2020

Smallest Worthwhile Effect Values for Pain and Function after a Total Knee Replacement

Nancy Wofford
Virginia Commonwealth University

Acknowledgement

I would like to thank my family, especially my mom and dad, brothers, step dad, step mom, step sister, and grandparents, for the love and support that you have given me over this eight- year journey. I would not have achieved this milestone without your love and support. I especially want to thank my older brother, David, for the countless hours of mentorship and assistance you provided me. I am so thankful that you went first in this PhD journey!

I would also like to thank my friends, especially Haley, Sara, and Tricia, and fiancé, Kevin, who encouraged and supported me during this entire process. You pushed me to continue and never gave up on me.

I would also like thank Penny Swindell for the assistance with data collection and for allowing me to attend your preoperative education class for two years. This study would not have occurred without your help.

I would also like to thank my dissertation committee for their help and guidance during this process: Dan Riddle (Chair), Robert Perera, Ben Darter, Spencer Wheeler. I am especially grateful to Dan Riddle for his advice, time, and mentorship during the entire program.

Smallest Worthwhile Effect Values for Pain and Function after a Total Knee Replacement

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

By

Nancy Chadwick Henderson
Transitional Doctorate of Physical Therapy, Elon University, 2004
Masters of Physical Therapy, Elon University, 2002
Bachelor of Arts Degree, Longwood University, 1998

Dr. Dan Riddle, PT, PhD, Chair
Physical Therapy

Virginia Commonwealth University
Richmond Virginia
March, 2020

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

SMALLEST WORTHWHILE EFFECT VALUES FOR PAIN AND FUNCTION AFTER A TOTAL KNEE REPLACEMENT

Nancy Henderson, PT

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2020

Dr. Dan Riddle, PT, PhD, Physical Therapy

Meaningful change is an important concept in healthcare as it allows providers to better understand the efficacy of an intervention. Three commonly used methods to assess meaningful change include the distribution-based, anchor-based, and benefit-harm trade-off methods. The benefit-harm trade-off method improves upon the limitations of the distribution-based and anchor-based methods. Unlike the other two methods, the benefit-harm trade-off method allows patients to determine clinically meaningful change by assessing the benefits and risks of an intervention compared to no intervention. The smallest worthwhile effect is a response measure which is derived using the benefit-harm trade-off method. The primary aim of the current study was to determine smallest worthwhile effect values for pain and function following a Total Knee Replacement.

This study enrolled 121 participants. A standardized script was used to determine smallest worthwhile effect values for the KOOS Pain and KOOS function, daily living subscales. A feasibility study was initially conducted to assess feasibility of subject recruitment, locations for data collection, and to assess wording and interpretability of the standardized script. Baseline and

psychosocial variables were assessed for their influence on smallest worthwhile effect values. A 6-month follow-up, via telephone, was conducted to determine if participants achieved their smallest worthwhile effect values and to determine patient satisfaction and complication rates.

The results of this study demonstrated that 90% of participants needed at least a 57 point improvement in their KOOS Pain subscale scores, over their baseline pain, to feel that the surgery was worthwhile as compared to not having surgery. Similarly, 90% of participants needed at least a 51 point change on their KOOS Function, daily living subscale scores to view the surgery was worthwhile when compared to not having a knee replacement. Baseline scores on the KOOS Pain and Function, daily living subscales were the only significant predictors of smallest worthwhile effect estimates. Eighty-three of the 121 participants completed the 6-month follow-up. Of the six month follow-up participants, 91% were satisfied with their current state. There was a significant association between meeting or exceeding the smallest worthwhile effect estimates for the KOOS Pain and Function, daily living subscales and patient satisfaction.

The smallest worthwhile effect estimates generated in the current study are higher than response measures calculated using the distribution- and anchor-based methods. The higher estimates in the current study may ultimately mean that individuals require more benefit after surgery, as previously thought, when costs and risks are factored into determining meaningful change following a knee replacement. Those individuals with higher baseline pain and worse preoperative function required a greater improvement after surgery to feel that the surgery was worth the associated costs and risks. Surgeons and patients should factor baseline pain and function into their decision whether or not to undergo a knee replacement. The association between smallest worthwhile effect estimates and patient satisfaction further strengthens the use

of the smallest worthwhile effect estimate as an accurate measure with which to estimate meaningful change.

CHAPTER 1

Introduction

The most common surgical intervention for knee osteoarthritis (OA) is a total knee replacement (TKR). (Cram, Lu, Kates, Singh, Li & Wolf, 2012). TKR substantially reduces pain and improves functional status for most persons with symptomatic knee OA who undergo the procedure (Losina, Walensky, Kessler, Emrani, Reichmann, Wright, Holt….., et al., 2009). Several research studies found that an approximate 50% improvement in pain and function can be expected 3-6 months after a TKR (Jones, Beaupre, Johnston, & Suarez-Almazar, 2005; Bachmeier, March, Lapsley, Tribe, Courtenay & Brooks, 2001; Nillsdotter, Toksvig-Larsen, & Roos, 2009; Kennedy, Stratford, Riddle, Hanna, & Gollish, 2007; Judge, Arden, Cooper, Javaid, Carr, Field & Diepp, 2012; Davis, Perruccio, Ibrahim, Hogg-Johnson, Wong, Streiner, Beaton….., et al., 2011). To fully understand outcome following TKR, it is imperative to determine if the outcome is considered worthwhile to those patients undergoing the procedure. In order to accurately assess the extent to which outcomes are judged to be worthwhile, a person has to weigh the potential benefits of the surgery against the associated risks and costs.

Self-reported outcome measures are useful tools for assessing change in pain and function after a TKR. An outcome measure can demonstrate statistically significant change and/or clinically meaningful change. Determining if clinically meaningful change has occurred is critical as it provides important information related to the treatment's effectiveness. When defining a clinically meaningful change, the degree of change the patient perceives worthwhile in light of associated risks and costs of the intervention should be assessed.

Several measures of responsiveness of self-reported outcome measures exist and these indexes indicate the extent to which changes in outcome reflect meaningful changes to patients. Responsiveness measures were derived using a variety of methods. According to Ferreira et al.,

the majority of studies examining measures of responsiveness used either an anchor- or distribution-based method to derive estimates of the responsiveness of outcome measures (Ferreira, Herbert, Ferreira, Latimer, Ostelo, Nascimento & Smeets, 2012). Both methods have important limitations which diminish confidence in the estimates. The distribution-based method assesses magnitude of change, but does not address whether this change is clinically meaningful. The anchor-based method aims to establish clinically meaningful change, but uses a self-report measure, most commonly the Global Rating of Change (GROC) as an external criterion (Norman, Sridhar, Guyatt, & Walter, 2001). Retrospective measures, such as the GROC, have several limitations which diminish their validity. Specifically, the GROC is a single-item instrument that is likely not comprehensive enough to capture meaningful change. Secondly, the GROC has an arbitrary cut-point established by clinicians, not patients, that has been established to be indicative of meaningful change (Kamper, Maher, & MacKay, 2009; Schmitt. & Abbott, 2014). Additionally, research has shown that the GROC is subject to extensive recall bias (Norman, Stratford, & Regehr, 1997; Schmitt & Di Fabio, 2005). Aside from the use of the GROC as an external criterion, another limitation of the anchor-based approach is that it only accounts for the perceived benefits of an intervention, but without weighing them against the associated costs and risks.

In response to the limitations of both the distribution and anchor-based methods, Barrett et al. described a novel method, called the benefit-harm trade-off method, to estimate the magnitude of change necessary for patients to consider as meaningful (Barrett, Brown, Mundt & Brown, 2005). This method, used to develop both the sufficiently important difference (SID) and the smallest worthwhile effect (SWE), addresses the limitations of previous methods. The benefit-harm trade-off method, unlike previous methods, allows for patients to make judgments

regarding care and to weigh the benefits of an intervention against the associated risks, costs, and inconveniences of the treatment (Ferreira, Herbert, Ferreira, Latimer, Ostelo, Grotle, & Barrett, 2013). The SID is defined as, "the smallest amount of patient-valued benefit that an intervention would require to justify the associated costs, risks, and other harms" (Barrett et al., 2005, p. 254). SWE, a closely related index of responsiveness, refers to the smallest amount of improvement that is deemed worthwhile by the patient (Ferreira et al., 2013). The SWE is a modification of the SID in that it asks patients to determine the cut-point for improvement which is meaningful to them. One of the most notable differences between the SID and SWE is that the SWE allows for the patient to estimate the magnitude of meaningful improvement with the treatment of interest relative to what could be expected with no treatment. When assessing meaningful change of an intervention, the patient must compare this change to what degree of positive or negative change that would occur without the intervention. Without comparing the change with and without the intervention, the true efficacy of the intervention is unknown.

The benefit-harm trade-off method has advantages over both distribution- and anchor-based methods for the assessment of responsiveness. By assessing SWE values, clinicians will have a better understanding of the amount of improvement patients see as worthwhile, which reflects on whether the intervention was effective. Therefore, the purpose of this study is to estimate SWE values for self-reported pain and function in patients undergoing TKR. Additionally, baseline and psychosocial variables will be assessed for their association with SWE estimates for both pain and function.

The study aims are as follows:

- Study aim #1: Determine the SWE for KOOS Pain subscale scores following a TKR.

- Study aim # 2: Determine the SWE for KOOS Function, daily living subscale scores following a TKR.
- Study aim #3: Determine whether initial SWE values for pain and function are congruent with pain and function scores six months after the TKR.
- Study aim #4: Determine if baseline scores on a variety of variables are associated with SWE estimates of KOOS Pain and KOOS Function, daily living scores. These variables are the following: baseline KOOS Pain and KOOS Function, daily living scores, income level, educational level, pain catastrophizing, depressive symptoms, anxiety, self-efficacy, and pre-surgical outcome expectations.

CHAPTER 2

Literature Review

Chapter two is organized as follows: the first section provides a description of a Total Knee Replacement (TKR); the second section provides an overview of outcome measures; the third section discusses the three derivation methods and their associated measures of responsiveness; the final section outlines the research to support the chosen baseline and psychosocial variables used in this study.

Total Knee Replacement

Osteoarthritis (OA) is a debilitating disease process that can lead to pain and dysfunction and affect quality of life (Jones, Beaupre, Johnston, & Suarez-Almazor, 2007). A TKR is a commonly performed surgical procedure used to treat knee OA when conservative measures fail. This surgical procedure is performed by an Orthopedic Surgeon and entails replacing the distal end of the femur and proximal end of the tibia with a prosthesis. The TKR is an effective intervention, with low mortality rates, to relieve pain and improve function (Jones et al., 2007). Potential complications include deep vein thrombosis, superficial infections, peripheral nerve damage, pulmonary embolism, and deep infections. A 2009 study examined the cost effectiveness of a TKR in the United States and found that it is a highly cost-effective surgery for the management of end-stage knee OA among Medicare-aged persons compared to non-operative management (Wang, Olson-Kellog, Shamliyan, Choi, Ramakrishnan, & Kane, 2012). According to Peer and Lane (2013), pain and regaining functional abilities are the most common reasons to undergo a TKR. In a study by Mahomed et al., 76% of subjects expected to have no pain after their TKR and 84% expected 90% or greater chance of complete success with their surgery. Seventy-five percent of the subjects expected a 10% or less chance of complication

from their surgery and 40% expected to have no limitations in their usual activities after their TKR. (Mahomed, Liang, Cook, Daltroy, Fortin, Fossel, & Katz, 2002).

With the rising incidence of knee OA, the number of TKR procedures is also rapidly increasing. According to Cram, Lu, Kates, Singh, Li, & Wolfe (2012), approximately 600,000 TKRs are performed every year and cost approximately $15,000 per procedure, producing an impact of $9 billion in healthcare dollars. The demand for TKRs is projected to grow to 3.48 million procedures per year by the year 2030 (Healy, Rana, & Iorio, 2010).

Outcome Measures

Given the increasing number of TKRs and the potential impact on quality of life, an accurate determination of the effectiveness of a TKR is paramount. According to Jacobs and Christensen (2009), an evaluation of outcomes after treatment of OA is necessary to completely understand the efficacy of the intervention. Historically, clinicians have used impairments to gauge improvement after a treatment. Improvements in range of motion or strength would inform clinicians as to whether the applied intervention was a success. Or, with TKRs, they were labeled as successful based on technical details, such as whether the prosthesis was surgically fixated appropriately (Jones et al., 2007).

More recently, clinicians are utilizing outcome measures to determine intervention effectiveness. Outcome measures were developed out of the necessity to have standardized assessment of patient status or progress. Prior to outcome measure development, assessment was often limited to labeling a patient as “improving,” or “discharged” (Partridge, 1982). Outcome measures come in many forms, including self-report questionnaires and performance measures.

The goal of many patients undergoing a TKR is to decrease pain and improve physical function (Stratford & Kenney, 2006). Outcome measures allow clinicians and patients to determine how much improvement in pain and function has been achieved by the surgery. After

a TKR, pain is customarily assessed using patient-report, or self-report, pain scales; whereas performance measures and self-report questionnaires are used to assess changes in physical function. Performance measures are tests which provide objective data regarding a patient's physical function. Examples used with patients after a TKR include the timed up and go test (TUG), stair climbing test (SCT) and the six-minute walk test (6MW) (Mizner, Petterson, Clements, Zenl, Irrgang, & Snyder-Mackeler, 2011). These tests yield quantitative data that is easy to track and measure and provide a picture of the patient's functional status. A disadvantage of using performance tests to judge the effectiveness of a TKR is that the healthcare professionals, not the patients, are determining whether the treatment was a success.

Self-report questionnaires are instruments that reflect patients' perceptions of their own health status (Carr, Hewlett, Huges, Mitchell, Ryan, Carr, & Kirwan, 2003; Peer et al., 2013). Health status encompasses many domains of health, including pain and physical function. They show whether patients' expectations, who have undergone a TKR, have been fulfilled; therefore, speaking to the efficacy of the TKR as an intervention. Mizner et al. states that self-report, or patient-report, measures, when compared to performance based measures, are less expensive, less time intensive, and reduce the number of patients lost to follow-up (Mizner et al., 2011). Per Peer et al., physical function is best characterized by the patients themselves versus the clinicians (Peer et al., 2013). Self-report outcome measures are an integral component to evaluating the efficacy of a TKR as they provide information, from a patient perspective, on the treatment's success.

Responsiveness, also called sensitivity to change, is one of the most important qualities of an outcome measure and relates to assessing outcome over time (Tuley, Mulrow & McMahan, 1991; Peer et al., 2013). Measures of responsiveness provide information regarding magnitude of

change after an intervention and can vary depending on who is determining the change. They ultimately ask the question whether the degree of change from an intervention is important. Responsiveness of an outcome measure is directly related to the amount of change in a person's score which is clinically meaningful (Guyatt, Walter & Norman, 1987). Ultimately, measures of responsiveness guide clinicians and patients in understanding magnitude of change, thus having a better understanding of the efficacy of an intervention. Secondly, they may have utility in interpreting research and informing sample size calculations in future clinical trials (Ferreira, Herbert, Ferreira, Latimer, & Ostelo, Grotle & Barrett, 2013).

Methods used to estimate the importance of change scores

There are three primary methods for establishing measures of responsiveness: distribution-based method, anchor-based method, and benefit-harm trade-off method.

Distribution-based method

The distribution-based method relates the difference in scores between a treatment and control group on an outcome measure to some form of variability (Norman, Sridhar, Guyatt, & Walter, 2001). It only provides information regarding minimum change on a self-report outcome measure, or a measurement tool, that is likely due to measurement error (Gatchel, Lurie, Mayer, & 2010). A significant limitation of this method is that it assesses magnitude of change, but does not address whether the change is clinically meaningful (Gatchel et al., 2010). It only indicates whether a given change exceeds the variability of a measurement instrument (Deyo & Patrick, 1995). Additionally, this method does not reflect a patient's perception of the magnitude of change and is not linked to an intervention. These estimates cannot tell us if the magnitude of effect in light of the benefits and/or costs and risks of an intervention (McNamara, Elkins, Ferreira, Spencer, & Herbert, 2015). This method is used to calculate several measures of

responsiveness, including effect size, standard error of the mean, Guyatt's responsiveness index, and reliability change index (RCI).

Effect size is one of the earliest response measures and relates to responsiveness at a group level (Schmitt & Di Fabio, 2004). It is used to evaluate change only and cannot assess change of different degrees (Stratford, Binkley, & Riddle, 1996). The effect size is an absolute measure of change and does not take into account variability of scores (Sullivan, G. & Fein, R., 2012). Effect size is beneficial as it allows for change to be converted to a standard unit of measurement which permits comparison among different outcome measures and interventions (Deyo et al., 1995). This standardized measure of change assesses change within a group or the difference in amount of change between groups (Kazis, Anderson, & Meenan, 1989). One mathematical equation used to calculate effect size is as follows: $ES= (m_1-m_2)/s_1-m_1$ where m_1= pre-treatment mean, m_2= post-treatment mean, and s_1= standard deviation (Kazis et al., 1989). These values are compliments of p-values, which provide a measure of statistical significance, but effect sizes provide information about the magnitude of change. These values are commonly reported as the following standardized effect sizes: .2-.3= small effect, .5= medium effect, .8 and higher= large effect. The disadvantage of effect sizes is that, even though effect sizes provide useful information as to the magnitude of change, these changes may not be clinically meaningful (Jacobson, Roberts, Berns, & McGlinchey, 1999). A large effect size is more likely to be clinically meaningful, but this is not guaranteed as even large effects can be clinically insignificant (Jacobson et al., 1999).

The standard error of the mean is also a commonly used measure derived with the distribution-based method (Schmitt & Di Fabio, 2004). Similar to an effect size, the standard error of the mean (SEM) evaluates change only and cannot assess change of different degrees

(Stratford & Kennedy, 1996). It is considered to be a fixed characteristic and should remain constant if repeated samples are drawn from the same population. The amount of error in this value is associated with an individual's, versus a group's, assessment. Mathematically, SEM= SD $(1\text{-}R)^{1/2}$, with SD= baseline standard deviation and R= test-retest reliability coefficient (Wyrwich, Nienabaer, Tierney, & Wolinsky, 1999). Similar to the SEM, the standardized response mean (SRM) is a mean change score of people who have improved on an outcome measure divided by the standard deviation of their change scores (Schmitt et al., 2004). Guyatt's responsiveness index is the mean change score of patients determined to have improved on an outcome measure, usually the Global Rating of Change (GROC), divided by the standard deviation of change scores among stable patients over a 3-month time frame (Schmitt et al., 2004). Lastly, the reliability change index is a response measure for outcome measures which are nondevelopmental in nature (Ottenbacher, Johnson, & Hojem, 1988). Mathematically, RCI= $(X_2\text{-}X_1)/S_E$, where X_2= post-test score, X_1= pre-test score and S_e is the standard error of measurement (Ottenbacher et al., 1988).

Anchor-based method

The anchor-based method assesses the relationship of an outcome measure to an external criterion, known as an anchor (Norman et al., 2001). This method varies from the distribution-based method as it provides change values which are deemed clinically meaningful. The GROC is a commonly used anchor or external criterion. The response measures in this section were developed using the GROC as the external criterion. The GROC is a 15-point Likert scale which asks patients to determine the amount of change they experienced using descriptors which range from worse (-7) to a better (+7) with a score of (0) indicating no change (Schmitt & Abbott, 2014). Since the patient chooses the amount of change on the GROC, the response measure estimates derived using this methodology appear to be derived from the patient; however, it is

the researcher who generally decides the threshold of meaningful change (McNamara, Elkins, Ferreira, Spencer, & Herbert, 2015).

One of the greatest limitations of the anchor based method is the use of a retrospective outcome measure as the external criterion by which to establish that meaningful change has occurred. The GROC, which is the most frequently used retrospective measure with this method, has several significant limitations. Similar to most retrospective measures, the GROC is a self-report measure which is prone to significant recall bias and is strongly correlated with current state versus baseline values, as well as, it is variable between different time periods (Norman, Stratford, & Regehr, 1997; Garrison & Cook, 2012; Schmitt & Di Fabio, 2005; Schmitt & Abbott, 2014; Biome, C. & Augustin, M., 2015). Norman et al. was one of the first studies to question the validity of retrospective estimates of change. The authors found that these estimates of change, such as the GROC, are invalid as they correlate more highly with a patient's current status versus their baseline status. They found that retrospective or transitional measures are strongly related to the health status at the time the patient completes the scale. These authors ultimately concluded that retrospective scales are biased and not valid measures of change over time (Norman et al., 1997).

Garrison and Cook assessed the consistency of GROC values over an eight-week time frame. They concluded that the GROC outcomes are not linear or progressive, but instead varied greatly over the eight weeks. A decline in GROC values over time was also noted and led the authors to conclude that the chance of identifying self-perceived improvement decreases over the treatment time frame (Garrison & Cook, 2012). A study by Schmitt and Di Fabio examined the validity of retrospective measures of change compared to more objective measures of change. The results of the study showed moderate correlation with the current status of the patient and

low correlation with the baseline status of the patient. The authors determined that retrospective global change measures are not accurate measures of change over time and that baseline patient status had little-no influence on the retrospective change values (Schmitt & Fabio, 2005). Schmitt and Abbott found similar results as the Schmitt and Fabio study. These authors found that GROC scores are unrelated to a patient's status at baseline, but are strongly related to a patient's status at discharge (Schmitt and Abbott, 2014).

Even though effect sizes and other related measures demonstrate magnitude of change, it is unknown whether this change is clinically meaningful. In response to the limitations of the distribution-based method, other response measures which directly relate to meaningful change were developed. The minimum clinically important difference (MCID) is one of the more common response measures which uses the anchor based approach to assess change that is clinically relevant. The MCID allows patients to determine the amount of change which is significant. Jaeschke, Singer, & Guyatt, (1989) defines this measure as, "the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troubling side effects and excessive cost, a change in the patient's management." To calculate MCID values, subjects are followed over time and asked to complete a patient-reported outcome measure (PROM) and the GROC after an intervention has been administered. The GROC is used as the external criterion to determine the cutoff score on the PROM which is clinically meaningful. An arbitrary cut-off point, determined by clinicians and researchers, serves as the measure by which meaningful change has occurred on the GROC. Change values on the PROM, for which patients have met the established cut-off point on the GROC, then serve as the MCID value.

