ed. Oxford: Oxford University Press.

Hunter, D., Schofield, D. & Callander, E., 2014. The individual and socioeconomic impact of osteoarthritis. *Nat Rev Rheumatol,* Volume 10, pp. 437-41.

Hutton, C., 1989. Osteoarthritis: the cause not result of joint failure?. *Ann Rheum Dis,* Volume 48, pp. 958-1.

Iglesias, C. & Torgerson, D., 2000. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. *Journal of Health Services Research and Policy,* 5(4), pp. 219-221.

Jenkinson, C. et al., 1997. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies?. *J of Public Health,* 19(2), pp. 179-186.

Karandikar, N. & Vargas, O. O. O., 2011. Kinetic Chains: A Review of the Concept and Its Clinical Applications. *American Academy of Physical Medicine and Rehabilitation,* Volume 3, pp. 739-45.

Kellgren, J. & Kawrence, J., 1963. *The Epidemiology of Chronic Rheumatism: Atlas of Standard Radiographs.* 1963 ed. Oxford: Blackwell Scientific Publications.

King, M. T., 2011. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Review of Pharmacoeconomics & Outcomes Research,* 11(2), pp. 171-184.

Koo, T. K. & Li, M. Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine,* Volume 15, pp. 155-163.

Leoser, R., Collins, J. & BO, D., 2016. Ageing and the pathogenesis of osteoarthritis.. *Nat Rev Rheumatol,* Volume 12, pp. 412-20.

Liang, Y. et al., 2014. Validity and reliability of questionnaires measuring physical activity self-efficacy, enjoyment, social support among Hong Kong Chinese children. *Preventive Medicin Reports,* Volume 1, pp. 48-52.

Lim, C. et al., 2015. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open,* p. 5:e007765.

Lim, J., Luscombe, K., Jones, P. & White, S., 2006. The effect of preoperative symptom severity on function outcome of total kne replacement: patients with the lowest preoperative scores achieve the lowest marks.. *Knee,* Volume 13, pp. 216-19.

Litwic, A., Edwards, M., Dennison, E. & Cyrus, C., 2013. Epidemiology and burden of osteoarthritis. *British Medical Bulletin,* Volume 105, pp. 185-99.

Maillot, C. et al., 2020. Moderate relationship between function and satisfaction of total hip arthroplasty patients: a cross sectional study. *Hip International,* p. Publised online.

Malterud, K., 2001. Qualitative research: standards, challenges, and guidelines. *The Lancet,* Volume 358, pp. 483-8.

Marquis, P., Fayol, C., Joire, J. & Leplege, A., 1995. Psychometric properties of a specific quality of life questionnaire in anggina pectoris patients. *Qual Life Res,* Volume 4, pp. 540-546.

Marshall, M., Watt, F., Vincent, T. & Dziedzic, K., 2018. Hand osteoarthritis: clinical phenotypes, molecular mechanism and disease management. *Nat Rev Rheumatol,* 14(11), pp. 641-656.

Martinelli, N. et al., 2011. Cross-cultural adaptation and validation with reliability, validity, and responsiveness of the Italian version of the Oxford Hip Score in patients with hip osteoarthritis. *Qaul Life Res,* 20(6), pp. 923-9.

McConnell, S., Kolopack, P. & Davis, A., 2001. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): A Review of Its Unitilty and Measurement Properties. *Arthrits care and res,* Volume 45, pp. 453-461.

McQuay, H., 2004. The measures of pain. In: P. Pynsent, J. Fairbank & A. Carr, eds. *Outcome measures in orthopaedics and orthopaedic trauma (2nd Ed).* Oxford: Arnold, pp. 63-69.

Messier, S., Loeser, R. & Miller, G., 2004. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. *Arthritis Rheum,* Volume 50, pp. 1501-10.

Meyer, K. et al., 1994. Monitoring dialysis pateints health status. *Am J Kidney Dis,* Volume 24, pp. 268-279.