Gatchel et al. wrote a 2010 editorial outlining the distribution- and anchor-based methods and the methodological issues related to the MCID response measure. The authors reported that it is erroneous methodology to use one subjective measure as an external criterion for another subjective measure. They also stated that, statistically speaking, the correlation between two related, subjective measures on the same individual will be almost absolute. The authors propose using two independent measures, which measure the same construct, when assessing meaningful change (Gatchel et al., 2010).

Other measures, such as minimal clinically important change and minimal important change, have also been used to describe longitudinal change in a person's score (Gatchel et al., 2010). The minimal important change (MIC) is a similar response measure to the MCID as it demonstrates amount of change based on what patients perceive as important (Wiebe, Matijevic, Eliasziw, & Derry, 2002). Another measure is the minimum detectable change. The MDC, also known as the reliable change or smallest real difference, is calculated by multiplying SEM by the z-score associated with the desired confidence level and square root of 2 (Schmitt et al., 2004). This response measure "represents the smallest change in score that likely reflects true change rather than measurement error alone" (Stratford et al., 1996). The MDC ultimately represents the magnitude of change needed to be confident that the change is not due to measurement error alone. The advantage of this method is that it incorporates reliability into calculations of responsiveness (Schmitt et al., 2004). The lower the reliability coefficient, the greater the SEM and in turn, the greater the MDC value. The disadvantage of the MDC is that it is based on a statistically reliable change and may not reflect clinically meaningful change (Kolber & Hanney, 2010).

Lastly, the minimal important difference (MID), also known as the minimal clinically important improvement (MCII) is defined as the "smallest change in measurement that signifies an important improvement in a patient's symptoms" (Tubach, Ravaud, Baron, Falissard, Logeart, Bellamy, & Bombardier….., et al., 2005). The MCII is expressed as a percentage of patients who improved on the measurement instrument. Similar to the MDC, the MCII or MID, is also expressed in the same units as the outcome measure (Schmitt et al., 2004).

Benefit-harm trade-off method

The benefit-harm trade-off method was developed in 2005 and used to develop the sufficiently important difference (SID) and smallest worthwhile effect (SWE) response measures (Barrett, Brown, Mundt, & Brown, 2005). This method overcomes several of the limitations of the anchor- and distribution-based methods, including assessing meaningful change based on weighing the potential harms of an intervention against its benefits. Additionally, this method assesses treatment efficacy based on patient perceptions. According to a 2001 study, distribution- and anchor-based approaches rely on either clinician's perspectives or the properties of a measurement instrument to establish worthwhile change (Wells, Beaton, Shea, Simon, Strand, Brooks, & Tugwell, 2001). This method also allows for comparison of pre- and post-intervention values to determine if meaningful change has actually occurred after an intervention (Ferreira, Herbert, Ferreira, Latimer, Ostelo, Grotle, & Barrett, 2013). Lastly, the benefit-harm trade-off method does not utilize a subjective, flawed external criterion, such as the GROC, to judge change with the intervention.

One of the more recent response measures, developed in 2005, is the SID. This response measure is defined as, "the smallest amount of patient-valued benefit that an intervention would require justifying the associated costs, risks, and other harms" (Barrett et al., 2005). The SID

takes into account both the potential harms and benefits of an intervention. Benefits may be in the form of improvements in quality and/or length of a person's life, as well as, reductions in functional impairments, and decreased risks of adverse events. Harms include both costs and risks. Risks are defined as frequent, but less severe, side effects, as well as, less frequent and but more severe adverse events. Costs contain both monetary and opportunity costs (Barrett et al., 2005).

Barrett and colleagues described three important properties of the SID measure: the SID value will be equal or greater than the MID value for a given treatment, potential harms and benefits will be assessed when determining the SID, and benefits and harms are qualified using the term probable versus absolute. The SID response measure is derived from a patient's perspective and includes both probable benefits and potential harms when deriving the value (Barrett et al). There are a few considerations to keep in mind when reviewing the SID measure. Monetary costs are highly variable among different persons. A person with a lower income and/or an uninsured person may require greater benefits to offset the added financial burden that accompanies the chosen intervention. Additionally, these costs may vary over time as the price of healthcare changes. As Barrett et al report, the SID is a relative measure and inherently unstable as it may change over time (Barrett et al., 2005). As with many domains of healthcare, the SID will likely continue to change with evolving healthcare and varying monetary costs.

Also in 2005, Barrett and colleagues developed the SID value for the common cold (Barrett, Brown, Mundt, Dye, Alt, Safdar, & Maberry, 2005). The authors determined and compared SID values for the following four treatments for the common cold: vitamins, herbal supplements, lozenges, and prescription pills. Each of these treatments had different associated side effects, effectiveness, and monetary costs. The subjects were informed of the probable

benefits and potential harms of each treatment and then asked if they would choose to undergo the proposed treatment. If the answer was yes, the researchers sought to determine the SID for the duration benefit of the intervention. Subjects were presented with a scenario outlining the length of time for which the intervention would decrease the person's cold symptoms. The researchers continued to decrease the length of time until the subject stated that they would no longer choose the intervention if it only decreased their cold symptoms by the stated amount. The minimum duration of reduced cold symptoms then served as the SID value for duration of each treatment.

The smallest worthwhile effect is similar to the SID response measure. The SWE is an effect of an intervention as "it refers to the hypothetical difference between the outcome a patient would experience with and without intervention" (Ferreira, Ferreira, Herbert, & Latimer, 2009). Thus far, the SWE has been established for breast cancer, leg ulcers, low back pain, chronic obstructive pulmonary disease, and exercise programs for fall prevention (Duric, Fallowfield, Houghton, Coates, & Stockler, 2005; Petherick, O'Meara, Spilsbury, Iglesias, Nelson & Togerson, 2006; Ferreira, Herbert, Ferreira, Latimer, Ostelo, Nascimento, & Smeets, 2012; Ferreira et al., 2013; McNamara et al., 2015; Franco, Sherrington, Ferreira, Ferreira, 2016). Ferreira and colleagues initially described the SWE response measure in 2009. They reported that the SWE has three characteristics. First, the SWE can only be evaluated by the person receiving the intervention. Second, the SWE is intervention-specific as subjects are asked to determine the amount of hypothetical improvement, by weighing out the potential harms, of the specific intervention. Finally, the SWE is the difference in a subject's outcome with the intervention and the outcome they would have had without the intervention (Ferreira et al., 2009). Although the SWE measure is very similar to the SID measure, one significant difference

is that the SWE is defined in terms of between-group difference. Subjects are asked to determine SWE estimates beyond what improvement can be expected without the intervention. One potential use of SWE estimates is to power future randomized controlled trials (RCTs). Since RCTs compare outcomes with and without an intervention, it only makes sense that the magnitude of effect used to power the trial should also be derived from between-group difference (Ferreira et al., 2013).

Ferreira et al established a SWE estimate for low back pain by presenting subjects with hypothetical scenarios about the potential effects of an intervention (physiotherapy and nonsteroidal anti-inflammatory drugs) and then determining the smallest degree of improvement that is necessary for subjects to choose that intervention (Fereira et al., 2013; Ferreira et al., 2012). These authors used a similar protocol to Barrett et al. The protocol began with assessing pain and disability at baseline. Next, a trained interviewer, using a standardized script, described how much improvement in pain and function could be expected with no intervention. Then, the interviewer outlined how the two interventions were administered, how much each one cost, potential adverse effects, and the degree of additional improvement which could be expected if the subject underwent each treatment. Subjects were then asked if they would choose the intervention, in light of the outlined benefits and harms. If the answer was yes, then the interviewer would incrementally decrease the degree of improvement until the subject answered no, indicating that the benefits no longer outweighed the potential harms. This estimate then served as the smallest worthwhile effect. This same protocol was used to determine SWE value for both physiotherapy and nonsteroidal anti-inflammatory drugs. After the determination of both SWE values, each subject underwent the intervention previously chosen for them prior to the

study. The interviewer administered the same protocols again four weeks later to determine if the subjects assigned similar SWE values before and after the intervention.

Most recently, Franco et al. determined the SWE value of an exercise program to prevent falls among older individuals. The authors used the same benefit-harm trade-off protocol as Ferreira et al to establish the estimates. In addition to assessing SWE values using the benefit-harm trade-off methodology, this study also assessed SWE values using the discrete choice method. The authors then compared the SWE estimates between the two methods. One disadvantage, pointed out by the authors, of the benefit-harm trade-off method was that only one domain of interest can be changed at a time. For example, only the degree of potential improvement in pain is assessed, while holding the potential harms constant. The authors also outlined the advantages of this methodology, including that the benefit-harm trade-off method can be applied easily and quickly (Franco et al., 2016).

In conclusion, the benefit-harm trade-off method overcomes many of the limitations of the distribution- and anchor-based approaches. The SWE is a viable response measure that has potential to provide meaningful information as to a treatment's efficacy. The proposed study aims to establish the smallest worthwhile effect value for both pain and function, using the benefit-harm trade-off method, for the Total Knee Replacement.

Financial costs and operative risks of total knee arthroplasty

A TKR can be a potentially costly surgery, encompassing both outpatient and inpatient costs. According to Bozic et al., the average monetary costs of a TKR associated with the 90-day perioperative period is $36,553.00 (Bozic, Stacey, Berger, Sadosky, & Oster, 2012). These authors also cited the mean total per-patient healthcare costs were typically higher 12 months post-operatively compared to the 12 months before surgery. Most of these increased costs were related to the inpatient stay. They also found that the outpatient costs were slightly lower in the

follow up period ($4338) compared to before surgery ($4571). Another 2012 study stated that a TKR costs approximately $20,000 per procedure (Cram, Lu, Kates, Jasvinder, Singh, Li, & Wolf, 2012). Typical hospital costs for a patient with a TKR on Medicare in Savannah, Georgia range from $11-14,000 ("Compare or hip replacement," 2011). According to the surgeons at Optim Orthopedics in Savannah, Georgia, the 90-day perioperative costs in Savannah range from $14-17,000. Medicare typically covers 100% of the inpatient costs and 80% of the outpatient costs. Private insurers vary in what they cover, but out of pocket expenses are typically more than Medicare recipients (Greengard & Kruick, 2012).

There are several potential complications that can occur during and/or after a TKR. Some of the more common complications include infection, pneumonia, deep vein thrombosis, and pulmonary embolism. These complications may result in a revision surgery of the TKR, further hospitalization, and, in the most severe situations, death. Juni et al cited an overall revision rate of 1% in patients after a TKR (Juni, Reichenbach & Dieppe, 2006). When examining the overall complication rate, one study demonstrated the risk of operative complications with knee replacement is estimated to be approximately 3% (Bozic, Grosso, Lin, Parzynski, Suter, Krumholz….., &Drye, 2014). Similarly, Cram et al found an unadjusted ninety-day complication rate of 2.1% in specialty hospitals and 3.8% in general hospitals in the United States (Cram, Vaughan-Sarrazin, Wolf, Katz, & Rostenthal, 2007). However, a 2008 study by Hamel et al found higher overall complication rates in a small cohort of 51 patients. Complications in this study were defined as anemia, pulmonary embolism, pulmonary edema, deep wound infection, peripheral neuropathy, and delirium. The authors stratified complication rates after a TKR into two categories: patients aged 60-74 and patients aged 75 years and older. The overall

complication rate was 16% for those aged 60-74 and 18.8% for patients aged 75 and older (Hamel, Toth, Legedza, & Rosen, 2008).

When examining the rate of infection status post TKR, the rates vary from 0.67-2.2% (Merle-Vincent, Couris, Schott, Conrozier, Piperno, Matthieu, & Vignon, 2011; Bozic et al., 2014; Rasouli, Restrepo, Maltenfort, Purtill, & Parvizi, 2014). Bozic et al. examined Medicare files of patients 65 years and older who underwent a TKR. Of the 626, 781 patients who had a TKR from 2008-2010, 0.67% developed a wound infection after their surgery. Another study assessed complication rates in a cohort of 2,549 patients who had a TKR from April 2010-June 2012. These authors observed a similar infection rate to Bozic et al. of 0.90% (95% CI, 0.54%-1.27%) (Rasouli et al., 2014). This study also examined risk factors contributing to surgical site infection in their cohort of patients. Preoperative anemia was one of the strongest modifiable risk factors for infection. The authors proposed that the higher incidence of blood transfusions with these patients may be a principle reason why they are more likely to develop a surgical site infection. The male gender was one of the strongest nonmodifiable risk factors in the study, possibly due to the effect of sex hormones on the immune system (Rasouli et al., 2014). Men had a significantly greater likelihood of developing an infection compared to females. The authors hypothesized that testosterone may decrease immune function while estrogen may improve it. In another study, Merle-Vincent et al found higher infection rates in their cohort of 264 patients. At a two-year follow-up, six patients (2.2%) had developed a post-operative infection (Merle-Vincent et al., 2011). Surgical site infection is a potentially serious complication after a TKR. According to Juni et al, infection can result in a need for an early revision surgery (Juni, Reichenbach & Dieppe, 2006).

Another potential complication after a TKR is a deep vein thrombosis (DVT) and/or pulmonary embolism (PE). The incidence of a DVT and PE after a TKR ranges from 9-34% and the 0.75-3%, respectively (Bozic et al., 2014; Merle-Vincent et al., 2011; Watanabe, Sekiya, Kariya, Hoshino, Sugimoto, & Hayasaka, 2011). The study by Bozic et al. assessed 626,781 Medicare beneficiaries, aged 65 and older, who underwent a TKR, and found an incidence for PE of 0.75% (Bozic et al., 2014). Merle-Vincent et al demonstrated a higher incidence of PE (2.2%) and 9% of a DVT. The authors in this study conducted a multi-site study examining 299 patients, with a two- year follow-up, who underwent a TKR. They also found a 10.6% admission rate for a problem related to their surgery (Merle-Vincent et al., 2011). Watanabe et al conducted a study assessing the rate of preoperative and postoperative DVT and PE in 64 patients undergoing a TKR. The authors concluded a postoperative incidence of a DVT was 34%, PE 3%, and DVT with a PE 13%. It is important to note that none of these patients demonstrated signs or symptoms consistent with either a DVT or PE. However, antithrombotic medication was initiated with these patients who demonstrated a DVT and/or PE.

Bozic et al also examined the incidence of pneumonia in their cohort of patients after a TKR and found an incidence rate of 0.85% (Bozic et al., 2014). (Bozic et al., 2014). Pneumonia can occur due to prolonged bed rest and can be mitigated with the use of a spirometer and early ambulation.

Additionally, a few studies researched mortality rates related to a TKR. Mortality after a TKR is estimated to be approximately 1% in the first three months after surgery (Juni et al., 2006). Similarly, Cram et al found an unadjusted mortality rate after discharge to be 0.3% from 2007-2010 (Cram et al., 2012).

Demographic and Psychosocial Variables related to SWE values

Several baseline and psychosocial variables will be assessed, in the current study, for their potential influence on SWE values. These variables include the following: baseline pain, baseline function, income level, education level, depression, pain catastrophizing behavior, anxiety, patient expectations, and self-efficacy. The above variables were chosen because published research supports their association with post-operative outcome after a TKR.

Previous studies have shown that preoperative pain scores are associated with postoperative pain (Brander, Stulberg, Adams, Harden, Bruehl, Stanos & Houle, 2003; Judge, Arden, Cooper, Javaid, Carr, Field, & Diepp, 2012; Lingard, Katz, Wright, & Sledge, 2004). A recent meta-analysis found preoperative pain to be one of the strongest independent predictors of postoperative outcome three months after surgery (Lewis, Rice, McNair & Kluger, 2014). A study by Kennedy et al demonstrated that subjects with higher baseline scores on the six-minute walk test had higher scores on the same measure at one week after surgery (Kennedy, Hanna, Stratford, Wessel & Gollish, 2006). Similarly, pre-operative pain at rest, as well as, pain catastrophizing and depression were shown to be correlated with high levels of continued pain at six months after a TKR (Noiseux, Callaghan, Clark, Zimmerman, Sluka & Rakel, 2014).

According to numerous studies, pain catastrophizing is a strong predictor of postoperative pain and function, as far as 2 years after surgery (Forsythe, Dunbar, Hennigar & Sullivan, 2008; Riddle, Wade, Jiraneck & Kong, 2010; Noiseux et al., 2014; Hirakawa, Hara, Fujiwara, Hanada & Morioka, 2014; Burns, Ritvo, Ferguson, Clarke, Zeltzer & Katz, 2015). Riddle et al. demonstrated that patients with a score of 16 or higher on the Pain Catastrophizing Scale (PCS) were 2.67 (95%CI = 1.2, 6.1) times more likely to have a poor pain outcome 6 months after knee arthroplasty compared to those who scored 15 or below. Additionally, patients were 2.18 (95%CI = 0.91,5.19) times more likely to have a poor functional outcome if they scored 16 or higher

compared to those who scored less than 16 (Riddle et al., 2010). A recent systematic review concluded that pain catastrophizing behavior affected outcomes after a TKR, specifically the intensity of chronic pain (Burns et al., 2015).

Judge et al. found that lower income levels predicted outcomes after a knee arthroplasty. Specifically, these researchers found that people living in poorer geographical areas had worse pain and function 6 months after a TKR compared to those living in more affluent areas (Judge et al., 2012). A 2014 study also examined the impact of income on post-operative outcome and discovered that subjects from a household with an annual income of less than $25,000 were more likely to have functional limitations post-operatively compared to subjects from higher-income households (Barrack, Ruh, Chen, Lombardi, Berend, Parvizi, Valle….., et al., 2014). Mahomed et al demonstrated that higher education levels were predictive of better postsurgical functional outcomes (Judge et al., 2012; Mahomed et al.,2002). Additionally, a 2011 study found that low education levels, specifically less than high school, was significantly correlated with greater pain scores on the Western Ontario McMaster (WOMAC) scale at six-months after a TKR (Lopez-Olivo, Landon, Siff, Edelstein, Pak, Kallen, Stanley….., et al., 2011).

Anxiety and depressive symptoms also have been shown to predict post-operative outcome after a TKR. Several studies have shown that preoperative depressive symptoms and anxiety are associated with increased pain and decreased function following a knee arthroplasty (Brander et al., 2003; Judge et al., 2012; Noiseux et al., 2014). Noiseux and colleagues concluded that depression was highly correlated with continued pain at six months after surgery (Noiseux et al., 2014). Similarly, Brander et al found that high levels of pre-operative depression predicted more pain at 1 year after surgery when compared to lower levels of depression. However, moderate depressive symptoms were not statistically correlated with pain, although the

authors did find that moderate preoperative anxiety scores predicted pain at one year post-operatively (Brander et al., 2003).

Lastly, patient expectations and self-efficacy have also demonstrated significant associations with postoperative status. A 2010 systematic review found that self-efficacy was a strong predictor of disability in persons with OA at six-months with individuals having lower self-efficacy scores yielding greater disability compared to those with higher self-efficacy (Benyon, Hill, Zadurian, & Mallen, 2010). A more recent systematic review yielded low association between pre-surgical self-efficacy and post-operative functional outcome; whereas postoperative self-efficacy was strongly associated with functional outcomes after a TKR (Magklara, Burton & Morrison, 2014). Wylde et al examined the role of preoperative self-efficacy on pain and function after a TKR. The researchers concluded that preoperative self-efficacy did not predict postoperative pain, but was a significant predictor of postoperative function on the WOMAC (Wylde, Dixon, & Blom, 2012).

Many studies have demonstrated significant associations between high preoperative patient expectations and improved postoperative function at 6 months and one year compared to those with lower expectations (Mahomed et al., 2002; Tejada, Escobar, Bilbao, Herrera-Espineira, Garcia-Perez, Aizpuru, & Sarasqueta, 2014). In one study by Mahomed et al, individuals who expected complete pain relief after their TKR had higher function and pain scores on the WOMAC at 6 six months (Mahomed et al., 2002).

CHAPTER THREE
Methods

This study determined the smallest worthwhile effect, utilizing a benefit-harm trade-off method, for both pain and function prior to Total Knee Replacement. A 6-month follow-up was conducted to determine satisfaction with outcome and the extent to which actual 6-month outcomes align with expected outcome assessed at baseline. The first section of this chapter presents the methods of the feasibility study, followed by the results of the feasibility study. The second section outlines the methods for the main study, including subject recruitment, data collection, formation of the standardized script, and the 6-month follow-up. The third and final section describes the analytic approach for the main study.

Feasibility Study

The aims of the feasibility study were to: 1) assess the feasibility of using total joint classes at two local hospitals for subject recruitment, 2) assess the practicalities of three locations to conduct data collection, and 3) assess the wording and interpretability of the standardized script as well as time to administer the script.

Subjects

A sample size calculation was conducted, using an equation by Viechtbauer et al. to detect a 10% probability, with 95% confidence, that a subject may misunderstand the script's wording (Viechtbauer, Smits, Kotz, Spigt, Serroyen, & Crutzen, 2015). The calculation yielded thirty subjects. A sample of 30 subjects was recruited to participate in the feasibility study. Inclusion criteria were men and women aged 50-90 with a diagnosis of advanced symptomatic knee osteoarthritis (OA), as determined by the participating surgeons, and who elected to undergo a total knee replacement (TKR) by one of six participating surgeons at one of the following two surgeon practice groups: Chatham Orthopedics or Optim Orthopedics. Exclusion

criteria included the following: knee arthroplasty revision, simultaneous bilateral knee arthroplasty, unicondylar knee arthroplasty, or knee arthroplasty for reasons other than OA including inflammatory arthropathy or cancer, currently participating in a randomized trial, or unwilling or unable to sign informed consent.

Procedures

The primary investigator (PI) attended the total joint classes 3x/week at Saint Joseph's Hospital (SJO) and Memorial University Medical Center (MUMC). Each data collection session started with subjects reading and signing the consent form, completion of a demographic questionnaire, and administration of the standardized script. Additionally, subjects completed five self-report measures; the Knee Injury and Osteoarthritis Outcome Score (KOOS) Pain and Function, daily living subscales, Patient Health Questionnaire Depression Scale (PHQ-8), Pain Catastrophizing Scale (PCS), GAD-7 Anxiety Scale, Hospital for Special Surgery (HSS) Knee Replacement Expectations Survey, and the Arthritis Self Efficacy-8 item (ASES-8 item) Scale. Time to complete only the standardized script, as well as, the overall data collection session was recorded. A debriefing questionnaire was administered after the standardized script. The debriefing questionnaire was designed to assess the script's wording and interpretability. All data from the demographic questionnaire, self-report measures, and the standardized script was into an excel spreadsheet.

Pilot study timeframe

Data collection began on July 12th, 2016 and finished on December 22nd, 2016. Given a three-week hiatus, the total time to complete the feasibility study was 4 ½ months.

Study Aims

Aim #1: Assess the feasibility of using the total joint educational classes at SJO and MUMC for subject recruitment.

The study began with the PI attending the joint replacement education classes at both hospitals (SJO and MUMC). Subject recruitment was more successful at SJO versus MUMC. SJO joint classes had consistently higher numbers of participants and more surgeons performing joint replacement surgeries. Prior to each class, the PI read a standardized recruitment script to the class participants, remained in the classroom during the class, and then remained afterwards to administer the study to willing class participants. In December, the PI began arriving at SJO hospital an hour prior to the joint education class to recruit subjects who had finished their pre-operative visit and were waiting for the joint class to begin. Pre-operative visits at SJO are generally scheduled on the same day as the class with time slots of 8:00 am, 9:00 am, and 10:00 am. As the pre-operative visit only takes an hour, there are several class participants who had finished their pre-operative work-up and sat in the waiting room until the joint class began. This left a cohort of subjects who were present and available for study recruitment prior to the joint class. Recruiting subjects to participate in the study both before and after the joint class significantly improved enrollment. Of the last seven enrolled subjects, five of them were recruited prior to the joint class.

After assessing the feasibility of utilizing both hospitals for subject recruitment in this study, it was concluded that MUMC was not a good option to use in the main study for subject recruitment. Subject recruitment from MUMC was slow because of a small number of class participants. From feasibility study inception until mid-September, only three subjects from MUMC enrolled in the study. The MUMC joint replacement education classes had poor attendance, as well as. low numbers of surgeons performing joint surgeries. Therefore, after discussion with the advisor, the PI ended recruitment at MUMC on September 15th.

Based on the feasibility study, SJO was the sole recruitment site for the main study. However, there were two problems noted with subject recruitment at SJO. In October, the Savannah area's highest volume surgeon stopped doing surgeries on Wednesdays at SJO because of a disagreement with the operating room staff. Prior to October, he performed 5-7 joint surgeries each on Mondays, Wednesdays, and Thursdays. In late October, he made the decision to stop doing surgeries on Wednesdays, which drastically decreased class enrollment at SJO on these days. However, because of increased operating room availability on Wednesdays, other surgeons began doing knee replacement surgeries at SJO on these days. Three out of four of these surgeons agreed to have their patients recruited for the study, therefore increasing the number of participating surgeons from six to nine.

Another problem noted during the feasibility study was the restricted age range of the subjects. Two potential subjects were lost because they fell below the 50-year age minimum (one person was 45 years old and the other was 49 years old). In response, the age range will be expanded to 45-90 years old in the main study.

Aim #2: Assess availability of two different locations within SJO and one external location for data collection

Two specific sites within SJO and one external site to administer the data collection session were assessed during this study. The SJO sites included the hospital-affiliated classroom where the joint class was held and an exam room in the hospital-affiliated Physical Therapy clinic. The external site was a conference room at the Armstrong Center. The Armstrong Center was offered as an option to class participants who agreed to participate in the study, but were

unable to remain after the class. If this option was used, the participants provided the PI with their contact information to schedule a day and time to administer the study.

The hospital-affiliated classroom, where the joint replacement educational class was held, worked well to administer the data collection session. At the beginning of the feasibility study, SJO's conference room was not an option. However, this changed after a subject agreed to participate in the study, but refused to travel to the hospital affiliated PT department. The PI was then given permission to use the SJO classroom, which is where the remainder of the data collection sessions were held. Each hospital classroom had ample seating and provided a confidential and comfortable location to administer the study.

Two subjects agreed to participate in the study, but were unable to remain on site immediately following the joint class. They consented to schedule a time and date to meet at the Armstrong Center. The PI was unable to successfully contact one subject. Contact was established with the second subject, but scheduling conflicts prohibited the collection of data. Experiences with these two subjects suggested that scheduling patients to come to the Armstrong Center was not feasible. Therefore, the hospital affiliated classroom served as the sole location for data collection in the main study.

Aim #3: Assess wording and interpretability of the standardized script and time to administer the standardized script and self-report measures.

Average time to administer the entire data collection session was 35 minutes (range of 20-105 minutes). The standardized script took an average of 10 minutes to complete. A debriefing questionnaire was used to assess the wording and interpretability of the standardized script. Overall, subjects reported no confusion with the script's wording and offered no

recommendations for improvement, based on the debriefing questionnaire. However, while administering the script, the PI noted several areas of wordiness which could be made more concise. Specifically, the word "functional" was deleted when referring to "functional activity" and run on sentences were separated and condensed.

Self-report measures

Six self-report measures were assessed during this study. The Patient Health Questionnaire depression scale (PHQ-8) and Pain Catastrophizing Scale (PCS) scales were quick and easy to administer with no problems noted by the subjects. The PI incidentally left the PHQ-8 self-report measures out of five of the subject packets. This omission reinforced the importance of assuring that all data packets are complete prior to collecting data on a subject. The PHQ-8 sample mean was 7.7 (sd=7.1, range=0-24). There was no missing data for the PCS measure and the sample mean for the total PCS was 19.9 (sd =14.1, range = 0-52).

The Arthritis Self Efficacy-8 item (ASES) was easy to administer. However, there were two points of confusion. The first point of confusion occurred because several items on the scale included the word "fibromyalgia", in addition to, "arthritis." The second point of confusion centered around whether the subject should answer the questions as they were currently feeling (no intervention) or anticipation of their upcoming TKR. These issues will be clarified in the main study by removing the word "fibromyalgia" from the questionnaire and by providing additional instruction to indicate the forms should be applied to the patient's current state. It is unknown how revising the wording on this questionnaire affects it's measurement properties.

The HSS Knee Surgery Expectations Survey generated the most problems of the self-report measures. All subjects completed the self-report measure; however, 47% had missing items. The number of missing items ranged from 1-6 for each subject, with questions #1 and 2

being the most commonly missed questions. Additionally, several subjects commented on how a few of the questions did not apply to them. The sample mean for this survey was 29.0 (sd=6.2, range=19-40). In response to the problems with this self-report measure, a 4-item Patient Expectations Survey (Mohomed, Liang, Cook, Daltroy, Fortin, Fossel, & Katz, 2002) was substituted to decrease confusion and time to complete the survey.

The GAD-7 self-report measure was understood by patients and was easy to administer. The sample mean was 7.4 (sd=6.0, range=0-20). The KOOS-pain subscale was also easy to administer with no missing data. The raw data for both the pain and function KOOS subscales was converted to a 0-100 point scale. The sample mean of the pain subscale, using the 100-point scale, was 33.2 (sd=15.4, range=11-67). All subjects completed the KOOS-function subscale; however, 13% of the subjects (4 subjects) did not complete all items on the subscale. One of these four subjects did not complete at least 50% of the items, therefore this data was excluded from the mean calculation. The sample mean of the function subscale, using the 100-point scale, was 37 (sd=15.9, range=12-71). In the main study, the PI will review each self-report measure, including both KOOS subscales, for completeness prior to ending the study session.

Results

Demographics

A total of 63% of the sample of 30 subjects was female with a mean age of 67 years old (49-85). Highest degree of education and approximate yearly income were also assessed using the demographic questionnaire. 26.7% of the sample had earned post-baccalaureate degrees, while a similar number of subjects (23.3%) were high school graduates. 40% of the sample had a yearly income of $50-100,000. There was no missing education data, but one subject refused to answer the income question.

Smallest Worthwhile Effect Results

The sample mean for the 6-month estimated KOOS Pain subscale was 78.4 (sd=16.3, range=29-99) while the average for the function subscale was 80.8 (sd=11.4, range=47-100). The average smallest worthwhile effect (SWE) for the KOOS Pain subscale was 45.3 (sd=20.6, range=9-78) and 42.1 (sd=19.2, range=12-76) for the KOOS Function, daily living subscale. (Table 1) To determine the smallest worthwhile effect, the PI changed responses on up to five questions, with the lowest pain scores, on the baseline KOOS subscales. It was hypothesized that five questions would have been enough questions to capture the SWE estimate. The PI asked the subject if they would still consider their surgery worthwhile if their response on one of the KOOS questions was shifted to either a greater level of pain with activity or a greater level of difficulty, depending on the subscale. If the subject answered "yes," indicating that they still considered the surgery worthwhile, then their response was shifted on a second question. This protocol continued until either the subject answered "no" when asked if the surgery would still be worthwhile or the subject answered "yes" to all five questions on the subscale. The same procedure was used for both the pain and function subscales. Sixty percent of the sample answered "yes" to all five KOOS pain and function items when asked if they would still consider their surgery worthwhile if their responses were changed to indicate a higher level of pain or difficulty with functional tasks. Because more than half of the subjects answered yes to all five questions, the SWE was likely not found in this sample. This may have occurred because the subjects didn't recall that they should determine the acceptable amount of pain or difficulty with activity by weighing the associated costs and risks of the surgery against the benefits of the surgery. In the current script, was stated early on that subjects should determine amount of pain or difficulty with activity by weighing the costs and benefits of the surgery. However, no additional reference is made throughout the remainder of the script. Additionally, the script was

capped at (5) questions for each subscale, which may have led to a ceiling effect. To combat both problems, the script wording and protocol will be modified. The statement "considering the 10% to 15% worsening you would experience without the knee replacement surgery" will be added two additional times throughout the script. Also, the protocol for establishing the SWE value will be modified so that the PI will continue to ask all the questions on each subscale until the subject answers "no" when asked if the surgery will still be worthwhile.

Table 3.1. KOOS Subscale

KOOS Subscale	Baseline Scores	Estimated 6-month Scores	SWE Values
Pain Subscale	33.2 (sd=15.4, range=11-67)	78.4 (sd=16.3, range=29-99)	45.3 (sd=20.6, range=9-78)
Function, daily living Subscale	38.6 (sd=16.7, range=12-71)	80.8 (sd=11.4, range=47-100)	42.1 (sd=19.2, range=12-76)

Summary of key issues and associated modifications from the feasibility study

The following provides a bulleted list of issues identified during the feasibility study and their associated modifications for the main study

- Subject recruitment
    - Subject recruitment occurring only at the beginning of the joint class
        - PI will recruit subjects both before and after the 11:00 am joint class at SJO
    - MUMC's joint class as a place for subject recruitment
        - Delete the use of MUMC in main study
    - Narrow age range for inclusion criteria
        - Broaden the age range from 50-90 to 45-90 years old

    - Savannah area's highest volume surgeon stopped doing surgeries on Wednesdays at SJO.
        - Utilize three additional surgeons, increasing the number of participating surgeons from six to nine

- Data collection location
    - Physical Therapy affiliated clinic was not conducive to administration of the study
        - Delete the use of the Physical Therapy affiliated clinic from the main study
    - Armstrong Center was not a viable option for data collection
        - Delete the use of the Armstrong Center from the main study
- Standardized script wording
    - Five questions are not sufficient to determine the smallest worthwhile effect value using the standardized script
        - PI will continue to change responses on KOOS-pain and function subscales until the subject answers "no," indicating that the surgery would no longer be worthwhile
            - Modify the script's wording to reflect changing subjects' responses on two items at one time, versus one item
    - Missing data for self-report measures and age data from the demographic questionnaire

- Wording on the demographic questionnaire will be changed from "approximate age" to "birth date: ___day ____month ____ year
- PI will carefully review all outcome measures and demographic questionnaires to ensure completeness, prior to the subject leaving the session.

- Confusion as to the wording on both the ASES and the HSS Knee Surgery Expectations Survey
  - PI will delete the word, "fibromyalgia" from the ASES self-report measure
  - Expectations self-report measure will be changed from the HSS Knee Expectations Survey to a 4-item Patient Expectations Survey

Main Study

Subjects

A sample of 120 subjects scheduled to undergo total knee replacement at SJO in Savannah, Georgia were recruited through convenience sampling to participate in this study. Approximately 90 knee arthroplasty surgeries are performed monthly at St. Joseph's Hospital. Chatham Orthopedics and Optim Orthopedics are Savannah based orthopedic surgery practices with a collective eight surgeons who perform most of knee arthroplasty surgeries. These eight surgeons, in conjunction with an independent Orthopedic Surgeon, have provided verbal consent to allow their patients to participate in the proposed study. All nine of these surgeons perform their knee replacement surgeries at SJO.

Inclusion Criteria

Inclusion criteria for the study include the following: men and women 45-90 years of age with a diagnosis of advanced symptomatic knee osteoarthritis (OA), as determined by the

participating surgeons, and who have elected to undergo a primary TKR by one of the nine participating surgeons.

Exclusion Criteria

Exclusion criteria include the following: knee arthroplasty revision, simultaneous bilateral knee arthroplasty, unicondylar knee arthroplasty, or knee arthroplasty for reasons other than OA including inflammatory arthropathy or cancer, currently participating in a randomized trial, or unwilling or unable to sign informed consent.

Procedures

Subjects were recruited both before and after the knee replacement educational class at SJO. St. Joseph's Hospital hosts classes three times per week with an average of seven knee replacement patients per week who attend the class. Class participants attend the total joint replacement education class approximately 1-2 weeks prior to their surgery. The nine participating Orthopedic Surgeons in the current study strongly encourage their patients to attend one of these classes. Prior to the start of each class, the PI delivered a brief synopsis of the study and asked for volunteers. Additionally, the PI arrived to the hospital approximately two hours prior to the start of the class to recruit subjects who had finished their pre-operative work-up and were waiting for the class to begin.

The data collection sessions occurred in the St. Joseph's Hospital classroom, which is the same location as the class. If subjects were recruited and agreed to participate prior to the class, they were escorted to the classroom for the data collection session. After completing the study, these subjects either returned to the waiting room or waited in the classroom until the class began. For those subjects who were recruited immediately prior to the joint class and chose to participate, they remained in the classroom afterwards to complete the data collection session.

Once a subject agreed to participate in the study, the PI provided them with the consent form and allowed time for the subject to read it and ask questions. If the subject agreed to participate, the subject and PI signed the informed consent. Each subject completed a demographic questionnaire, the KOOS Pain subscale, the KOOS Function, daily living subscale, and five other self-report measures to assess for depression, anxiety, self-efficacy, expectations, and pain catastrophizing scales.

Following completion of the self-report forms, the PI administered the standardized script to assess the smallest worthwhile effect values for both pain and function, utilizing the KOOS Pain and Function, daily living subscales. Depending on the number of participating subjects on a particular day, the script was either administered before or after completion of the self-report measures and demographic questionnaire, excluding the baseline KOOS Pain and Function, daily living subscales, which all subjects initially completed. Using this protocol allowed two subjects to complete the study simultaneously. When two subjects agreed to participate in the study, the following sequence occurred. After reading and signing the informed consent and completing the baseline KOOS subscales, subject #1 completed the self-report measures and demographic questionnaire while subject #2 answered questions via the standardized script. Then, the subjects switched roles and completed the remaining portion of the study. All subjects provided consent and completed the KOOS Pain and Function, daily living subscale prior to answering questions via the standardized script or completion of the self-report measures/demographic questionnaire. Data for all subjects was collected in a quiet and private environment.

Self-report Measures Obtained at Baseline

The five self-report measures along with income level, education level, and baseline pain and function served as predictor variables and be assessed for their association with SWE values.

The five self-report measures include: PCS, PHQ-8, GAD-7 anxiety scale, a 4-question Patient Expectation Survey, and the ASES-8 item scale scale (Forsythe, Dunbar, Hennigar & Sullivan, 2008; Riddle, Wade, Jiraneck, & Wong, 2010; Kroenke, Strine, Spitzer, Williams, Berry & Mokdad, 2009; Spitzer, Kroenke, Williams & Lowe, 2006; Mohomed et al., 2002; Brady, 2011; Wilcox, Schoffman, Dowda & Sharpe, 2014). Annual income level and highest level of education was assessed using the demographic questionnaire.

Baseline pain and function was assessed using the KOOS Pain subscale and KOOS Function, daily living subscales respectively (Collins, Misra, Felson, Crossley, & Roos, 2011). Each subscale was scored individually on a 5-point Likert scale and then transformed to a 0-100 point scale, where 0 equals extreme knee problems and 100 equals no knee problems. The KOOS demonstrates convergent and divergent construct validity and strongly correlates with the SF-36. In patients with knee osteoarthritis, ICC values for KOOS Pain subscale range from 0.80-0.97 and KOOS Function, daily living subscale 0.84-0.94 (Collins et al., 2011). Internal consistency values in patients with osteoarthritis for KOOS Pain range from 0.65-0.94 and for KOOS Function, daily living subscale 0.78-0.97 (Collins et al., 2011). The KOOS has been shown to be responsive in patients after total knee replacement (Collins et al., 2011 & Roos & Lohmander, 2003). Roos and Lohmander found a Minimal Clinically Important Difference (MCID) of 8-10 to be appropriate (Roos et al., 2003).

Pain catastrophic thinking was measured using the PCS scale. This scale has thirteen statements of emotion related to different pain-related emotions. A five-point Likert scale is used to annotate the frequency. For each of the thirteen statements, the subject is asked how frequently they feel the statement of emotion. The total score ranges from 0 (no catastrophizing) to 52 (severely catastrophizing). The PCS contains three subscales, rumination, magnification,

and helplessness, all of which have adequate to high internal consistency (Chronbach's alpha ranging from .66-.87) (Forsythe et al., 2008; Riddle et al., 2010).

The PHQ-8 measure was used to quantify the extent of depressive symptoms. The PHQ-8 was developed from the Patient Health Questionnaire depression scale (PHQ-9) and was made available in 2006. The PHQ-8 consists of eight of the nine criteria for depression from the Diagnostic and Statistical Manual for Mental Disorders, fourth edition (DSM-IV). The PHQ-8 consists of eight questions which have a maximum score of 24. Kroeneke et al found that a cutpoint >10 points on the PHQ-8 accurately determined the presence of major depression with sensitivity of 70% and specificity of 98% when compared to the established diagnostic algorithm (Kroenke et al., 2009). A five-point change on this scale is considered clinically significant (Kroenke, K., 2012).

GAD-7 anxiety scale was used to assess anxiety-related symptoms. The GAD-7 consists of seven items with a maximum scale of 21 and was developed from the original 13-item GAD scale. Higher values indicate greater feelings of anxiety. A cut-point of 10 or greater demonstrated 89% sensitivity and 82% specificity when compared to the 13-item scale. According to Spitzer et al, the GAD-7 has good reliability, criterion and construct validity, as well as, good factorial and procedural validity (Spitzer et al., 2006). A score of 0-4 indicates a minimal level of anxiety, 5-9 indicates a mild level, 10-14 indicates a moderate level, and 15-21 indicates a severe level of anxiety (Spitzer et al., 2006)

The 4-question Patient Expectations Survey was used to assess pre-operative expectations. This questionnaire consists of four items which assess preoperative expectations prior to a TKR. The questions assess expectations related to pain relief, limitations in activities

of daily living (ADLs), overall success of the TKR, and likelihood of post-operative joint related complications. The responses for pain relief and ADLs are graded using a 4-point Likert scale. The other two questions, which relate to overall surgery success and joint complications, are assessed using a 100-point visual analog scale (Mohomed et al., 2002).

Self-efficacy was measured using the ASES-8 item scale. This scale is an eight-item scale with scores for each item ranging from 1-10, with 1 indicating "very uncertain" and 10 indicating "very certain." Responses are averaged, yielding a score ranging from 10 to 80 with higher values indicating greater levels of self-efficacy (Brady, 2011; Wilcox, Schoffman, Dowda & Sharpe, 2014). Wilcox et al recently assessed the reliability and validity of the English version of ASES-8 item scale and found that it demonstrated high internal consistency and concurrent validity (Wilcox et al., 2014).

Self-report Measures Obtained at 6-months Follow-up

During the 6-month follow-up, the PI administered, via telephone, both the KOOS Pain and KOOS Function, daily living subscales. Subjects also verbally completed The Patient Acceptable Symptom State (PASS) questionnaire. The PASS questionnaire is an absolute measurement of a subject's current satisfactory state (Tubach, Ravaud, Baron, Falissard, Logeart, Bellamy, Bombardier, et al., 2005). Using a yes/no question, subjects are asked if they consider their current state, taking into account, functional limitations, impairments, and pain, as satisfactory.

At the conclusion of the initial data collection session, the PI collected contact information from subjects in order to conduct the 6-month follow-up. At the time of follow-up, the PI called each subject up to 6 times during the 2-week period before and after the 6-month timepoint following TKR.

Smallest Worthwhile Effect

A standardized script was used to calculate smallest worthwhile effect values for both pain and function. Subjects were asked to estimate the magnitude of improvement in pain or function over their current baseline pain/function is necessary to be considered a worthwhile change. Baseline pain was assessed using the KOOS Pain subscale. Baseline function was assessed using the KOOS Function, daily living subscale.

While administering the standardized script, the PI asked subjects to complete additional KOOS Pain and Function, daily living subscales to indicate the level of pain and function they would need to have at six months to categorize the improvement as worthwhile. The following protocol, outlined in the next paragraph, was first administered to assess SWE values for pain and then again to assess SWE values for function

After completion of the 6-month KOOS Pain subscale, the PI shifted the subject's response on two items by one absolute value on their 6-month pain scale, indicating a higher pain level, and asked if the subject would consider the surgery worthwhile if they received this revised amount of improvement. For example, if a subject indicated "no pain" on the "pain at night while in bed" question, the PI shifted the subject's response from "no pain" to "mild pain" and asked the subject if they would still consider the surgery worthwhile if they had "mild pain" while in bed at 6 months.

If the subject answers "no," then the PI asked the subject if they would still consider the surgery worthwhile if their response was only changed on the first question. If the answer was still "no," then the PI stopped and the smallest degree of improvement in pain was recorded.

However, If the subject answered "yes" to both questions, the pain response on two additional questions were shifted one point to indicate increased pain with all four activities and the subject was asked if this amount of improvement is still worthwhile. If the subject answered "no," then the PI asked the subject if they would still consider the surgery worthwhile if their response was only changed on the three questions. If the answer was still "no," then the PI stopped and the smallest degree of improvement in pain was recorded.

If the answer was "yes" to all four questions, the pain responses on an additional two questions were shifted to one point and the question asked again. This protocol continued until the subject either answered "no" when asked if the amount of change was still worthwhile or he or she answered "yes" to all items on the pain subscale.