Mokkink, L. et al., 2019. *COSMIN Study Design checklist for Patient-reported outcome measurement instruments.* [Online]
Available at: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf#
[Accessed October 2019].

Murray, D. W. et al., 2007. The use of the Oxford hip and knee scores. *J Bone Joint Surge [Br],* Volume 89, pp. 1010-14.

Nelson, A. et al., 2014. A systematic review of recommendations and guidelines for the management of osteoarthritis: the chronic osteoarthritis management initiative of the US bone and joint initiative.. *Semin Arthritis Rheum,* Volume 43, pp. 701-12.

Nelson, E. et al., 2015. Patient reported outcome measures in practice. *The BMJ,* Volume 350:g7818.

Neogi, T., 2013. The epidemiology and impact of pain in osteoarthritis. *Osteoarthritis and Cartilage,* Volume 21, pp. 1145-53.

NHS England, 2018. *PROMs guide.* [Online]
Available at: https://www.england.nhs.uk/wp-content/uploads/2018/08/proms-guide-aug-18-v3.pdf
[Accessed January 2020].

NHS UK, 2019. *Patient Reported Outcome Measures (PROMs).* [Online]
Available at: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms
[Accessed January 2020].

NICE, 2014. *Osteoarthritis: care and management, Clinical guideline [CG177].* [Online]
Available at: https://www.nice.org.uk/guidance/cg177
[Accessed January 2020].

Nilsdotter, A. & Bremander, A., 2011. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons. *Arthritis Care Res,* 63(11), pp. 200-7.

NJR Annual Report, 2019. NJR 16th Annual Report.

NJR, 2020. *NJR Centre.* [Online]
Available at: http://www.njrcentre.org.uk/njrcentre/
[Accessed January 2020].

Nyvang, J., Hedstrom, M. & Gleissman, S. A., 2016. It's not just a knee, but a whole life: A qualitative descriptive study on pateints' experiences of living with knee osteoarthritis and their expectations for knee arthroplasty. *Int J of Qaulitative studies on Health and Well-being,* Volume 11, p. 30193.

Patel, A. A., Donegan, D. & Albert, T., 2007. The 36-item Short Form. *J Am Acad Orthop Surgery,* Volume 15, pp. 126-135.

Paulsen, A., Odgaard, A. & Overgaard, S., 2012. Translation, cross-cultural adaptation and validation of the Danish version of the Oxford hip score: Assessed against generic and disease-specific questionnaires. *Bone Joint Res,* 1(9), pp. 225-33.

Polit, D. & Beck, C., 2006. The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health,* Volume 29, pp. 489-97.

Prieto-Alhambra, D. et al., 2014. Incidence and risk factors for clinically dianosed knee, hip and hand osteoarthritis: influence of age, gender and osteoarthritis affecting other joints.. *Ann Rhuem Dis,* Volume 73, pp. 1659-64.

Prince, M. et al., 2015. The burden of disease in older people and implications for health policy and practice. *Lancet,* Volume 385, pp. 549-562.

Pynsent, P., 2004. Choosing an outcome measure. In: P. Pynsent, J. Fairbank & A. Carr, eds. *Outcome measures in orthpaedics and orthopaedics trauma (2nd Ed).* Oxford: Arnold, pp. 1-7.

Rabin, R. & De Charro, F., 2001. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med,* Volume 33, pp. 337-43.

Ramkumar, P. N., Harris, J. D. & Noble, P. C., 2015. Patient-reported outcome measures after total arthroplasty. *Bone and joint research,* Volume 4, pp. 120 - 127.

RAND Corp., n.d. *36-Item Short Form Survey (SF-36).* [Online]
Available at: https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html
[Accessed December 2019].

RAND Corp, n.d. *36-Item Short Form Survey (SF-36) Scoring Instructions.* [Online]
Available at: https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form/scoring.html
[Accessed January 2020].

Reilly, K., Beard, D. & Barker, K., 2005. Efficacy of an accelerated recovery protocol for Oxford unicompartmental knee arthroplasty: a randomised controlled trial. *Knee,* Volume 12, pp. 351-7.