The same protocol was used to establish the smallest degree of improvement in function. For example, if a subject rated "none" for functional limitation with "rising from bed," the PI shifted the subject's response to the right from "none" to "mild" limitation and asked whether the subject would still consider the surgery worthwhile if they had "mild" difficulty with rising from bed at 6 months. This same protocol that used for the KOOS Pain scale was also used to determine the SWE for the KOOS Function, daily living subscale.

The order of KOOS items that were shifted to a one-point higher score was standardized based on the subjects' responses. The protocol began with changing the first two items with the lowest degree of pain or function on the 6-month subscales.

Standardized script

The intent of the script was to standardize and quantify both the risks and benefits of a TKR to determine the extent of improvement needed for the subject to rate the upcoming surgery

as worthwhile. The script also outlined the potential financial and operative risks associated with a total knee replacement.

Script for Six-month Follow-up Telephone Session

A 6-month follow-up data collection session via telephone was conducted to determine whether initial SWE values were congruent with pain and function status at six months after surgery. Each 6-month follow-up occurred within ± 2 weeks of the 6-month surgery date for each subject. The PI read and administered both the KOOS Pain and Function, daily living subscales during this phone follow-up. The PI also administered the PASS questionnaire and asked each patient if they sustained any postoperative complications for which they required an additional hospitalization. If the subject answered yes to having a postoperative complication, the reason for the additional hospitalization was asked and recorded.

Sample Size

Sample size was calculated by estimating the change score for pain using the following formula: theta +/ t*se. The difference between baseline and 6-month KOOS Pain subscale scores from the feasibility study served as the parameter of interest (theta) in the equation. The standard deviation (se) in the calculation was the standard deviation of the difference between baseline and 6-month KOOS Pain subscale scores mentioned previously. Using a 90% CI in the above calculation, a sample of 70 subjects was needed to estimate the change score for pain +/- 5 points. This five-point margin of error is sufficient because it falls below the established MCID value of the KOOS (Roos et al., 2003).

Sample size was also calculated using the number of predictor variables for the multiple regression analysis. According to Field (2009), 10-15 cases of data are needed for each predictor

variable. Using the above equation for the (9) predictor variables in the current study, the PI recruited a sample of 120 subjects for the current study.

Data Analysis

Data was double entered into excel spreadsheets to reduce risk of transcription error and then transferred to SPSS for analysis. Baseline and 6-month KOOS Pain and Function, daily living subscales were transformed to a 0-100 point scale as recommended by the scale developers (Roos, Roos, Lohmander, Ekdahl, & Beynnon, 1998). SWE values, as determined by the standardized script, were also be transformed to a 0-100 point scale and assessed for normality visually using histograms and with the Shapiro-Wilk and/or Kolmogorov-Smirnov test.

Descriptive statistics

Baseline and six-month follow-up data was described with measures of central tendency. Variables were assessed for normality using the Komorgov-Smirnov test and histograms and either the mean or median, based on normality, were used to describe the following variables: age, baseline and 6-month KOOS Pain and Function, daily living subscales, PCS, GAD-7, ASES, and PHQ-8 self-report measures. The following variables are categorical and were expressed using frequency data: gender, income level, and education level. The first two questions (listed below) on the Patient Expectation Survey are categorical and were also described using frequency data while the second two questions are continuous and were expressed using either means or medians, based on normality.

1) How painful do you expect your knee to be in one year? (not at all painful, slightly painful, moderately painful, very painful)

2) How limited do you expect to be in your usual activities in one year? (not at all limited, slightly limited, moderately limited, very limited)

Smallest worthwhile effect of baseline data

The distribution of SWE values were visually displayed using histograms. Data was described in terms of quartiles. Based on a study by Ferreira, Herbert, Ferreira, Latimer, Ostelo, Grotle, & Barrett (2013), 50^{th} and 90^{th} percentiles were marked on the histograms.

Regression analysis of baseline data

A multiple linear regression was used to assess associations between demographic and psychosocial variables and SWE values on pain and function. SWE values served as the dependent, or outcome, variable and are measured on the continuous scale. All predictor, or independent, variables are continuous or categorical in nature. (Table 2)

After testing for the assumptions of the regression analysis, univariate analyses were conducted to assess the significance of each predictor variable. Significant variables, $p<.05$, were entered into a backward stepwise regression. Variables were removed from the model if they did not significantly contribute to the model ($P>.05$). A goodness of fit test was then used to assess how well the predictor variables left in the model were associated with the SWE values. The predictor variables in the model were also assessed for their individual contribution to the model.

Table 3.2. Predictor Variables

Variable	Type	Type of Data
Income level	Predictor	Categorical
Education	Predictor	Categorical
Baseline pain	Predictor	Continuous
Baseline Function	Predictor	Continuous
Depression	Predictor	Continuous
Anxiety	Predictor	Continuous
Self-efficacy	Predictor	Continuous

Pain Catastrophizing	Predictor	Continuous/Categorical
Smallest worthwhile effect on pain	Criterion	Continuous
Smallest worthwhile effect on function	Criterion	Continuous

Six-month follow-up

Descriptive statistics were used to summarize KOOS Function, daily living subscale scores, KOOS Pain subscale scores, PASS scores, and post-operative complications. Comparisons between SWE estimates and the actual 6-month KOOS Pain and Function, daily living subscale scores were assessed using paired t-tests, with a Bonferroni correction for multiple comparisons. Additionally, PASS scores were used to dichotomize the sample into those who are satisfied with their current state and those who are not in order to determine if outcome is related to whether the subject achieved their SWE. Lastly, a Pearson chi square test was used to assess whether achievement of the SWE (yes/no) was associated with satisfaction with the actual outcome at 6-months (yes/no).

CHAPTER 4
Results

Characteristics of the Sample

A total of 121 participants provided informed consent to participate in the study and were already scheduled to undergo a unilateral Total Knee Replacement (TKR). The majority of the sample (54%) was female and the mean age was 67 (47-83) years. Participant characteristics are presented in Table 4.1.

Table 4.1. Participant characteristics at baseline

Sample Characteristics	N=121
Age (yr), *mean (sd)*	67 (9.7)
Gender (women), *n* (%)	65 (54)
Educational level, *n* (%)	
Less than high school graduate	3 (2.5)
High school degree	16 (13.4)
Some college	26 (21.8)
College degree	35 (29.4)
Some graduate school	5 (4.2)
Graduate degree	34 (28.6)
Yearly household income, *n* (%)	
<$10K	3 (2.6)
$10-<25K	6 (5.2)
$25-50K	23 (19.8)
$50-100K	50 (43.1)
$100K or greater	34 (29.3)
KOOS Pain Score (0-100), *mean (sd)*	44 (16.9)
Anticipated pain severity at 6-months (0-100), *median* (IQR)	89 (53-100)
KOOS ADL Score (0-100), *mean (sd)*	46 (19.1)
Anticipated disability level at 6-months (0-100), *median* (IQR)	82 (50-100)
Depressive symptoms (0-24), *median* (IQR)	4 (0-21)
Anxiety (0-21), *median* (IQR)	3 (0-22)
Pain catastrophizing (0-52), *median* (IQR)	13 (0-52)
Self efficacy (8-80), *median* (IQR)	26 (8-80)

Knee Injury and Osteoarthritis Score (KOOS) pain subscale- baseline pain (0-100 with lower numbers indicating greater pain)
Knee Injury and Osteoarthritis Score (KOOS) function subscale- baseline function (0-100 with lower numbers indicating greater difficulty with functional activities)
Generalized Anxiety Disorder (GAD)- baseline anxiety (0-21 with higher numbers indicating greater feelings of anxiety)
Personal Health Questionnaire (PHQ) Depression Scale- depressive feelings (0--24 with higher numbers indicating greater feelings of depression)

Pain Catastrophizing Scale (PCS)- pain catastrophizing feelings (0-52 with higher numbers indicating greater feelings of pain catastrophizing)
Arthritis Self-Efficacy Scale (ASES)- feelings of self-efficacy (10-80 with higher numbers indicating greater levels of self-efficacy)

Using the Kolmorgov-Smirnov test, baseline Knee Injury and Osteoarthritis Outcome Scale (KOOS) pain and function, daily living subscales were normally distributed; however, anticipated 6-month KOOS pain and function, Arthritis Self Efficacy-8 item (ASES), Pain Catastrophizing Scale (PCS), Generalized Anxiety Disorder (GAD-7), and Patient Health Questionnaire depression scale (PHQ-8) data was not normally distributed. Data from the 4-question expectations survey was also not normally distributed. The variables which were not normally distributed were described using the median with an interquartile range or with frequencies.

The 4-question Patient Expectations questionnaire indicated that the majority of participants (74%) expected no pain at six months and 61% of the participants expected to have no functional limitations, in their operative knee, six months after surgery. Baseline expectation data can be found in Table 4.2. The median score and interquartile range on the visual analog scale (VAS), which assessed how likely a participant felt their surgery would be a success, was 97 (66-100) points. A score of 100 would indicate that the participant felt that their surgery would be a complete success. On average, participants indicated a score of 7 (0-33) on the VAS, indicating how likely it was that they would develop a post-operative complication. Lower scores indicated a lower likelihood that the participant felt they would develop a post-operative complication.

Table 4.2. Patient Expectations at Baseline

Sample Characteristics	*n* (%)
Pain Expectation	
Not at all painful	84 (73.7)
Slightly painful	26 (22.8)

Moderately painful	3 (2.6)
Very painful	1 (.9)
Function Expectation	
Not at all limited	69 (60.5)
Slightly limited	41 (36.0)
Moderately limited	3 (2.6)
Very limited	1 (.9)
Visual Analog Scale (0-100), *mean* (IQR)	
Likelihood of postoperative success	97 (66-100)
Likelihood of postoperative complication	7 (0-33)

Pain Expectations: How painful do you expect your knee to be in one year?
Function Expectations: How limited do you expect to be in your usual activities in one year?

An independent samples t-test was used to assess differences in continuous level baseline scores (baseline pain and function, anxiety, depression, pain catastrophizing, self-efficacy, likelihood of post-operative success, and likelihood of a post-operative complication) between those who completed the 6-month follow-up and those who did not. Baseline comparisons are presented in Table 4.3. There was a statistically significant difference between those lost to follow-up and those who competed the 6-month questionnaire on depressive symptoms ($p=0.01$) and anxiety ($p=0.03$) with those who completed the follow up reporting higher levels of depressive symptoms and anxiety. The average depressive symptom score for those who completed the 6-month follow-up was 5.7 points, indicating a higher level of depression, when compared to those who did not complete the follow-up and had an average score of 3.4 points. Similarly, the average anxiety score was 5.4 points for those who completed the 6-month follow-up, indicating a higher level of anxiety, when compared to those who did not complete the follow-up and had an average score of 3.6 points. A Chi-square test was used to assess differences in categorical baseline scores (education, income, expectations related to pain, and expectations related to function). No between group differences were noted ($p>0.05$) for all categorical variables. All subjects were scheduled to undergo a TKR; however, it is unknown if

those participants who were lost to follow-up after the baseline assessment underwent the surgery.

Table 4.3. Baseline comparisons between those who completed and did not complete follow-up

	Completed follow-up (N=83) Mean (sd)	Lost to follow-up (N=38) Mean (sd)	P value
KOOS pain (0-100)	43.8 (17.4)	45.1 (16.0)	0.71
KOOS function (0-100)	45.5 (19.5)	46.9 (18.7)	0.71
GAD score (0-21)	5.4 (5.3)	3.6 (3.2)	0.03*
PHQ score, (0-24)	5.7 (5.7)	3.4 (3.9)	0.01*
PCS score, (0-52)	15.3 (13.5)	17.8 (12.7)	0.34
ASES, (10-80)	31.9 (18.5)	30.1 (18.5)	0.61
Likelihood of postoperative success, (0-100)	93.6 (7.5)	94.9 (6.6)	0.36
Likelihood of postoperative complication, (0-100)	7.4 (8.5)	8.9 (9.9)	0.41
Expectations related to pain, *n* (%)			0.84
Not at all painful	57 (74.0)	28 (73.7)	
Slightly painful	18 (23.4)	8 (21.1)	
Moderately painful	2 (2.6)	1 (2.6)	
Very painful	0 ()	1 (2.6)	
Expectations related to function, *n* (%)			0.37
Not at all limited	49 (63.6)	20 (52.6)	
Slightly limited	25 (32.5)	16 (42.1)	
Moderately limited	3 (3.9)	1 (2.6)	
Very painful	0 (0)	1 (2.6)	
Educational level, *n* (%)			0.29
Less than high school graduate	2 (2.5)	1 (2.5)	
High school degree	11 (13.9)	4 (10.0)	
Some college	20 (25.3)	6 (15.0)	
College degree	21 (26.6)	14 (35.0)	
Some graduate school	3 (3.8)	2 (5.0)	
Graduate degree	22 (27.8)	13 (32.5)	
Yearly household income level, *n* (%)			0.59
<$10K	1 (1.3)	2 (5.1)	
$10-<25K	6 (7.9)	0 (0)	
$25-50K	17 (22.4)	5 (12.8)	
$50-100K	28 (31.6)	22 (56.4)	
$100K or greater	24 (31.6)	10 (25.6)	

*statistically significant difference between groups; p<0.05

Knee Injury and Osteoarthritis Score (KOOS) pain subscale- baseline pain (0-100 with lower numbers indicating greater pain)

Knee Injury and Osteoarthritis Score (KOOS) function subscale- baseline function (0-100 with lower numbers indicating greater difficulty with functional activities)

Generalized Anxiety Disorder (GAD)- baseline anxiety (0-21 with higher numbers indicating greater feelings of anxiety)
Personal Health Questionnaire (PHQ) Depression Scale- depressive feelings (0--24 with higher numbers indicating greater feelings of depression)
Pain Catastrophizing Scale (PCS)- pain catastrophizing feelings (0-52 with higher numbers indicating greater feelings of pain catastrophizing)
Arthritis Self-Efficacy Scale (ASES)- feelings of self-efficacy (10-80 with higher numbers indicating greater levels of self-efficacy)
Pain Expectations: How painful do you expect your knee to be in one year?
Function Expectations: How limited do you expect to be in your usual activities in one year?

The average anticipated 6-month KOOS Pain subscale scores obtained at baseline (i.e., the score on the KOOS Pain subscale indicating the lowest degree of pain which the participant felt was necessary to justify TKR relative not having surgery) was 74 points and the actual 6-month KOOS score was 90 points. A paired t-test was calculated to assess differences between anticipated 6-month KOOS Pain scores obtained at baseline and actual 6-month KOOS Pain value. This t-test was statistically significant ($t=8.1$; $p<0.001$) with the actual 6-month KOOS Pain scores being, on average, 16 points higher than the anticipated 6-month KOOS-pain score.

The average anticipated 6-month KOOS Function, daily living subscale scores obtained at baseline was 73 points and the actual 6-month KOOS score was 92 points. A paired t-test was calculated to assess differences between anticipated 6-month KOOS Function, daily living scores obtained at baseline and actual 6-month KOOS function, daily living value. This t-test was statistically significant ($t=8.6$; $p<0.001$) with the actual 6-month KOOS Function, daily living scores being, on average, 20 points higher than the anticipated 6-month KOOS Function, daily living score. Table 4.4 outlines the comparisons between anticipated and actual 6-month KOOS scores.

Table 4.4. Comparisons between anticipated and actual 6-month KOOS scores

	Anticipated 6-month KOOS score Mean (sd)	Actual 6-month KOOS score Mean (sd)	Mean difference	Significance value

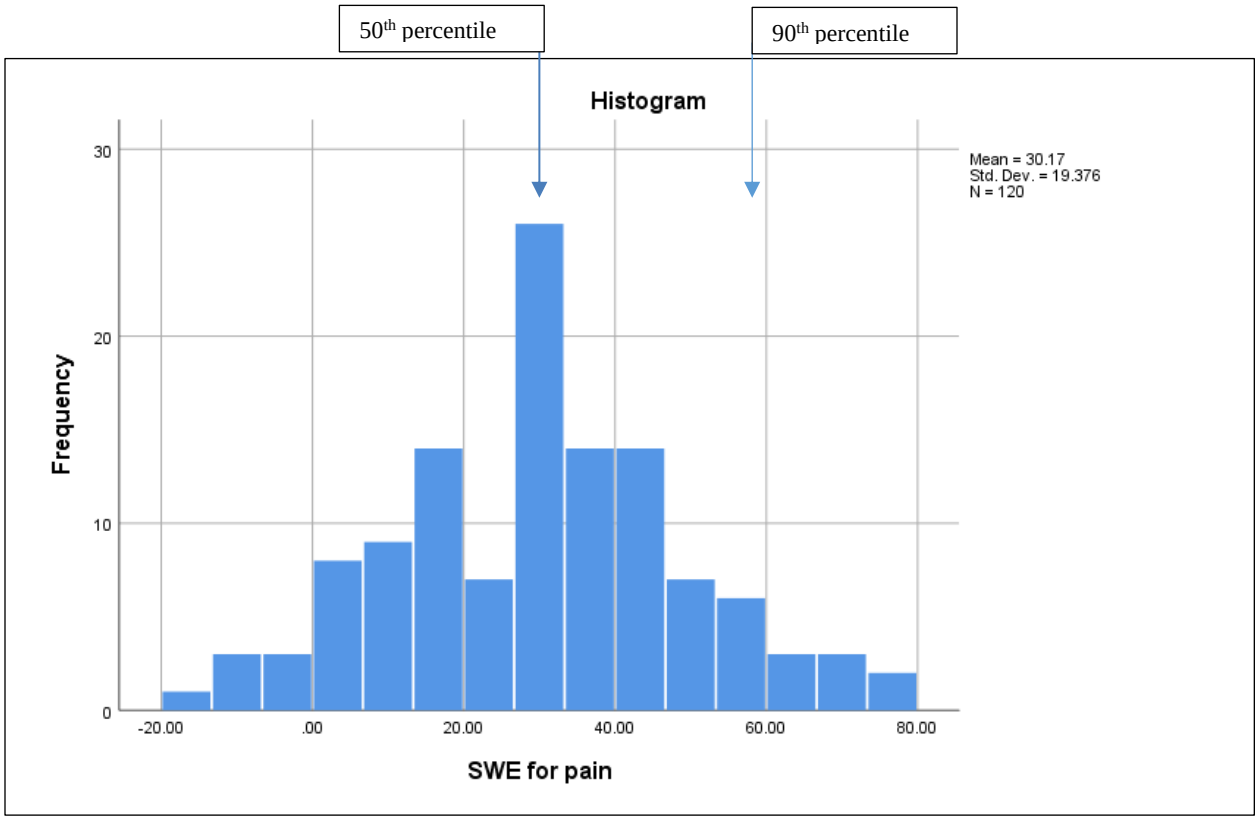
KOOS pain subscale	74.1 (15.1)	90.2 (14.4)	16.1	<0.001*
KOOS function subscale	72.5 (17.6)	92.2 (13.5)	19.7	<0.001*

*statistically significant difference between groups; p<.05
Knee Injury and Osteoarthritis Score (KOOS) pain subscale- baseline pain (0-100 with lower numbers indicating greater pain)
Knee Injury and Osteoarthritis Score (KOOS) function subscale- baseline function (0-100 with lower numbers indicating greater difficulty with functional activities)

Smallest worthwhile effect for KOOS Pain Subscale

The smallest worthwhile effect (SWE) for pain was defined as the lowest amount of improvement on the anticipated 6-month KOOS Pain subscale, over the baseline score, that a participant would need in order to justify the associated costs and risks of knee replacement as compared to not having a knee replacement. On average, 4.6 items (sd =2.9) were altered to achieve the SWE estimate for pain. A total of 50% of the participants reported that they would need at least a 31 (IQR: 17-42) point improvement in pain on the 100-point KOOS Pain subscale, over their baseline score, to make the costs and risks of a TKR worthwhile as compared to not having surgery. Additionally, 90% of the participants in the sample reported that they would need at least a 57 point improvement in pain, over their baseline pain, to feel that the surgery was worthwhile as compared to not having surgery. When comparing the SWE value for pain to the baseline pain value, on average, 90% participants needed a 77% improvement in pain. Figure 1 displays the SWE estimates for pain.

Figure 4.1 SWE estimates for KOOS Pain Subscale

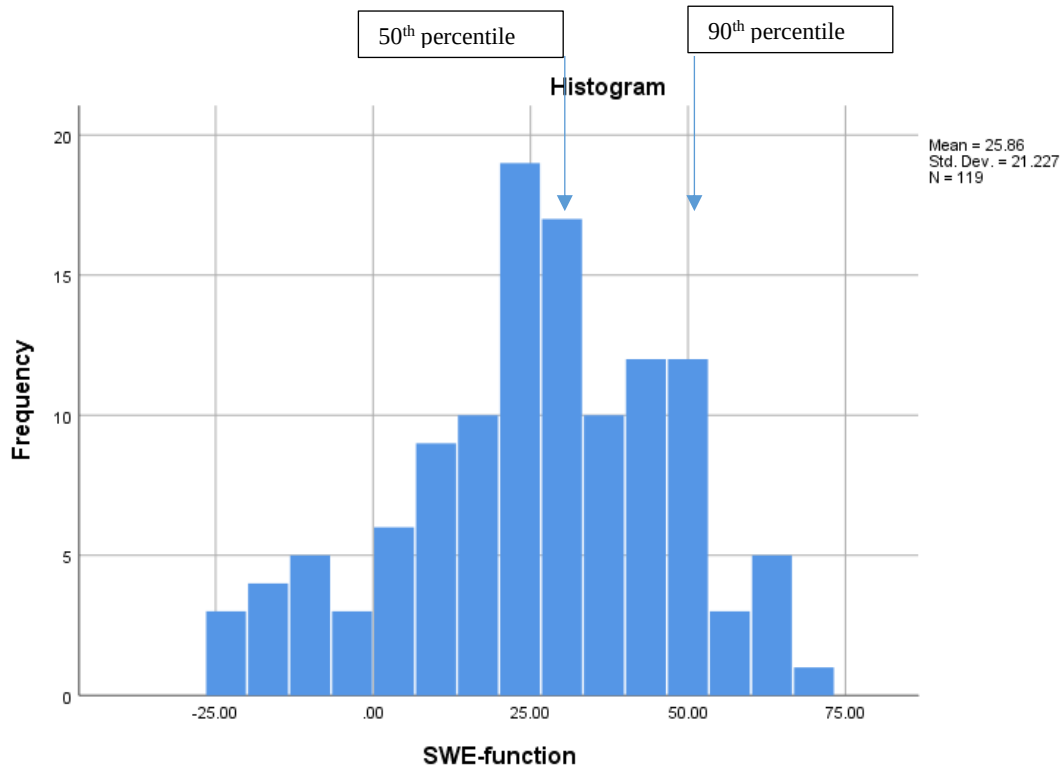


Smallest worthwhile effect for KOOS Function, daily living Subscale

The SWE for function was defined as the lowest amount of improvement on the anticipated 6-month KOOS Function, daily living subscale, over the baseline score, that patients would need in order to justify the associated costs and risks of knee replacement as compared to not having a knee replacement. On average, 6.8 questions (sd =6.9) were altered to achieve the SWE estimate for function. A total of 50% of the sample would need 28 (interquartile range [IQR]: 13-44) points improvement in function on the 100-point KOOS Function, daily living subscale, over their baseline score, in order to feel that the TKR was worth the costs and risks as compared to not having the surgery. Additionally, 90% of the participants in the sample would need a 51 point change in function to view the surgery was worthwhile when compared to not having a knee replacement. When comparing SWE value for function to baseline function, on average, 90% participants needed a 88% improvement in function, over their baseline function,

to feel that the surgery was worthwhile as compared to not having the surgery. Figure 2 displays SWE estimates for the KOOS Function, daily living subscale.

Figure 4.2 SWE estimates for the KOOS Function, daily living Subscale



Regression results for smallest worthwhile effect on pain

A backward multiple linear regression analysis was performed with baseline pain and psychosocial variables as the independent variables and the SWE estimate for pain as the dependent variable. The model with the highest adjusted R^2 value was retained, consistent with previous research (Van Onsem, Van Der Straeten, Arnout, Deprez, Van Damme, & Victor, 2016). Prior to conducting the analysis, all assumptions were analyzed and found to be met. Dummy coding was used for all categorical predictor variables with more than two categories. The following variables were significant at ($p <= 0.05$) and inputted into the backward linear regression: baseline pain, self-efficacy score, anxiety score, pain catastrophizing score,

depressive symptom score, and income (10-25,000$ and >100,000$). The multiple regression yielded seven models with a sequential decrease in number of predictor variables within each model. Model #6 had the highest adjusted R^2 value and consisted of the following predictor variables: baseline pain and anxiety score. The R^2 for this model was 0.499 and the adjusted R^2 was 0.490. An ANOVA was used to assess whether the R^2 value was significantly greater than 0, indicating that the model was a good fit. The ANOVA was statistically significant, ($f=56.2$; $p<0.001$) indicating that the R^2 value was significantly different than 0. Model statistics can be found in Table 4.5. When assessing these predictor variables individually, only baseline pain was statistically significant. Coefficient data can be found in Table 4.6. Baseline pain had a negative relationship with the SWE value as the unstandardized coefficient for baseline pain was -0.77 ($SE=0.08$; $p<0.001$). Baseline pain was statistically significant in all seven models.

Table 4.5. Multiple Regression for SWE for Pain

Regression Model	Variables within the Model	R	R^2	Adjusted R^2
1	Baseline pain, self efficacy, pain catastrophizing, depression, anxiety, income ($10-25,000), income ($100,000)	.715	.511	.479
2	Baseline pain, self efficacy, pain catastrophizing, depression, anxiety, income ($10-25,000))	.714	.510	.483
3	Baseline pain, self efficacy, pain catastrophizing, anxiety, income ($10-25,000)	.713	.509	.486
4	Baseline pain, self efficacy, pain catastrophizing, anxiety	.711	.505	.487
5	Baseline pain, self efficacy, anxiety	.708	.502	.488
6	Baseline pain, anxiety	.706	.499	.490*
7	Baseline pain	.702	.493	.488

*Model with the highest adjusted R^2

Table 4.6. Coefficients for SWE for Pain

Model	Variable	Beta Coefficient	Sd	Significance value
1	Baseline pain	-.759	.092	>0.001*
	Self efficacy	-.079	.081	0.334
	Pain catastrophizing	-.112	.131	0.393
	Depression	-.180	.394	0.648
	Anxiety	.468	.435	0.284
	Income ($10-25K)	5.46	6.31	0.389
	Income (>$100K)	-1.33	3.03	0.661
2	Baseline pain	-.767	.090	<0.001*
	Self efficacy	-.072	-.079	0.369
	Pain catastrophizing	-.108	.130	0.410
	Depression	-.200	.390	0.609
	Anxiety	.493	.429	0.253
	Income ($10-25K)	5.69	6.26	0.366
3	Baseline pain	-.764	.090	<0.001*
	Self efficacy	-.077	.078	0.325
	Pain catastrophizing	-.124	.126	0.325
	Anxiety	.350	.326	0.285
	Income ($10-25K)	5.39	6.22	0.387
4	Baseline pain	-.773	.089	<0.001*
	Self efficacy	-.076	.078	0.335
	Pain catastrophizing	-.109	.124	0.383
	Anxiety	.391	.322	0.228
5	Baseline pain	-.748	.084	<0.001*
	Self efficacy	-.064	.077	0.409
	Anxiety	.272	.292	0.354
6	Baseline pain	-.767	.081	<0.001*
	Anxiety	.328	.284	0.250
7	Baseline pain	-.799	.076	<0.001*

*Statistically significant variable <0.05

Regression for smallest worthwhile effect on function

A backward multiple linear regression analysis was performed with baseline function and psychosocial variables as the independent variables and the SWE estimate for function as the dependent variable. The model with the highest adjusted R^2 value was retained, consistent with previous research (Van Onsem et al, 2016). Prior to conducting the analysis, all assumptions were analyzed and found to be met. Dummy coding was used for all categorical predictor variables with more than two categories. Prior to conducting the multiple regression, univariate analyses were performed. The following variables were significant and inputted into the backward stepwise linear regression: baseline pain, self-efficacy score, anxiety score, pain catastrophizing score, and depressive symptom score. The multiple regression yielded five models with a sequential decrease in number of predictor variables within each model. Model #5 had the highest adjusted R^2 value and only included baseline function. The R^2 and adjusted R^2 for this model was 0.39. An ANOVA was used to assess whether the R^2 value was significantly greater than 0, indicating that the model was a good fit. The ANOVA test was statistically significant, ($f=73.2$; $p<0.001$). Baseline function had a negative relationship with the SWE function value as the unstandardized coefficient for baseline function was -0.70 (SE=0.08; $p<0.001$). Baseline function was a significant predictor in all five models. Coefficient data is presented in table 4.8.

Table 4.7. Multiple Regression for SWE for Function

Regression Model	Variables within the Model	R	R^2	Adjusted R^2
1	Baseline function, self efficacy, pain catastrophizing, anxiety, depression	.605	.366	.377
2	Baseline function, self-efficacy, pain catastrophizing, anxiety	.605	.366	.381
3	Baseline function, catastrophizing, anxiety	.601	.362	.383
4	Baseline function, anxiety	.597	.356	.383
5	Baseline function	.594	.353	.386*

*Model with the highest adjusted R^2

Table 4.8. Coefficients for SWE for Function

Model	Variable	Beta Coefficient	Sd	Significance value
1	Baseline function	-.748	.113	>0.001*
	Self efficacy	-.081	.100	0.417
	Pain catastrophizing	-.176	.163	0.283
	Depression	-.015	.490	0.976
	Anxiety	.428	.540	0.430
2	Baseline function	-.748	.112	<0.001*
	Self efficacy	-.082	-.099	0.408
	Pain catastrophizing	-.177	.157	0.260
	Anxiety	.417	.406	0.307
3	Baseline function	-.767	.110	<0.001*
	Pain catastrophizing	-.155	.154	0.318
	Anxiety	.462	.402	0.253
4	Baseline function	-.725	.102	<0.001*
	Anxiety	.279	.358	0.439
5	Baseline function	-.752	.095	<0.001*

*Statistically significant variable <0.05

Descriptive Statistics for 6-month Follow-up Sample

A total of 83 of the 121 participants were successfully contacted for the 6-month follow-up assessment (68% response rate). Of the 83 participants, (1) person reported that she did not undergo a TKR, as previously scheduled, and was therefore removed from the 6-month assessment data. For the remaining (82) participants, the median 6-month KOOS Pain score was 94 (range = 22-100) and 6-month KOOS function, daily living scores was 97(range = 31-100) points. Using the Kolmorgov-Smirnov test, 6-month KOOS pain and function, daily living subscales were not normally distributed; therefore, the median and interquartile range was used to describe the data. A total of 5% (4 participants) of the sample had to be re-hospitalized after

discharge due to post-operative complications, and these included kidney infection (n = 1), knee infection, low potassium, and (1) manipulation under anesthesia.

Using the Patient Acceptable Symptom State (PASS) questionnaire, 91% of participants (n = 75) reported that they were satisfied with their current state six months after surgery. When comparing KOOS pain subscale scores for participants who were satisfied with their current state to those who were not, the median value for the KOOS Pain subscale was 100 points for those who were satisfied versus 51 points for those who were not. Results for the KOOS Function, daily living subscale were similar with a median of 99 points for those who were satisfied versus 53 points for those who were not.

A Pearson's chi-square analysis was performed to assess whether those who met or exceeded their SWE value for pain was correlated with whether or not they were satisfied with their current state. Of those participants who met or exceeded their SWE estimates for pain, 97% (66 participants) were also satisfied with their current state. There was a significant association between meeting/exceeding SWE estimate for pain and satisfaction with current state $X^2=15.9$, $p<0.001$.

A Pearson's chi square analysis was also conducted for the KOOS Function, daily living subscale. Similar results were found with this subscale with 99% (70 participants) who met their SWE estimate for function also being satisfied with their current state. There was a significant association between meeting/exceeding SWE estimate for function and satisfaction with current state $X^2=30.1$, $p<0.001$.

CHAPTER 5
Discussion

This chapter provides a discussion and interpretation of the results of this study of the smallest worthwhile effect (SWE) values for the Knee Injury and Osteoarthritis Outcome Score (KOOS) Pain and Function, daily living subscales. Additionally, the study investigated the association between baseline scores on a variety of variables and SWE estimates for KOOS Pain and Function, daily living subscales. Lastly, the study determined whether initial SWE estimates for KOOS Pain and Function, daily living subscales met or exceeded actual KOOS Pain and Function, daily living subscales six months after a Total Knee Replacement (TKR), as well as, whether participants were satisfied with their outcome six months after surgery. This research fills a gap in the literature as prior research has not attempted to determine meaningful change, using the benefit-harm trade-off method, after a TKR.

This chapter is organized as follows: the first section discusses SWE estimates for the KOOS Pain and Function, daily living subscales and compares them with meaningful change estimates from the literature. The next section includes a comparison of demographic variables in the current study with previous research. The third section explores the implications of associations between baseline scores on psychosocial and demographic variables and SWE estimates. The fourth section compares initial SWE estimates with KOOS Pain and Function, daily living scores six months after surgery. This section also discusses the participants postoperative satisfaction rates. Limitations and recommendations for future research are discussed in the final section.

SWE estimates for KOOS Pain and Function, daily living subscales

The benefit-harm trade-off method indicated that 90% of participants needed a change of up to 58 points on the 100-point KOOS Pain subscale, to conclude that the benefits of a TKR outweighed the associated costs and risks when compared to not having the surgery. Similarly, 90% of the participants cited a change of up to 51 points on the 100-point KOOS Function, daily living subscale. The benefit-harm trade-off method was used to estimate these SWE values. Alternate methods used to determine change on an outcome measure include the distribution- and anchor-based methods. The distribution-based method provides information about minimum change on a self-report outcome measure, or a measurement tool, that is likely due to measurement error (Gatchel, Lurie, & Mayer, 2010). Using the distribution-based method, participants are not provided an opportunity to determine the meaningful change and there is no consideration given to the potential costs and risks associated with the intervention. (McNamara, Elkins, Ferreira, Spencer, & Herbert, 2015). The distribution-based method assesses magnitude of change, but does not address whether the change is clinically meaningful to the patient (Gatchel, et al., 2010).

The anchor-based method uses an external criterion to calculate meaningful change (Norman, Sridhar, Guyatt, &Walter, 2001). The anchor-based method, unlike the distribution-based method, does provide change estimates which are clinically meaningful. When using the anchor-based method, the first step is to use the scores on an external criterion as a cut-off to determine which patients experienced a meaningful change. Next, the magnitude of improvement (pre-post intervention change scores) reported on the patient reported outcome (PRO) measure, for those patients who were categorized as having a meaningful change on the external criterion, is then calculated and serves as the response estimate of meaningful change. The validity of this method hinges on the accuracy of the external criterion for determining

which patients experienced meaningful change. The most commonly used external criterion is the Global Rating of Change (GROC) which has several limitations, most notably is that it is a retrospective measure which is subject to substantial recall bias (Norman et al., 2001; Garrison & Cook, 2012; Schmitt & Di Fabio, 2005). Because of this recall bias, scores on the GROC correlate more highly with a patient's current status versus their baseline status, (Norman, Stratford, & Regehr, 1997), which likely diminishes it's accuracy. The diminished accuracy of the GROC for determining meaningful change negatively impacts the validity of change estimates derived using the anchor-based method.

When compared to the anchor- and distribution-based methods, the benefit-harm trade-off method has improved upon several of the limitations of these two methods. One improvement is that the benefit-harm trade-off method assesses meaningful change based on the participant's weighing of the potential harms of an intervention against it's benefits in relation to not undergoing the intervention. The other two methods assess change over time rather than the difference with and without the intervention. Assessing change over time is a limitation because several other factors, other than the intervention, may influence that change (Ferreira, Herbert, Ferreira, Latimer, Ostelo, Grotle, & Barrett, 2013; Herbert, Mead, & Hagen, 2005). The benefit-harm trade-off method also varies from the anchor-based method because it is intervention specific. Because each intervention has different associated costs and risks, the response measure should be linked to the actual intervention versus to the outcome measure, as with the anchor-based method (Ferreira, Ferreira, Herbert, & Latimer, 2009). Since the benefit-harm trade-off method improves upon several limitations of the other two methods, it may be a more accurate way to calculate meaningful change.

The SWE estimates in the current study are substantially larger as compared to response measures calculated using distribution- and anchor-based approaches associated with a TKR (Lyman, Lee, McLawhorn, Islam, & MacLean, 2018; Berliner, Brodke, Chan, SooHoo, & Bozic, 2017). A 2018 retrospective study by Lyman calculated the minimum clinically important difference (MCID) values for the KOOS Pain and Function, daily living subscales after undergoing a TKR. The authors assessed meaningful change using both the distribution- and anchor-based methods. With the anchor-based method, the authors used the Quality of Life (QOL) item on the HSS satisfaction survey as the external criterion. The responses on the QOL item were dichotomized into the following two groups: those who answered "moderate improvement" and those who answered "a little improvement" or "no improvement." Once the responses were categorized, the authors concluded that meaningful change had occurred with the group who answered "moderate improvement" on the QOL item. Using the distribution method, the researchers calculated the MCID values as 8 points for the KOOS Pain subscale and 9 points for the KOOS Function, daily living subscale for those individuals undergoing a total joint replacement. Using the anchor-based method, MCID values were 18 and 16 points, respectively. Similarly, Berliner et al calculated MCID values using a distribution-based method, defined as half of the standard deviation of outcome change scores, and found the MCID for the aggregate score on the KOOS was 10 points after a TKR (Berliner et al., 2017). The SWE estimates in the current study are larger than the MCID values in the previous two studies, which is likely attributed to the differences in derivation methods. Because the SWE estimates in the current study were determined using the benefit-harm trade-off method, which improves upon several limitations of the other derivation methods, they may portray a more accurate reflection of meaningful change after a TKR. The SWE estimates generated in the current study indicate that

patients actually require more improvement in pain and function after a TKR, compared to what was previously thought, to experience a meaningful change. Surgeons can use these SWE estimates to establish more realistic and accurate patient expectations prior to a TKR.

Previous studies which assessed MCID values, calculated using both the distribution and anchor-based methods, for the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) demonstrated that the distribution-based method consistently underestimated MCID values when compared to anchor-based methods (Chesworth, Mahomed, Bourne, & Davis, 2008; Escobar, Quintana, Bilbao, Arostegui, Lafuente, Vidaurreta, 2007; SooHoo, Li, Chenok, & Bozic, 2015). Clement et al recently examined MCID values for the WOMAC in a large cohort of patients one year status post TKR. In the study, the MCID value for the WOMAC function scale was 9 points, which is the same value that Lyman et al found for the KOOS Function subscale using a distribution-based method (Clement, Bardgett, Weir, Holland, Gerrand, & Deehan, 2018; Lyman et al., 2018). A 2019 systematic review assessed MCID values for the WOMAC in patients who underwent a TKR. These authors found a wide variety of MCID values for the WOMAC function subscale, ranging from 1.8-33.0 points. The authors reported that not only did they find variability in the different methods used to calculate MCID values, but also within each of the methodologies themselves. This variability in methodology likely contributed to the large range of MCID values for the WOMAC function subscale (MacKay, Clements, Wong, & Davis, 2019). Authors in a recent systematic review summarized articles which calculated MCID values, using an anchor-based approach, for commonly used lower extremity PRO measures (Celik, Coban, & Kilicoglu, 2019). In these studies, the MCID values for the KOOS Pain subscale ranged from 11.5-16.7 points and 8.1-18.2 for the KOOS Function, daily living subscale (Harris, Dawson, Jones, Beard, & Price, 2013; Huang, Chen, Tsai, & Wang,

2017; Mills, Naylor, Eyles, Roos, & Hunter, 2016; Monticone, Ferrante, Salvaderi, Motta, & Cerri, 2013).

The MCID values reported in the McKay and Celik studies had considerable variability. The large variability in the MCID estimates may indicate that this response measure does not accurately reflect meaningful change. Not only are there multiple methods used to determine the MCID values, but there are also several names used to refer minimal clinically important values (Cepeda, Polascik, & Ling, 2019).

Considering the SWE values in the current study were higher than previously established MCID values for the KOOS and WOMAC, it could signify that both distribution- and anchor-based methods underestimate meaningful change. When asked to weigh the costs and risks associated with an intervention against the potential benefits, people are more inclined to require a greater level of improvement.

Considering some of the methodological concerns associated with the distribution and anchor-based approaches, the benefit-harm trade-off method may be a more accurate method to assess meaningful change. The benefit-harm trade-off method allows patients to determine meaningful change by assessing the associated benefits and risks of the intervention and allows for comparison of values before and after the intervention to determine if meaningful change has actually occurred after an intervention. It is well-known that the decision- making process to undergo a TKR is complex. The results of this study could have usefulness in this decision-making process as the SWE value potentially provides patients and medical providers with a more accurate assessment of TKR effectiveness and degree of meaningful change needed after the intervention. Using the SWE estimates, patients will better understand the magnitude of meaningful change that

may occur after the surgery. Several studies have assessed the influence of meeting preoperative expectations and satisfaction after surgery. These studies concluded that meeting preoperative expectations is positively linked to patient satisfaction after a TKR (Riddle, Golladay, Hayes, & Ghomrawi, 2017; Gandhi, Davey, & Mahomed, 2009). By having a more accurate estimate of meaningful change and preoperative expectations, satisfaction after surgery may improve. Additionally, these results could influence future randomized controlled trials (RCTs) by providing a valid change value to include in sample size calculations to detect meaningful change of the TKR.

Comparison of demographic and baseline variables

The demographic breakdown in the current study was similar to that of other studies; thereby, increasing the generalizability of the SWE findings. The majority of the participants in the current study were female and of Medicare age, which is consistent with several previous studies (Hamel, Toth, Legedza, & Rosen, 2008; Van Onsem, Van Der Straeten, Arnout, Deprez, Van Damme, & Victor, 2016; Bourne, Chesworth, Davis, Mahomed, & Charron, 2010; Berliner, Brodke, Chan, SooHoo, & Bozic, 2017; Maratt, Lee, Lyman, & Westrich, 2015; Mahomed, Liang, Cook, Daltroy, Fortin, Fossel, & Katz, 2002). Additionally, a study by Maratt et al had a similar educational breakdown compared to the current study. The Maratt study included a sample of participants, 48.4% of whom had a higher level of education, defined as having had some college, if not a college degree (Maratt et al., 2015). The current study was similar with 51.2% of the participants having had at least some college or a college degree.

When reviewing the preoperative expectations in the current study, 74% of the study expected to have no pain at six months and 61% expected to have no functional limitations. These values are consistent with previous literature (Mahomed et al. 2002; Nilsdotter, Toksvig-

Larsen, & Roos, 2009). A 2002 study by Mahomed et al, 76% of patients expected to have no pain after recovery from a TKR and 40% expected to have no functional limitations (Mahomed et al, 2002). In a 2009 study by Nillsdotter et al., the authors found that 98% of patients, when questioned preoperatively, expected much less or less pain postoperatively ((Nilsdotter et al., 2009). In the previous two studies, as well as the current study, participants expected to have greater pain relief compared to fewer functional limitations. However, when looking at the six month results in the current study, the median scores on the KOOS Function, daily living score (97 points) was higher than that of the KOOS Pain score (94 points).

When looking at baseline pain and function values, average baseline scores on the KOOS for Pain and Function, daily living subscales (43 and 45 points, respectively) were found to be consistent with previous literature (Lyman, Lee, McLawhorn, Islam, & MacLean, 2018; Berliner et al., 2017; Uyen-Sa, Ayers, Li, Harrold, & Franklin, 2016; Nilsdotter, Toksvig-Larsen, & Roos, 2009). The fact that the baseline scores are similar to previous studies helps to improve the generalizability of the SWE findings. The authors in a 2018 study by Lyman et al calculated baseline KOOS scores, prior to a TKR, as 51 points for the pain subscale and 55 points for the function, daily living subscale (Lymann et al., 2018). Likewise, in the Nilsdotter study, the average preoperative, baseline score was 40 points for the pain subscale and 50 points for the function, daily living subscale. Additionally, Berliner et al determined the average preoperative KOOS score for all five subscales was 50 points (Berliner et al., 2017).