Renskers, L. et al., 2018. Comparison of the construct validity and reproducibility of four different types of patient-reported outcome measures (PROMs) in patients with rheumatoid arthritis. *Clinical Rheumatology,* Volume 37, pp. 3191-9.

Rolstad, S., Adler, J. & Ryden, A., 2011. Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value Health,* 14(8), pp. 1101-8.

Roos, E. M. & Toksvig-Larsen, S., 2003. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes,* Volume 17.

Roth, P. L., 1994. Missing Data: A Conceptual Review for Applied Psychologist. *Personnel Psychology,* 47(3), pp. 537-60.

Sijtsma, K., 2009. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika,* 74(1), p. 107–120.

Smith, M. V. et al., 2012. Upper Extremity-Specific Measures of Disability and Outcomes in Orthopaedic Surgery. *JBJS Am,* 94(3), pp. 277-85.

Stubbs, B., Aluko, Y., Myint, P. & Smith, T., 2016. Prevalence of depressive symptoms and anxiety in osteoarthritis: a systematic review and meta-analysis. *Age Ageing,* 45(2), pp. 228-35.

Taber, K. S., 2018. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res Sci Educ,* Volume 48, p. 1273–1296.

Tavakol, M. & Dennick, R., 2011. Making sense of Cronbach's alpha. *Int J Med Educ.,* Volume 2, pp. 53-55.

Theis, K., Murphy, L., Hootman, J. & Wilkie, R., 2013. Social praticipation restriction among US adults with arthritis:a population-based study using the International CLassifcation of Functioning, Disability and Health. *Arthritis Care Res,* Volume 65, pp. 1059-69.

Tuan, H. L., Chin, C. C. & Shieh, S. H., 2005. The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education,* 27(6), pp. 639-654.

van der Waal, J., Terwee, C. & van der Windt, D., 2005. Health-related and overall quality of life of patients with chronic hip and knee complaints in general practice. *Qual Life Res,* Volume 14, pp. 795-803.

Ware-JE, J. & Sherbourne, C., 1992. The MOS 36-item short-form health survery (SF-36). I. Conceptual framework and item selection. *Med Care,* Volume 30, pp. 473-83.

Ware, J. E., Kosinski, M. & Keller, S. D., 1996. A 12-item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Valildity. *Medical Care,* Volume 34, pp. 220-233.

Ware, J., Keller, S. & Kosinski, M., 1995. *SF-12: How to score the SF-12 physical and mental health summary scales.* Boston, MA: New England Medical Center.

Weale, A., Lee, A. & MacEachern, A., 2001. High tibial osteotomy using a dynamic axial external fixator. *Clin Orthop,* Volume 382, pp. 154-67.

Wilkie, R., Peat, G., Thomas, E. & Croft, P., 2007. Factors associated with restricted mobility outside the home in community-dwelling adults ages fifty years and older with knee pain: an example of use of the International Classification of Functioning to investigate participation restriction. *Arthritis Rheum,* Volume 57, pp. 1381-9.

Wolfe, F. & Kong, S., 1999. Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Ann Rheum Dis,* 58(9), pp. 563-568.

Wood, A. et al., 2013. A Review on the Management of Hip and Knee Osteoarthritis. *International Journal of Chronic Diseases,* 2013(ID845015).

World Health Organisation, 2020. *Health topics: Disabilities.* [Online]
Available at: https://www.who.int/topics/disabilities/en/

Wylde, V., Learmonth, I. & Cavendish, V., 2005. The Oxford hip score: the patient's perspective. *Health Qual Life Outcomes,* Volume 3, p. 66.

Yu, D., Peat, G., Bedson, J. & Jordan, K., 2015. Annual consultation incidence of osteoarthritis estimated from population-based health care data in England. *Rheumatology,* Volume 54, pp. 2051-2060.

Zhang, W. et al., 2010. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis,* Volume 69, pp. 483-89.

Zhang, Y. et al., 2002. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly. *Am J Epidemiol,* Volume 156, pp. 1021-7.