Sensitivity Analysis for the KOOS Pain Subscale

While exploring the data, negative SWE values were found for some participants. A post-hoc analysis revealed that 4.2% (n=5) of the sample reported negative SWE values for the KOOS Pain subscale, indicating that they were willing to have an increase in pain at 6 months after

surgery compared to their baseline pain score. A sensitivity analysis was conducted to examine the influence of the negative SWE estimates of KOOS Pain on the overall study results. When looking solely at SWE values for those individuals who reported some improvement, as indicated by positive SWE values, 50% of the participants needed up to a 33 point (17-42) improvement in pain, over their baseline score, to feel that their surgery was worthwhile when compared to not having the surgery. Additionally, 90% of participants, with positive SWE estimates, needed up to a 58 point (17-42) improvement in pain scores. Comparing these SWE estimates to the entire sample, there was minimal difference (1-2 points) between those individuals who had only positive SWE estimates and the entire sample. This finding indicates that the presence of the (5) participants with negative SWE estimates for the KOOS Pain subscale did not impact the SWE estimates for the overall sample.

Baseline scores were also analyzed for the sample who reported only positive SWE estimates for the KOOS Pain subscale. No significant differences were noted for all baseline demographic and psychosocial variables between the entire sample and those participants who only reported positive SWE estimates. See tables 5.1 and 5.2 for baseline data for those participants who reported positive SWE estimates for the KOOS Pain subscale. Similar to the comparative findings for the SWE estimates, these results indicate that the sample demographics were not affected by the presence of the negative SWE estimates.

Table 5.1. Participant characteristics at baseline for those who only reported positive SWE estimates for the KOOS Pain subscale

Sample Characteristics	N=116
Age (yr), *mean (sd)*	67 (7.7)
Gender (women), *n* (%)	63 (55)
Educational level, *n* (%)	
Less than high school graduate	3 (2.6)
High school degree	14 (12.3)
Some college	27 (23.7)

College degree	31 (27.2)
Some graduate school	5 (4.4)
Graduate degree	34 (29.8)
Yearly household income, *n* (%)	
<$10K	3 (2.7)
$10-<25K	6 (5.4)
$25-50K	22 (19.8)
$50-100K	47 (42.3)
$100K or greater	33 (29.7)
KOOS Pain Score (0-100), *mean (sd)*	43 (16.6)
Anticipated pain severity at 6-months (0-100), *median* (IQR)	89 (53-100)
Depressive symptoms (0-24), *median* (IQR)	3 (0-22)
Anxiety (0-21), *median* (IQR)	4 (0-21)
Pain catastrophizing (0-52), *median* (IQR)	14 (0-52)
Self efficacy (8-80), *median* (IQR)	26 (8-80)

Knee Injury and Osteoarthritis Score (KOOS) pain subscale- baseline pain (0-100 with lower numbers indicating greater pain)
Generalized Anxiety Disorder (GAD)- baseline anxiety (0-21 with higher numbers indicating greater feelings of anxiety)
Personal Health Questionnaire (PHQ) Depression Scale- depressive feelings (0--24 with higher numbers indicating greater feelings of depression)
Pain Catastrophizing Scale (PCS)- pain catastrophizing feelings (0-52 with higher numbers indicating greater feelings of pain catastrophizing)
Arthritis Self-Efficacy Scale (ASES)- feelings of self-efficacy (10-80 with higher numbers indicating greater levels of self-efficacy)

Table 5.2. Patient Expectations at Baseline for those who only reported positive SWE estimates for the KOOS Pain subscale

Sample Characteristics	*n* (%)
Pain Expectation	
Not at all painful	82 (74.5)
Slightly painful	24 (21.8)
Moderately painful	3 (2.7)
Very painful	1 (.9)
Function Expectation	
Not at all limited	67 (60.9)
Slightly limited	39 (35.5)
Moderately limited	3 (2.7)
Very limited	1 (.9)
Visual Analog Scale (0-100), *mean* (IQR)	
Likelihood of postoperative success	94 (67-100)
Likelihood of postoperative complication	8 (0-33)

Pain Expectations: How painful do you expect your knee to be in one year?
Function Expectations: How limited do you expect to be in your usual activities in one year?

A regression analysis was also repeated to assess the influence of baseline demographic and psychosocial variables on SWE estimates for the KOOS Pain subscale, excluding those participants who had negative SWE estimates. This analysis generated similar results to the overall sample with baseline pain being the only significant predictor of SWE estimates, which indicates that those participants who reported negative SWE estimates did not influence the overall study results for the regression analysis. Table 5.3 includes the regression model results and coefficient results are found in table 5.4.

Table 5.3. Multiple Regression for SWE for KOOS Pain Subscale for those who only reported positive SWE estimates

Regression Model	Variables within the Model	R	R^2	Adjusted R^2
1	Baseline pain, self efficacy, pain catastrophizing, depression, anxiety, income ($10-25,000), income ($100,000)	.703	.495	.460
2	Baseline pain, self efficacy, pain catastrophizing, depression, anxiety, income ($10-25,000))	.703	.494	.465
3	Baseline pain, self efficacy, pain catastrophizing, anxiety, income ($10-25,000)	.701	.491	.466
4	Baseline pain, self efficacy, pain catastrophizing, anxiety	.699	.488	.469
5	Baseline pain, self efficacy, anxiety	.697	.485	.470
6	Baseline pain, anxiety	.694	.481	.472*
7	Baseline pain	.689	.475	.470

*Model with the highest adjusted R^2

Table 5.4. Coefficients for SWE for KOOS- Pain Subscale for those who only reported positive SWE estimates

Model	Variable	Beta Coefficient	Sd	Significance value
1	Baseline pain	-.692	.090	>0.001*
	Self efficacy	-.064	.077	0.409
	Pain catastrophizing	-.112	.123	0.365
	Depression	-.071	.364	0.846
	Anxiety	.366	.401	0.364

	Income ($10-25K)	5.44	5.85	0.355
	Income (>$100K)	-2.33	2.89	0.421
2	Baseline pain	-.690	.089	<0.001*
	Self efficacy	-.066	.076	0.386
	Pain catastrophizing	-.118	.119	0.323
	Income (>$100K)	-2.39	2.86	0.406
	Anxiety	.316	.307	0.307
	Income ($10-25K)	5.33	5.79	0.359
3	Baseline pain	-.705	.087	<0.001*
	Self efficacy	-.053	.074	0.474
	Pain catastrophizing	-.111	.118	0.349
	Anxiety	.335	.306	0.276
	Income ($10-25K)	5.68	5.77	0.328
4	Baseline pain	-.714	.086	<0.001*
	Income ($10-25K)	5.64	5.76	0.330
	Pain catastrophizing	-.093	.115	0.420
	Anxiety	.358	.304	0.242
5	Baseline pain	-.692	.081	<0.001*
	Income ($10-25K)	4.98	5.69	0.321
	Anxiety	.252	.274	0.384
6	Baseline pain	-.706	.079	<0.001*
	Anxiety	.304	.267	0.258
7	Baseline pain	-.736	.075	<0.001*

*Statistically significant variable <0.05

Sensitivity Analysis for the KOOS Function, daily living Subscale

Similar to the SWE estimates for the KOOS Pain Subscale, 12.5% (n=15) reported negative SWE values for the KOOS Function, daily living subscale; therefore, a sensitivity analysis was also conducted to examine the influence of the negative SWE estimates on overall study results. For SWE values for function, 50% of the participants, with positive SWE values, needed up to a 29 point improvement in function, over their baseline score, to feel that their surgery was worthwhile compared to not having the surgery. Moreover, 90% of participants, with positive SWE values, needed up to a 53 point improvement in function scores. The SWE estimates for the KOOS Function, daily living subscale for those participants with only positive SWE values were 1-2 points greater, indicating a slight worsening of baseline function,

compared to estimates of the entire sample. This change in SWE estimates was so minimal that it likely did not impact the overall sample.

When comparing baseline scores of the entire sample to those participants who only had positive SWE estimates for the KOOS Function, daily living subscale, there were no significant differences. This finding indicates that the presence of negative SWE estimates for the KOOS Function, daily living subscale scores did not affect the overall study results. See tables 5.5 and 5.6 for baseline data for those participants who reported positive SWE estimates for the KOOS Function, daily living subscale.

Table 5.5. Participant characteristics at baseline for those who only reported positive SWE estimates for the KOOS Function, daily living subscale

Sample Characteristics	N=105
Age (yr), *mean (sd)*	67 (9.7)
Gender (women), *n* (%)	60 (57)
Educational level, *n* (%)	
Less than high school graduate	3 (2.9)
High school degree	12 (11.4)
Some college	26 (24.8)
College degree	29 (27.6)
Some graduate school	5 (4.8)
Graduate degree	30 (28.6)
Yearly household income, *n* (%)	
<$10K	3 (2.9)
$10-<25K	6 (5.9)
$25-50K	20 (19.6)
$50-100K	43 (42.2)
$100K or greater	30 (29.4)
KOOS Function, daily living Score (0-100), *mean (sd)*	43 (17.8)
Anticipated disability level at 6-months (0-100), *median* (IQR)	82 (50-100)
Depressive symptoms (0-24), *median* (IQR)	5 (0-22)
Anxiety (0-21), *median* (IQR)	5 (0-21)
Pain catastrophizing (0-52), *median* (IQR)	13 (0-52)
Self efficacy (8-80), *median* (IQR)	29 (8-80)

Knee Injury and Osteoarthritis Score (KOOS) function subscale- baseline function (0-100 with lower numbers indicating greater difficulty with functional activities)
Generalized Anxiety Disorder (GAD)- baseline anxiety (0-21 with higher numbers indicating greater feelings of anxiety)

Personal Health Questionnaire (PHQ) Depression Scale- depressive feelings (0--24 with higher numbers indicating greater feelings of depression)
Pain Catastrophizing Scale (PCS)- pain catastrophizing feelings (0-52 with higher numbers indicating greater feelings of pain catastrophizing)
Arthritis Self-Efficacy Scale (ASES)- feelings of self-efficacy (10-80 with higher numbers indicating greater levels of self-efficacy)

Table 5.6. Patient Expectations at Baseline for those who only reported positive SWE estimates for the KOOS Function, daily living subscale

Sample Characteristics	*n* (%)
Pain Expectation	
Not at all painful	76 (75.2)
Slightly painful	21 (20.8)
Moderately painful	3 (3.0)
Very painful	1 (1.0)
Function Expectation	
Not at all limited	59 (58.4)
Slightly limited	38 (37.6)
Moderately limited	3 (3.0)
Very limited	1 (1.0)
Visual Analog Scale (0-100), *mean* (IQR)	
Likelihood of postoperative success	94 (75-100)
Likelihood of postoperative complication	8 (0-30)

Pain Expectations: How painful do you expect your knee to be in one year?
Function Expectations: How limited do you expect to be in your usual activities in one year?

A regression analysis was also repeated, excluding participants who had negative SWE estimates for the KOOS Function, daily living subscale, to assess the influence of baseline demographic and psychosocial variables on SWE estimates. These results were similar to that of the overall sample with baseline function being the only significant predictor of SWE estimates. This finding further supports the idea that the negative SWE estimates did not influence the overall study results. Table 5.7 includes the regression model results and coefficient results are found in table 5.8.

Table 5.7. Multiple Regression for SWE for KOOS Function, daily living Subscale for those who only reported positive SWE estimates

Regression Model	Variables within the Model	R	R^2	Adjusted R^2
1	Baseline function, self efficacy, pain catastrophizing, anxiety, depression	.491	.241	.201

2	Baseline function, self-efficacy, pain catastrophizing, anxiety	.491	.241	.209
3	Baseline function, catastrophizing, anxiety	.490	.240	.217
4	Baseline function, anxiety	.484	.234	.219
5	Baseline function	.479	.230	.222*

*Model with the highest adjusted R^2

Table 5.8. Coefficients for SWE for the KOOS Function, daily living subscale for those who only reported positive SWE estimates

Model	Variable	Beta Coefficient	Sd	Significance value
1	Baseline function	-.445	.098	>0.001*
	Self efficacy	-.085	.092	0.358
	Pain catastrophizing	-.001	.149	0.997
	Depression	-.093	.432	0.829
	Anxiety	.352	.467	0.453
2	Baseline function	-.445	.090	<0.001*
	Self efficacy	-.085	-.090	0.349
	Depression	-.094	.417	0.822
	Anxiety	.351	.456	0.442
3	Baseline function	-.439	.087	<0.001*
	Self efficacy	.084	.090	0.352
	Anxiety	.277	.314	0.379
4	Baseline function	-.456	.084	<0.001*
	Self efficacy	.067	.087	0.447
5	Baseline function	-.437	.080	<0.001*

*Statistically significant variable <0.05

The presence of negative SWE estimates was an unexpected finding. This finding is likely because some of the participants did not fully understand the SWE concept. Those participants with negative values essentially indicated that they would be willing to have more pain and worse function after surgery as compared to baseline. The intent of the standardized script, used to determine the SWE estimates, was to determine what degree of improvement was necessary, over the baseline score, to outweigh the associated risks and complications of the surgery. It is highly probable that these participants did not understand what was asked of them considering they indicated they would be willing to have more pain and worse function after surgery as compared to the pre-surgery pain and functional status. During data collection, when

it appeared that the participant did not understand the SWE concept, the researcher re-read the standardized script, including the examples, to reiterate what was being asked. However, the researcher did not vary or add additional wording outside of the standardized script.

Associations between SWE estimates of KOOS Pain and Function, daily living subscales and baseline predictors

Baseline pain was the strongest predictor of SWE estimates for the KOOS Pain subscale. Similar to the SWE estimates for the KOOS Pain subscale, baseline function was also the strongest predictor of SWE estimates for the KOOS Function, daily living subscale. Several previous studies also found that baseline pain and function are significant predictors of meaningful change (Ferreira et al., 2009; Berliner, Brodke, Chan, SooHoo, & Bozic, 2016; Berliner et al., 2017; Fortin, Clarke, Joseph, Liang, Tanzer, Ferland, Phillips….., et al, 1999).

The only other study that used the benefit-harm trade-off method to determine SWE estimates for an intervention was a 2009 study by Ferreira et al. The authors in this study found that initial symptom severity was the only significant predictor of SWE estimates in patients with low back pain. This study calculated SWE estimates for five different Physiotherapy interventions for low back pain. These authors concluded that initial symptom severity explained 9% of the total variance in SWE estimates for low back pain (Ferreira et al., 2009). In the current study, baseline pain and function were even stronger predictors of SWE estimates, explaining >40% of the variance in SWE values for KOOS Pain and Function, daily living subscales. The likely difference in baseline pain and function for a person undergoing a TKR versus a person with low back pain may explain why these variables were stronger predictors in the current study.

Similar to SWE values, baseline pain and function also predict MCID values after a TKR. A 2017 study, conducted by Berliner et al, found that those persons with KOOS Function scores above 58, indicating better preoperative function, were less likely to experience a clinically meaningful change, as determined by the MCID value, status post TKR (Berliner et al., 2017). Additionally, Fortin et al found that the single best predictor of pain and function at 6 months after a TKR was the person's baseline pain and function (Fortin et al., 1999). The results of the current study are similar to those of the previous three studies where baseline pain and function were the strongest predictor of meaningful change.

A post hoc analysis categorized baseline pain values, for only those participants who reported positive SWE estimates, and associated SWE values, into three tertiles. As baseline values decreased, SWE values increased. As baseline pain scores increased between the 1st and 2nd tertiles, indicating less baseline pain, there was a 29% decrease in the SWE estimates, indicating less change after the TKR was required to be clinically meaningful. Similarly, there was a 40% decrease in SWE estimates, between 2nd and 3rd tertiles, as baseline pain scores increased. See table 5.9 for tertile data for the KOOS Pain subscale. Both the apriori and post-hoc analyses demonstrated that persons with higher pain levels before surgery required more benefit after a TKR to justify associated costs and risks when compared to not having the surgery.

Table 5.9. SWE estimates for pain and baseline pain scores

Baseline scores, *range of scores*	1st tertile (0-36)	2nd tertile (37-50)	3rd tertile (50-75)
SWE estimates, *mean (sd)*	43.5 (17.0)	30.8 (13.5)	18.5 (11.7)

Knee Injury and Osteoarthritis Score (KOOS) Pain subscale- baseline pain (0-100 with lower numbers indicating greater pain)

A similar post hoc analysis was also calculated for baseline KOOS Function, daily living subscale scores, after excluding participants who reported negative SWE estimates. The analysis revealed the same results with SWE values increasing while baseline function values decreased. This finding demonstrates that those individuals with worse KOOS Function, daily activity scores before surgery required more benefit in order to feel that the surgery was worth the associated costs and risks. There was a 13% decrease in SWE estimates, between the 1st and 2nd tertile, as baseline function scores increased. Similarly, there was a 32% decrease in SWE estimates, between 2nd and 3rd tertiles, as baseline function scores increased. See table 5.10 for tertile data for function.

Table 5.10. SWE estimates for function and baseline KOOS function, daily living scale scores

Baseline KOOS scores, *range of scores*	1st tertile (0-37) baseline	2nd tertile (37-47)	3rd tertile (50-81)
SWE estimates, *mean (sd)*	38.2 (18.5)	33.3 (12.9)	22.8 (13.1)

Knee Injury and Osteoarthritis Score (KOOS) Function subscale- baseline function (0-100 with lower numbers indicating greater difficulty with functional activities)

Overall, these post hoc analyses demonstrate that persons with higher baseline pain and worse function require more improvement after surgery. These data may suggest that the knee pain and function in these individuals have worsened to the point that the participant would need more improvement after surgery to return to a quality of life that is satisfactory for them. Whereas, those with less pain and higher levels of function pre-operatively may already have a higher quality of life, as determined by pain severity and ability to perform functional activities, and therefore not require as much improvement in order to return to the activities that they would like to perform. These results are also important in terms of patient expectations. Is it important for both surgeons and patients to understand that persons with higher baseline pain and worse function require more improvement to feel that they had a clinically meaningful improvement.

Comparison of initial and 6-month SWE estimates for KOOS Pain and Function, daily living subscales

At six months post-surgery the vast majority of participants (91.5%) were satisfied with their current state after their TKR. Most studies found satisfaction rates between 80-90% six months after surgery despite using a variety of satisfaction measures (Bourne et al., 2010; Judge, Arden, Kiran, Price, Javaid, Beard, Murray, & Field, 2012; Williams, Price, Beard, Hadfield, Arden, Murray, & Field, 2013; Woolhead, Donovan, & Dieppe, 2005; Van Onsem et al., 2016).

When assessing the relationship between meeting or exceeding SWE estimates and satisfaction with current state, there was a significant association (X^2=30.1, p<0.001). This finding indicates that the SWE estimate is closely associated with patient satisfaction after a TKR, which strengthens the argument for the utility of the SWE estimate and the benefit-harm trade-off method. It is likely that this method is a reasonably accurate way to assess effects that patients deem worthwhile, prior to undergoing TKR. Additionally, this result is clinically intuitive because those persons who achieved at least the smallest amount of improvement that they felt was necessary to justify the costs and risks would be more likely be more satisfied at six months after surgery.

Previous studies have shown that 70-80% of patients after a TKR report functional improvement (Callahan, Drake, Heck, & Dittus, 1994; Jones, Voaklander, Johnston, Suarez-& Almazor, 2000; Jones, Voaklandr, & Suarez-Almazor, 2003). A lower percentage of patients experience pain relief after surgery. Beswick et al found that up to 34% of patients continue to experience long term pain after this surgery (Beswick, Wylde, Gooberman-Hill, & Blom, 2012). The median score on the KOOS Pain subscale 6-months post-surgery was 94 points and 97 points for the Function, daily living subscales. The baseline KOOS for these two subscales were 43 and 45 points, respectively. These values reflect a 114% improvement in pain and a 111%

increase in function, compared to baseline KOOS Pain and Function, daily living scores. These numbers are quite higher than several previous studies which found an approximate 50% improvement can be expected 3-6 months after a TKR (Bachmeier, March, Lapsley, Tribe, Courtenay & Brooks, 2001; Nillsdotter, Toksvig-Larsen, & Roos, 2009; Judge, Arden, Cooper, Javaid, Carr, Field & Diepp, 2012; Davis, Perruccio, Ibrahim, Hogg-Johnson, Wong, Streiner, Beaton….., et al., 2011).

Limitations

There were several limitations in this study. The most notable limitation is the presence of negative SWE estimates for the KOOS Pain subscale for 5 participants and 15 participants for KOOS Function, daily living scores. The presence of negative SWE estimates indicates that these participants are willing to have more pain and less function after surgery compared to baseline. Since this is unlikely, it is probable that those with negative SWE estimates did not understand that the intent of the script was to ascertain the lowest amount of improvement that would be needed, over baseline values, to justify the associated costs and risks of the TKR.

Another limitation is that the standardized script that was used to estimate SWE values did not account for patient perceived importance of the activities on the KOOS subscales. As part of the script, the participant completed the KOOS Pain and Function, daily living subscales, indicating the lowest degree of pain or difficulty with functional tasks they would be willing to have at 6 months after surgery to justify the associated costs and risks of the TKR. Next, the investigator progressively changed the item scores, indicating greater pain or difficulty with tasks, until the participant reported that the amount of benefit would no longer be worth the costs and risks. When changing the item scores, the investigator began with the question which had the best score. However, when there were multiple items with the same score, the question which

was listed first on the KOOS subscale was initially asked, followed by the other items. This order of questioning may have influenced the SWE estimates. For example, if a patient indicated "no pain" on both walking on flat surface and going up and down stairs, the investigator would first change the score on the walking on flat surface item, since it appears first on the KOOS subscale, followed by going up and down stairs. Some participants, for example, did not negotiate stairs on a regular basis and may not have been concerned with having greater pain with this activity; whereas other individuals lived in a two story home and may not have been willing to have increased pain with this activity. This difference in perception of task importance may have influenced the SWE estimates.

The substantial loss to follow-up is another limitation. Of the 121 participants who completed the baseline assessment, n = 82 participants (68%) completed the six month follow-up assessment. A comparative analysis of baseline and psychosocial variables demonstrated that those lost to follow-up had significantly lower anxiety and depressive symptoms compared to those who completed the follow-up. It is unknown how these lower scores may have impacted SWE estimates.

Future Research

The study should be validated on a different sample of TKA recipients. Additionally, the order of questioning on the standardized script could then be modified, on an individual basis, based on the perceive importance of the KOOS item. Lastly, the wording used to estimate SWE values should be revised in an effort to improve participant understanding of the SWE concept and decrease the risk of negative SWE estimates.

Conclusion

Most participants (90%) in the current study required up to a 74% reduction in pain and 88% improvement in function to feel that the benefit of a TKR outweighed the costs and risks when compared to not having the surgery. These values are substantially higher than previous estimates of meaningful change, derived using the distribution- and anchor-based methods. The benefit-harm trade-off method may be a more accurate way to assess meaningful change, compared to the distribution- and anchor-based methods. The higher SWE estimates in the current study may ultimately mean that individuals require more benefit after surgery, as previously thought, when costs and risks are factored into determining meaningful change after a TKR. The SWE estimates for the KOOS Pain and Function, daily living subscales in this study should help patients to better understand the degree of improvement would be considered meaningful after a TKR.

The strongest predictors of SWE estimates for the KOOS Pain and Function, daily living subscales, after a TKR, were baseline pain and function. Those individuals with higher baseline pain and worse preoperative function required a greater improvement after a TKR to feel that the surgery was worth the associated costs and risks. Patients should understand that they may need more improvement after surgery, if they are in a lot of pain before surgery, to feel that the TKR was worthwhile.

At 6-months after surgery, 97% of participants in the present study who were satisfied with their current state had met or exceeded their SWE estimate for the KOOS Pain subscale. Similarly, 99% of participants who met their SWE estimate for the KOOS Function, daily living subscale were also satisfied with their current state. Additionally, there was a significant association ($p<0.001$) between being satisfied at 6-months with their current state and meeting/exceeding their SWE estimates for the KOOS Pain and Function, daily living subscales.

This finding ultimately indicates that those individuals who have at least met their SWE estimate after surgery are more likely to be satisfied. The association between satisfaction and meeting/exceeding the SWE value further strengthens the use of the SWE estimate as an accurate measure with which to estimate meaningful change.

REFERENCES

Bachmeier, C.J.M., March, L.M., Lapsley, H.M., Tribe, K.L., Courtenay, B.G., & Brooks, P.M. (2001). A comparsion of outcomes in patients undergoing total hip and knee replacement surgery. *Osteoarthritis and Cartilage*, 9, 137-146.

Barrack, R.L., Rhu, E.L., Chen, J., Lombardi, A.V., Berend, K.R., Parvizi, J.…Nunley, R.M. (2014). Impact of socioeconomic factors on outcome of total knee arthroplasty. *Clinical Orthopaedics and Related Research*, 472, 86-97.

Barrett B., Brown D, Mundt M, & Brown R. (2005). Sufficiently Important Difference: Expanding the Framework of Clinical Significance. *Medical Decision Making*, 25, 250-261.

Barrett B, Brown R, Mundt, M, Dye L., Alt J., Safdar N., MaberryR.(2005). Using Benefit Harm Tradeoffs to Estimate Sufficiently Important Difference: The Case of the Common Cold. *Medical Decision Making*, 25, 47-55.

Berliner, J.L., Brodke, D.J., Chan, V., SooHoo, N.F., & Bozic, K.J. (2016). Preoperative patient-reported outcome measures predict clinically meaningful improvement in function after THA. *Clinical Orthopaedics and Releated Research*, 474, 321-329.

Berliner, J.L., Brodke, D.J., Chan, V., SooHoo, N.F., & Bozic, K.J. (2017). Can preoperative patient-reported outcome measures be used to predict meaningful improvement in function after TKA. *Clinical Orthopaedics and Related Research*, 475, 149-157.

Benyon, K., Hill, S., Zadurian, N. & Mallen, C. (2010). Coping strategies and self-efficacy as predictors of outcome in osteoarthritis: a systematic review. *Musculoskelet Care*, 8, 224-236.

Besnick, A.D., Wylde, V., Gooberman-Hill, R., Blom, A., & Dieppe, P. (2012). What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ Open*, 22(1).

Biome, C. & Augustin, M. (2015). Measuring change in quality of life: bias in prospective and retrospective evaluation. *Value Health*, 18(1), 110-115.

Bourne, R.B., Chesworth, B.M., Davis, A.M., Mahomed, N.N., & Charron, K.D.J. (2010). Patient Satisfaction after Total Knee Arthroplasty. *Clinical Orthopaedics and Related Research*, 468, 57-63.

Bozic, K.J., Stacey, B., Berger, A., Sadosky, A. & Oster, G. (2012). Resource utilization and costs before and after total joint arthroplasty. *BMC Health Services Research*, 12(73), 1472-1479.

Bozic, K.J., Grosso, L.M., Lin, Z., Parzynski, C.S., Suter, L.G., Krumholz, H.M., Lieberman, J.R., Berry, D.J., Bucholz, R., Han, L., Rapp, M.T., Bernheim, S. & Drye, E.E. (2014). Variation in hospital-level risk-standardized complication rates following elective primary total hip and knee arthroplasty. *Journal of Bone and Joint Surgery*, 96, 640-647.

Brady, T. (2011). Measures of self-efficacy. *Arthritis Care & Research*, 63(S11), S473-S485.

Brander V.A, Stulberg S.D, Adams A.D, Harden R.N, Bruehl S., Stanos S.P., & Houle T. (2003). Predicting Total Knee Replacement Pain. *Clinical Orthopaedics and Related Research*, 416, 27-36.

Burns, L., Ritvo, S., Ferguson, M., Clarke, H., Zeltzer, Z, & Katz, J. (2015). Pain catastrophizing as a risk factor for chronic pain after total knee arthroplasty: a systematic review. *Journal of Pain*

Research, 8, 21-32. Carr, A., Hewlett, S., Huges, R., Mitchell, H., Ryan, S., Carr, M., & Kirwan, J. (2003). Rheumatology outcomes: the patient's perspective. *J Rheumatol*, 30, 880-883.

Callahan, C.M., Drake, B.G., Heck, D.A., & Dittus, R.S. (1994). Patient outcomes following tricompartmental total knee replacement: A meta-analysis. *Journal of American Medical Association*, 271 (17), 1349-57.

Celik, D., Coban, O. & Kilicoglu, O. (2019). Minimal clinically important difference of commonly used hip-, knee-, foot-, and ankle-specific questionnaires: a systematic review. *Journal of Clinical Epidemiology*, 113, 44-57.

Cepeda, N.A., Polascik, B.S., & Ling, D.L (2019). A primer on clinically important outcome values. J Bone Joint Surg Am., 102, 262-8.

Chesworth B.M., Mahomed N.N., Bourne R.B., & Davis A.M. (2008). Willingness to go through surgery again validated the WOMAC clinically important difference from THR/TKR surgery. *Journal of Clinical Epidemiology*, 61, 907-918.

Clement, N., Bardgett, M., Weir, D., Holland, J., Gerrand, C., & Deehan, D. (2018). What is the minimum clinically important difference for the womac index after TKR? Clin Orthop Relat Res, 476, 2005-2014.

Collins, N., Misra, D., Felson, D., Crossley, K., & Roos, E. (2011). Measures of knee function: International knee documentation committee (IKDC), knee injury and osteoarthritis outcome score physical function short form (KOOS-PS), knee outcome survey activities of daily living scale (KOS-ADL), Lysholm knee scoring scale, oxford knee score (OKS), western Ontario and mcmaster universities osteoarthritis index (WOMAC), activity rating scale (ARS), tegner activity score (TAS). (2011). *Arthritis Care & Research*, 63, S208-228.

Collins, N., Prinsen, C., Christensen, R., Bartels, E., Terwee, C., & Roos, E. (2016). Knee injury and osteoarthritis outcome score (KOOS): systematic review and meta-analysis of measurement properties. *Osteoarthritis and Cartilage*, 24, 1317-1329.

Compare knee or hip replacement or reattachment surgery near GA (2011). Retrieved from:http://www.nerdwallet.com/health/hospitals/compare?proc=Knee+or+hip+replacement+or+reattachment+surgery+without+major+complications&loc=GA&yt0=Go&page=2.

Cram, P., Vaughan-Sarrazin, M.S., Wolf, B., Katz, J.N., & Rosenthal, G.E. (2007). A comparison of total hip and knee replacement in specialty and general hospitals. *J Bone Joint Surg*, 89(8), 1675-84.

Cram P, Lu X, Kates S.L., Singh J.A., Li Y, Wolf B.R. (2012). Total Knee Arhtroplasty Volume, Utilization, and Outcomes among Medicare Beneficiaries, 1991-2010. *Journal of American Medical Association*, 308(12), 1227-1236.

Davis, A.M., Perruccio, A.V., Ibrahim, S., Hogg-Johnson, S.,Wong, R., Streiner, D.L., Beaton, D.E., Cote, P., Gignac, M.A., Flannery, J., Schemitsch, E., Mahomed, N.N., & Badley, E.M. (2011). The trajectory of recovery and inter-relationships of symptoms, activity participation in the first year following total hip and knee replacement. *Osteoarthritis and Cartilage*, 19, 1413-1421.

Deyo, R.A. & Patrick, D.L. (1995). The Significance of Treatment Effects: The Clinical Perspective. *Medical Care*, 33(4), AS286-AS291.

Duric, V.M., Fallowfield, L.J., Saunders, C., Houghton, J., Coates, A.S., & Stockler, M.R. (2005). Patients' preferences for adjuvant endocrine therapy in early breast cancer: what makes it worthwhile? *British Journal of Cancer*, 93, 1319-1323.

Escobar, A., Garcia Perez, L., Herrera-Espineira, C., Aizpuru, F., Sarasqueta, C., Gonzalez Saenz de Tejada, M., Quintana, J.M., & Bilbao, A. (2017). Total knee Replacement: Are there any baseline factors that have influence in patient reported outcomes? *J Eval Clin Pract*, 23, 1232-1239.

Escobar, A., Quintana, J.M., Bilbao, A., Arostegui, I., Lafuente, I., & Vidaurreta, I. (2007). Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. *Osteoarthritis Cartilage*, 15, 273-280.

Ferreira, M.L., Ferreira, P.H., Herbert, R. D., & Latimer, J. (2009). People with low back pain typically need to feel "much better" to consider intervention worthwhile: an observational study. *Australian Journal of Physiotherapy*, 55, 123-127.

Ferreira M.L., Herbert R.D., Ferreira P.H., Latimer J., Ostelo R.W., Nascimento D.P., Smeets R.J. (2012). A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. *Journal of Clinical Epidemiology*, 65, 253-261.

Ferreira M.L., Herbert R.D., Ferreira P.H., Latimer J., Ostelo R.W., Grotle M., Barrett B. (2013). The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. *Journal of Clinical Epidemiology*, 2013, 66, 1397-1404.

Field, A. (2009) Discovering Statistics using SPSS. London, England: SAGE Publications Ltd.

Forsythe, M., Dunbar, M., Hennigar, A. & Sullivan, M. (2008). Prospective relation between catastrophizing and residual pain following knee arthroplasty: Two-year follow-up. Pain Res Manage, 13(4), 335-341.

Fortin, P.R., Clarke, A.E., Joseph, L., Liang, M.H., Tanzer, M., Ferland, D., Phillips, C., Partridge, A.J., Belisle, P., Fosel, A.H., Mahomed, N., Sledge, C.B, & Katz, J.N. (1999). Outcomes of total hip and knee replacement. *Arthritis and Rheumatism*, 42(8), 1722-1728.

Franco, M.R., Howard, K., Sherrington, C., Rose, J., Ferreira. P.H., & Ferreira, M.L. (2016). Smallest worthwhile effect of exercise programs to prevent falls among older people: estimates from benefit-harm trade-off and discrete choice methods. *Age and Ageing*, 45, 806-812.

Garrison C., Cook C. (2012). Clinimetrics corner: The Global Rating of Change Score (GRoC) poorly correlates with functional measures and is not temporally stable. *Journal of Manual and Manipulative Therapy*, 20(4), 178-181.

Gatchel, R.J., Lurie, J.D., & Mayer, T.G. (2010). Minimal Clinically Important Difference. *Spine*, 35(19), 1739-1743.

Greengard, S. & Kruick, G. (2012). Understanding Knee Replacement Costs: What's on the Bill? Retrieved February 12, 2015.

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *J Chron Dis*, 40(2), 171-178. *Arch Intern Med*, 168(13), 1430-1440.

Hamel, M.B., Toth, M., Legedza, A., & Rosen, M.P. (2008). Joint replacement surgery in elderly patients with severe osteoarthritis of the hip and knee. *Archives of Internal Medicine*, 168(13), 1430-40.

Harris, W.H. & Sledge, C.B. (1990). Total hip and knee replacement (1). *New England Journal of Medicine,* 323, 725-31.

Harris, W.H. & Sledge, C.B. (1990). Total hip and knee replacement (2). *New England Journal of Medicine*, 323, 801-7.

Harris, K.K., Dawson, J., Jones, L.D., Beard, D.J., & Price, A.J. (2013). Extending the use of PROMs in the NHS- using the oxford knee score in patients undergoing non-operative management for knee osteoarthritis: a validation study. *British Journal of Medicine*, 3, e003365.

Healy W.L., Rana A.J., Iorio R. (2010). Hospital economics of primary total knee arthroplasty at a teaching hospital. *Clinical Orthopaedic and Related Research*, 469, 87-94.

Herbert, R.D., Jamtvedt, G., Mead, J., & Hagen, K.B.. (2005). Outcome measures measure outcomes, not effects of intervention. *Aust J Physiother*, 51, 3-4.

Hirakawa, Y., Hara, M., Fujiwara, A., Hanada, H., & Morioka, S. (2014). The relationship among psychological factors,neglect-like symptoms and postoperative pain after total knee arthroplasty. *Pain Res Manag*, 19(5), 251-256.

Huang, C.C., Chen, W.S., Tsai, M.W., & Wang, W.T.J. (2017). Comparing the Chinese versions of the two knee-specific questionnnaires (IKDC and KOOS: reliability, validity, and responsiveness. *Health Quality of Life Outcomes*, 15, 238.

Jacobs, C.A & Christensen, C.P. (2009). Correlations between Knee Society Function Scores and Functional Force Measures. *Clin Orthop Relat Res*, 467, 2414-219.

Jacobson, N.S., Roberts, L.J, Berns, S.B., & McGlinchey, J.B. (1999). Methods for Defining and Determining the Critical Significance of Treatment Effects: Description, Application, and Alternatives. *Journal of Counseling and Clinical Psychology*, 67(3), 300-307.

Jaeschke R., Singer J. Guyatt G.H. (1989). Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, 10, 407-415.

Jones, C.A., Voaklander, D.C. Johnston, D.W., & Suarez-Almazor, M.E. (2000). Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. *Journal of Rheumatology*, 27, 1745-1752.

Jones, C.A., Voaklandr, D.C. & Suarez-Almazor, M.E. (2003). Determinants of function after total knee arthroplasty. *Physical Therapy*, 83(8), 696-706.

Jones, C.A., Beaupre, L.A., Johnston, D.W.C, & Suarez-Almazor, M.E. (2007). Total Joint Arthroplasties: Current Concepts of Patient Outcomes after Surgery. *Rheum Dis Clin N Am*,33,71-86.

Judge, A, Arden, N.K., Cooper, C. Javai,d M.K., Carr, A.J., Field, R.E.,& Diepp, P.A. (2012). Predictors of outcomes of total knee replacement surgery. *Rheumatology*, 5(1), 1804-1813.

Juni,, P., Reichenbach,, S., & Dieppe, P.A. (2006). Osteoarthritis: rational approach to treating the individual. *Best Practice & Research Clinical Rheumatology*, 20(4), 721-740.

Kamper, S.J., Maher, C.G., & Mackay, G. (2009). Global Rating of Change Scales: A review of the strengths and weaknesses and considerations for design. *Journal of Manual and Manipulative Therapy*, 17(3), 163-170.

Kazis, L.E., Anderson, J.J., & Meenan, R.F. (1989). Effect Sizes for Interpreting Changes in Health Status. *Medical Care*, 27(3),S178-S189.

Kennedy, D., Stratford, P., Riddle, D., Hanna, S., & Gollish, J. (2008). Assessing recovery and establishing prognosis following total knee arthroplasty. *Physical Therapy*, 88, 22-32.

Kennedy, D., Hanna, S., Stratford, P., Wessel, J. & Gollish, J. (2006). Preoperative function and gender predict pattern of functional recovery after hip and knee arthroplasty. *Journal of Arthroplasty*, 21(4), 559-566.

Kolber, M.J. & Hanney, W.J. (2010). The reliability, minimum detectable change and construct validity of a clinical measurement for identifying posterior shoulder tightness. *N Am J Sports Phys Ther*, 5(4), 208-219.

Kroenke, K., Strine, T., Spitzer, R., Williams, J., Berry, J., & Mokdad, A. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114, 163-173.

Kroenke, K. (2012). Enhancing the clinical utility of depression screening. *CMAJ*, 184(3), 281-282.

Losina, E., Walensky, R.P., Kessler, C.L.,Emrani, P.S., Reichmann, W.M., Wright, E.A., Holt, H.L., Solomon, D.H., Yelin, E., Paltiel, A.D. & Katz, J.N. (2009). Cost-effectiveness of Total Knee Arthroplasty in the United States. *Arch Intern Med*., 169(12), 1113–1122.

Lingard, E., Katz, J., Wright, E., & Slede, C. (2004). Predicting the outcome of total knee arthroplasty. *Journal of Bone and Joint Surgery*, 86-A (10), 2179-2186.

Lews, G.N., Rice, D.A., McNair, P.J., & Kluger, M. (2015). Predictors of persistent pain after total knee arthroplasty: a systematic review and meta-analysis. *British Journal of Anaesthesia*, 114(4), 551–61.

Lopez-Olivo, M.A., Landon, G.C., Siff, S.J., Edelstein, D., Pak, C., Kallen, M.A….Suarez-Almazor, M.E. (2011). *Ann Rheum Dis*, 70,1775-1781.

Lyman, S., Lee, Y., McLawhorn, A.S., Islam, W., & MacLean, C.H. (2018). What are the minimal and substantial improvements in the HOOS and KOOS JR versions after total joint replacement. *Clinical Orthopaedics and Related Research*, 476, 2432-2441.

Magklara, E., Burton, C.R. & Morrison, V. (2014). Does self-efficacy influence recovery and well-being in osteoarthritis patients undergoing joint replacement? A systematic review. *Clinical Rehabilitation*, 28(9), 835-846.

Mahomed, N., Liang, M., Cook, E., Daltroy, L., Fortin, P., Fossel, A., & Katz, J. (2002). The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *Journal of Rheumatology*, 29(6), 1273-1279.

MacKay, C., Clements, N., Wong, R. & Davis, A. (2019). A systematic review of estimates of the minimal clinically important difference and patient acceptable symptom state of the western Ontario and mcmaster universities osteoarthritis index in patients who underwent total hip and total knee replacement. *Osteoarthritis and cartilage*, 27, 1408-1419.

Maratt, J.D., Lee, Y., Lyman, S., & Westrich, G.H. (2015). Predictors of satisfaction following total knee arthroplasty. *The Journal of Arthroplasty*, 30, 1142-1145.

McNamara, R.J., Elkins, M.R., Ferreira, M.L., Spencer, L.M. & Herbert, R.D. (2015). Smallest worthwhile effect of land-based and water-based pulmonary rehabilitation for COPD. *ERJ Open Res*, 1, 1-8.

Merle-Vincent, F., Couris, C., Schott, A.M., Conrozier, T., Piperno, M., Matthieu, P.& Vignon, E. (2011). Factors predicting patient satisfaction 2 years after total knee arthroplasty for osteoarthritis. *Joint Bone Spine*, 78, 383-386.

Mills, K.A., Naylor, J.M., Eyles, J.P., Roos, E.M., & Hunter, D.J. (2016). Examining the minimal important difference of patient-reported outcome measures for individuals with knee osteoarthritis: a model using the knee injury and osteoarthritis outcome score. *Journal of Rheumatology*, 43, 395-404.

Mizner, R.L., Petterson, S.C., Clements, K.E., Zenl, J.A., rrgang, J. & Snyder-Mackeler, L. (2011)/ Measuring Functional Improvements after Total Knee Arthroplasty Requires both Performance-Based and Patient-Report Assessments: A Longitudinal Analysis of Outcomes. *J Arthroplasty*, 26(5). 728-737.

Monticone, M., Ferrante, S., Salvaderi, S., Motta, L., & Cerri, C. (2013). Responsiveness and minimal important changes for the knee injury and osteoarthritis outcome score in subjects undergoing rehabilitation after total knee arthroplasty. *American Journal of Physical Medicine and Rehabilitation*, 92, 864-70.

Nilsdotter, A.K., Toksvig-Larsen, S. & Roos, E.M. (2009). Knee arthroplasty: are patient's expectations fulfilled? *Acta Orthopaedica*, 80(1), 55-61.

Nilsdotter, A.K., Toksvig-Larsen, S. & Roos, E.M. (2009). A 5 year prospective study of patient-relevant outcomes after total knee replacement. *Osteoarthritis and Cartilage*, 17, 601-606.

Noiseux, N., Callaghan, J., Clark, C., Zimmerman, B., Sluka, K., & Rakel, B. (2014). Preoperative Predictors of Pain Following Total Knee Arthroplasty. *J Arthroplasty*, 29(7), 1383–1387.

Norman, G.R, Stratford P., & Regehr, G. (1997). Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol*, 50, 869-879.

Norman, G.R., Sridhar, F.G, Guyatt, G.H., &Walter, S.D. (2001). Relation of Distribution- and Anchor-Based Approaches in Interpretation of Changes in Health-Related Quality of Life. *Medical Care*,39(1),1039-1047.

Ottenbacher, K.J., Johnson, M.B., & Hojem, M. (1988). The Significance of Clinical Change and Clinical Change of Significance: Issues and Methods. *American Journal of Occupational Therapy*, 42(3), 156-163.

Patridge, C.J. (1982). The outcome of physiotherapy and its measurement. *Physiotherapy*, 68(11), 362-363.

Peer, M.A. & Lane, J. (2013). The Knee Injury and Osteoarthritis Outcome Score (KOOS): A Review of its Psychometric Properties in People Undergoing Total Knee Arthroplasty. *Journal of Orthopaedic & Sports Physical Therapy*, 43(1), 20-28.

Petherick, E.S., O'Meara, S.O., Spilsbury, K., Iglesias, C.P., Nelson, E.A., & Togerson, D.J. (2006). Patient acceptability of larval therapy for leg ulcer treatment: a randomized survey to inform the sample size calculation of a randomized trial. *BMC Medical Research Methodology*, 6(43), 1-4.

Rasouli, M.R., Restrepo, C., Maltenfort, M.G, Purtill, J.J., & Parvizi, J. (2014). Risk factors for surgical site infection following total joint arthroplasty. *Journal of Bone and Joint Surgery*, 96, e158(1)-e158(5).

Riddle, D., Wade, J., Jiraneck, W., & Kong, X. (2010). Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. *Clin Orthop Relat Res*, 468, 798-806.

Roos, E., Roos, H., Lohmander, L., Ekdahl, C., Beynnon, B. (1998). Knee Injury and Osteoarthritis Outcome Score- Development of a Self-Administered Outcome Measure. *J Orthop Sports Phys Ther*, 28(2):88-96.

Roos, E. & Lohmander, L. (2003). Knee injury and osteoarthritis outcome score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes*, 1, 64.

Schmitt, J.S. & Di Fabio, P. (2004). Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *Journal of Clinical Epidemiology*, 57, 1008-1018.

Schmitt, J.S. & Di Fabio, P. (2005). The validity of prospective and retrospective global change criterion measures. *Arch Phys Med Rehabil*, 86(12), 2270-2276.

SooHoo, N.F, Li, Z., Chenok, K.E., & Bozic, K.J. (2015). Responsiveness of patient reported outcome measures in total joint arthroplasty patients. *Journal of Arthroplasty*, 30, 176-191.

Spitzer, R., Kroenke, K., Williams, J. & Lowe, B. (2006). A Brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine*, 166, 1092-1097.

Stratford, P.W., Binkley, J.M., & Riddle, D.L. (1996). Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther*, 76(10), 1109-1123.

Stratford, P.W. & Kennedy, D.M. (2006). Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *Journal of Clinical Epidemiology*, 59, 160-167.

Sullivan, G. & Fein, R. (2012). Using effect size- or why the p-value is not enough. *Journal of Graduate Medical Education*, 4(3), 279-282.

Tejada, M.G., Escobar, A., Bilbao, A., Herrera- Espineira, C., Garcia-Perez, L., Aizpuru,F., & Sarasqueta, C. (2014). A prospective study of the association of patient expectations with changes in health-related quality of life outcomes, following total joint replacement. BMC Musculoskeletal Disorders, 15, 248.

Tubach, F., Ravaud, P., Baron, G., Falissard, B., Logeart, L., Bellamy, N., Bombardier, C., Felson, D., Hochberg, M., van der Heije, D., & Dougados, M. (2005). Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Am Rheum Dis*, 64, 29-33.

Tuley, M.R., Mulrow, C.D., & McMahan, C.A. (1991). Estimating and testing an index of responsiveness and the relationship of the index to power. *J Clin Epidemiol*, 44(4-5), 417-421.

Uyen-Sa, D.T.N., Ayers, D.C., Li, W., Harrold, L.R., & Franklin, P.D. (2016). Pre-operative pain and function: Profiles of patients selected for total knee replacement. *Journal of Arthroplasty*, 31(11), 2402-2407.

Van Onsem, S., Van Der Straeten, C., Arnout, N., Deprez, P., Van Damme, G., & Victor, J. (2016). A new prediction model for patient satisfaction after total knee arthroplasty. *Journal of Arthroplasty*, 31, 2660-2667.

Viechtbauer, W., Smits, L., Kotz, D., Bude, L., Spigt, M., Serroyen, J., & Crutzen, R. (2015). A simple formula for the calculation of sample size in pilot studies. *Journal of Clinical Epidemiology*. Accepted.

Wang S.W., Olson-Kellogg B, Shamliyan T.A., Choi J.Y., Ramakrishnan R., Kane R.L. (2012). Physical Therapy Interventions for Knee Pain Secondary to Osteoarthritis. *Annals of Internal Medicine*, 157(9), 632-644.

Watanabe, H., Sekiya, H., Kariya, Y., Hoshino, Y., Sugimoto, H., & Hayasaka, S. (2011). The incidence of venous thromboembolism before and after total knee arthroplasty using 16-row multidetector computed tomography. *Journal of Arthroplasty*, 26(8), 1488-1493.

Wells, G., Beaton, D.E., Shea, B., Simon, L., Strand, V., Brooks, P., & Tugwell, P. (2001) Minimal clinically important differences: reviews of methods. *Journal of Rheumatology*, 28, 406-412.

Wiebe, S., Matijevic, S. Eliasziw, M., & Derry, P.A. (2002). Clinically important change in quality of life in epilepsy. *J Neurol Neurosurg Psychiatry*, 73(2), 116-120.

Wilcox, S., Schoffman, D., Dowda, M., Sharpe, P. (2014). Psychometric properties of the 8-item English arthritis self-efficacy scale in a diverse sample. Arthritis.

Williams, D.P., Price, A.J., Beard, D.J., Hadfield, S.G., Arden, N.K., Murray, D.W. & Field, R.E. (2013). The effects of age on patient-reported outcome measures in total knee replacements. *The Bone and Joint Journal*, 95, B38-44.

Woolhead, G.M., Donovan, J.L., & Dieppe, P.A. (2005). Outcomes of total knee replacement: a qualitative study. *Rheumatology*, 44, 1032-1037.

Wylde, V., Dixon, S. & Blom, A.W. The role of preoperative self-efficacy in predicting outcome after total knee replacement. *Musculoskelet Care*, 10, 110-118.

Wyrwich, K.W., Nienabaer, N.A., Tierney, W.M., & Wolinsky, F.D. (1999). Linking Clinical Relevance and Statistical Significance in Evaluating Intra-Individual Changes in Health-Related Quality of Life. *Medical Care*, 37(5), 469-478.

APPENDIX A

Script for determining the smallest worthwhile effect of a Total Knee Arthroplasty

We are trying to determine how much your knee pain and function will need to improve for you to believe that your surgery is worthwhile to you.

Total Knee Replacement

Before we begin, I would like to review with you the costs and risks of a total knee replacement. You can use this information in whatever way you choose when I ask you the remaining questions. The surgery is an inpatient procedure performed by an orthopedic surgeon. After the surgery, you will remain in the hospital for approximately 1 to 4 days. An average of six weeks of formal physical therapy is usually necessary for recovery. The 90-day perioperative costs in Savannah range from $14-17,000. Of these costs, Medicare typically covers 100% of the inpatient patient costs and 80% of the outpatient patient costs, while private insurance costs are typically higher and they vary in the amount they cover. There is a 3-5% overall complication rate. According to research, of the 90 day post-surgical potential complications, there is a 1% chance of infection, less than 1% chance for pneumonia, less than 2% chance of a blood clot, and a 1% chance of requiring a revision of your total knee. Most of these complications will result in longer hospitalization for additional treatment.

Do you have any questions about the knee replacement surgery?

SWE in pain for total knee replacement

Think about the knee pain you have, based on the pain scale you completed. Without the knee replacement surgery, you could expect your pain to continue to worsen by approximately 10-15% over the next six months. Most of the improvement in pain occurs by 6 months after surgery. You scored a ___on your initial pain scale. Please complete this additional pain scale to show how much pain with activity is acceptable 6 months after surgery, considering the 10% to 15% worsening you would experience without the knee replacement surgery, in order for you to feel that the surgery was worthwhile, given the costs and risks we reviewed earlier.

Now, I am going to provide you with a scenario where I change your response to questions on the 6 month pain scale. For example, if you reported “no pain” with question # 7- “pain at night while in bed” and I shifted your response from “no pain” to “mild pain” with this activity, would you still consider the surgery worthwhile considering the associated costs and risks?

Now, let's begin. For questions #AA and #BB on the 6 month pain scale, you reported "_____" pain with these activities. If I shifted your response from "_____" pain to "_____" pain with these activities, indicating a higher level of pain, would you still consider the surgery worthwhile considering the associated costs and risks?

#AA= the item on the 6-month KOOS-Pain scale with the lowest, least painful, item score.

#BB= the item on the 6-month KOOS-Pain scale with the second lowest, least painful, item score.

If the answer is no:

Would you still consider the surgery worthwhile if I only shifted your response on question #AA from "_____" pain to "_____" pain with this activity, indicating a higher level of pain?

If the answer is yes:

Now, let us assume that you now have "_____" pain with questions #AA and BB, based on the previous scenario. If I shifted your response on two additional questions, #CC and #DD, from "_____" pain to "_____" pain with these activities, indicating a higher level of pain with both activities, would you still consider the surgery worthwhile considering the associated costs and risks?

#CC= the item on the 6-month KOOS-Pain scale with the third lowest, least painful, item score.CC

#DD- the item on the 6-month KOOS-Pain scale with the fourth lowest, least painful, item score.

If the answer is no:

Would you still consider the surgery worthwhile if I only shifted your response on question #CC from "_____" pain to "_____" pain with this activity, indicating a higher level of pain?

***This same procedure will continue until either the subject answers "no" to both activities or they answer "yes" to all nine items on the pain scale.**

SWE in function for total knee replacement

Think about the activities which you had difficulty with on the functional scale you completed. Without the knee replacement surgery, you could expect for your functional level to continue to

worsen by approximately 10-15% over the next six months. The majority of improvement in function can be expected by 6 months after surgery. You scored a ___ on the initial functional scale. Please complete this additional functional scale to show how much difficulty with activity, considering the 10% to 15% worsening you would experience without the knee replacement surgery, is acceptable in order for you to feel that the surgery was worthwhile, given the costs and risks we reviewed earlier?

I am going to provide you with a scenario where I change your response to two of the questions on the 6 month function scale. For example, if you reported "no difficulty" for question # 14- "difficulty with sitting or lying" and I shifted your response from "no difficulty" to "mild difficulty" with this activity, would you still consider the surgery worthwhile?

Now, let's begin. For questions #AA and #BB on the 6 month functional scale, you reported "_____" pain with these activities. If I shifted your response from "_____" difficulty to "_____" difficulty with these activities, indicating a higher level of difficulty, would you still consider the surgery worthwhile considering the associated costs and risks?

#AA= the item on the 6-month KOOS-Function scale with the lowest, least difficult, item score.

#BB= the item on the 6-month KOOS-Pain scale with the second lowest, least difficult, item score.

If the answer is no:

Would you still consider the surgery worthwhile if I only shifted your response on question #AA from "_____" difficulty to "_____" difficulty with this activity, indicating a higher level of difficulty?

If the answer is yes:

Now, let us assume that you now have "_____" difficulty with questions #AA and BB, based on the previous scenario. If I shifted your response on two additional questions, #CC and #DD, from "_____" difficulty to "_____" difficulty with these activities, indicating a higher level of difficulty with both activities, would you still consider the surgery worthwhile considering the associated costs and risks?

#CC= the item on the 6-month KOOS-Pain scale with the third lowest, least difficulty, item score.CC

#DD- the item on the 6-month KOOS-Pain scale with the fourth lowest, least difficulty, item score.

If the answer is no:

Would you still consider the surgery worthwhile if I only shifted your response on question #CC from "_____" difficulty to "_____" difficulty with this activity, indicating a higher level of difficulty?

***This same procedure will continue until either the subject answers "no" to both activities or they answer "yes" to all seventeen items on the functional scale.**

APPENDIX B

Date _____ Subject #_____

KOOS KNEE SURVEY

INSTRUCTIONS: This survey asks for your view about your knee. This information will help us keep track of how you feel about your knee and how well you are able to perform your usual activities.
Answer every question by ticking the appropriate box, only one box for each question. If you are unsure about how to answer a question, please give the best answer you can.

Pain

P1. How often do you experience knee pain?

Never	Monthly	Weekly	Daily	Always
☐	☐	☐	☐	☐

What amount of knee pain have you experienced the **last week** during the following activities?

P2. Twisting/pivoting on your knee

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P3. Straightening knee fully

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P4. Bending knee fully

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P5. Walking on flat surface

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P6. Going up or down stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P7. At night while in bed

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P8. Sitting or lying

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P9. Standing upright

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

APPENDIX C

Function, daily living

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 week** due to your knee.

A1. Descending stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A2. Ascending stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A3. Rising from sitting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A4. Standing

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

For each of the following activities please indicate the degree of difficulty you have experienced in the **last week** due to your knee.

A5. Bending to floor/pick up an object

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A6. Walking on flat surface

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A7. Getting in/out of car

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

A8. Going shopping

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

A9. Putting on socks/stockings

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

A10. Rising from bed

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

A11. Taking off socks/stockings

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

A12. Lying in bed (turning over, maintaining knee position)

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

A13. Getting in/out of bath

None ☐ Mild ☐ Moderate ☐ Severe ☐ Extreme ☐

For each of the following activities please indicate the degree of difficulty you have experienced in the **last week** due to your knee.

A14. Sitting

None Mild ☐ Moderate Severe Extreme

☐ ☐ ☐ ☐

A15. Getting on/off toilet

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A16. Heavy domestic duties (moving heavy boxes, scrubbing floors, etc)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A17. Light domestic duties (cooking, dusting, etc)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Collins, N. & Roos, E. (2012). Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a "good" measure. *Clinics in geriatric medicine*, 28(3), 367-394.

Thank you very much for completing all the questions in this questionnaire.

APPENDIX D

Date:_____ Subject #_____

Generalized Anxiety Disorder 7-item (GAD-7) scale

Over the last 2 weeks, how often have you been bothered by the following problems?	Not at all sure	Several days	Over half the days	Nearly every day
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it's hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3
Add the score for each column	+	+	+	
Total Score *(add your column scores)* =				

If you checked off any problems, how difficult have these made it for you to do your work, take care of things at home, or get along with other people? Check the appropriate response.

Not difficult at all_____
Somewhat difficult _____
Very difficult_____
Extremely difficult _____

Source: Spitzer RL, Kroenke K, Williams JBW, Lowe B. A brief measure for assessing generalized anxiety disorder. *Arch Inern Med*. 2006;166:1092-1097.

Thank you very much for completing all the questions in this questionnaire.

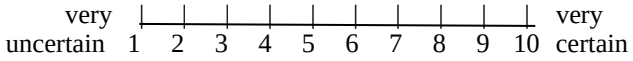
APPENDIX E

Date _____ Subject #_____

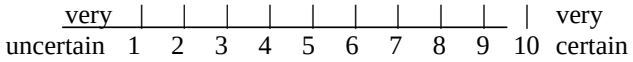
Arthritis Self-Efficacy 8-item Scale

For each of the following questions, please circle the number that corresponds to how certain you are that you can do the following tasks regularly at the present time.

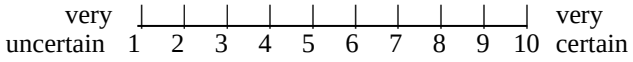
1. How certain are you that you can decrease your pain quite a bit?



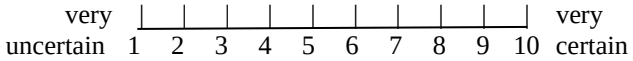
2. How certain are you that you can keep your arthritis pain from interfering with your sleep?



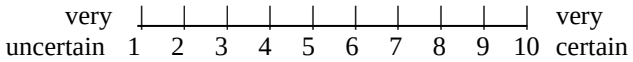
3. How certain are you that you can keep your arthritis pain from interfering with the things you want to do?



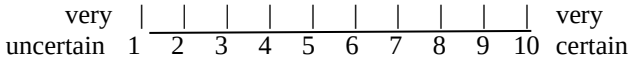
4. How certain are you that you can regulate your activity so as to be active without aggravating your arthritis?



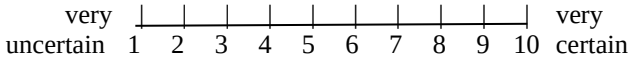
5. How certain are you that you can keep the fatigue caused by your arthritis from interfering with the things you want to do?



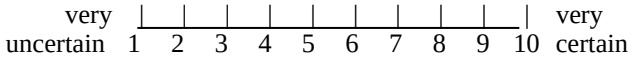
6. How certain are you that you can do something to help yourself feel better if you are feeling blue?



7. As compared with other people with arthritis like yours, how certain are you that you can manage pain during your daily activities?



8. How certain are you that you can deal with the frustration of arthritis?



Lorig K, Chastain RL, Ung E, Shoor S, & Holman HR: Development and evaluation of a scale to measure self-efficacy in people with arthritis. Arthritis and Rheumatism, 32, 1, 1989, pp. 37-44 (original scales).

Thank you very much for completing all the questions in this questionnaire.

APPENDIX F

Date _____ Subject # _____

Personal Health Questionnaire Depression Scale (PHQ-8)

Over the **last 2 weeks**, how often have you been bothered by any of the following problems?
*(circle **one** number on each line)*

How often during the past 2 weeks were you bothered by...	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself, or that you are a failure, or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have				

been moving around a lot more than usual 0 1 2 3

Kroenke K, Strine TW, Spritzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord. 2009; 114(1-3):163-73.

Thank you very much for completing all the questions in this questionnaire.

APPENDIX G

Date _____
Subject # _____

Pain Catastrophizing Scale

Everyone experiences painful situations at some point in their lives. Such experiences may include headaches, tooth pain, joint or muscle pain. People are often exposed to situations that may cause pain such as illness, injury, dental procedures or surgery.

Instructions:

We are interested in the types of thoughts and feelings that you have when you are in pain. Listed below are thirteen statements describing different thoughts and feelings that may be associated with pain. Using the following scale, please indicate the degree to which you have these thoughts and feelings when you are experiencing pain.

RATING	**0**	**1**	**2**	**3**	**4**
MEANING	Not at all	To a slight degree	To a moderate degree	To a great degree	All the time

When I'm in pain …

Number	**Statement**	**Rating**
1	I worry all the time about whether the pain will end.	
2	I feel I can't go on.	
3	It's terrible and I think it's never going to get any better	
4	It's awful and I feel that it overwhelms me.	
5	I feel I can't stand it anymore	
6	I become afraid that the pain will get worse.	
7	I keep thinking of other painful events	
8	I anxiously want the pain to go away	
9	I can't seem to keep it out of my mind	

10	I keep thinking about how much it hurts.	
11	I keep thinking about how badly I want the pain to stop	
12	There's nothing I can do to reduce the intensity of the pain	
13	I wonder whether something serious may happen.	

Source: **Sullivan** MJL, **Bishop** S, **Pivik** J. The **pain catastrophizing scale**: development and validation. *Psychol Assess*, **1995**, 7: 524-532

Thank you very much for completing all the questions in this questionnaire.

APPENDIX H

Subject #_____

Date:_____

PATIENT EXPECTATIONS SURVEY

1) How painful do you expect your knee to be in one year? (not at all painful, slightly painful, moderately painful, very painful) **Please circle an answer above.**

2) How limited do you expect to be in your usual activities in one year? (not at all limited, slightly limited, moderately limited, very limited) **Please circle an answer above.**

3) How likely will your surgery be a complete success? (0 to 100 VAS score) **Please place a mark on the line below as your answer to the above question.**

0_____100

Very likely Not at all likely

4) How likely is it that you will have a knee joint complication? (0 to 100 VAS score) **Please place a mark on the line below as your answer to the above question.**

0_____100

Very likely Not at all likely

APPENDIX I

DEMOGRAPHIC QUESTIONNAIRE

Date _____ Subject #_____

Please place an "x" to indicate your answer to the following question:

1. Gender: Male _____ Female_____

Please answer the following question by filling in the blank with your answer.

2. Birth date: _____day _____month _____ year

Please circle the following response which best answers each question.

3. What is the highest degree or level of school you have completed?

   Less than high school graduate — High school graduate

   Some college — College graduate

   Some graduate school — Graduate degree

4. What is your approximate yearly income before taxes? (Include total income of all adults living in your household)

   Less than $10K — $10K- <$25K

   $25-$50K — $50K-$100K

   $100K or greater

APPENDIX J

ST JOSEPHS/CANDLER HEALTH SYSTEM CONSENT FOR PARTICIPATION IN A RESEARCH STUDY

("I" or "You" refers to the person participating in the study).

Why are you being invited to take part in a research study?

You are being asked to participate in a research study because you are between the ages of 45-90 and have agreed to undergo a Total Knee Arthroplasty surgery on one of your knees. Your participation is voluntary. Your decision whether or not to participate will not affect the quality of medical care you receive. If you join the study, you can stop or leave at any time with no changes in the quality of the health care you receive. You will be told about any new information or changes in the study that could affect you.

Please ask questions if there is anything you do not understand.

PURPOSE OF STUDY

The purpose of this study is to learn how much improvement in pain and function is necessary to justify the costs and risks of Total Knee Replacement. Total Knee Replacement is a commonly performed surgery. This study will help us better understand the effects of this treatment.

You may or may not benefit from being in this research study. We hope to gather information that may help people in the future.

DESCRIPTION OF STUDY PROCEDURES

Approximately 120 subjects will participate in this research study at St. Josephs/Candler Health System. Your participation in this study is expected to last 30-45 minutes and is a one-time only commitment.

During the session, you will be asked to complete a demographic questionnaire and seven different surveys used to look at pain, function and expectations before surgery, as well as, depression, anxiety, pain catastrophizing, and self-efficacy. Next, the study staff will use a script to ask you questions that relate to how much improvement in pain would you need to have in order for you to feel that the Total Knee Replacement surgery was worthwhile, given the associated costs and risks of the surgery. Then, you will be asked questions related to the amount of difficulty with different functional activities that you are willing to have in order to feel that the Total Knee Replacement surgery was worthwhile, given the costs and risks of the surgery.

Significant new findings developed during the course of the research which may relate to your willingness to continue participation will be provided to you.

POTENTIAL RISKS/DISCOMFORTS

There a few risks associated with this study. The first is loss of time as it will take approximately 30-45 minutes to complete the data collection session. Secondly, there is a possibility of psychological risks as you will be asked to complete surveys related to depression, anxiety, and pain catastrophizing. There is minimal likelihood of the above outlined risks. No physical, legal, or financial risks are associated with this study.

You have the right to refuse participation or to stop participation any time during the study. Additionally, a list of community counseling resources will be provided to you, if needed.

One of the risks of being in this study is that your personal information could be lost or exposed. This is very unlikely to happen, and we will do everything we can to make sure that your information is protected.

REPRODUCTIVE RISKS

No reproductive risks are associated with this study.

POTENTIAL BENEFITS

Participation in this study may help to improve your condition, but it is also possible that your condition may worsen. There is no guarantee that you will personally benefit by participating in this research study. Your participation in this study may provide information that may help other people who have a similar medical problem in the future.

Additionally, you will receive a $15.00 gift card as compensation for your time in participating in this study.

TREATMENT ALTERNATIVES

You do not have to participate in this research study to receive treatment for your condition. You may choose not to be in this study. This will not jeopardize your care in any way.

HIPPPA AUTHORIZATION

Persons/organizations providing the information: Student investigator

Persons/organizations receiving the information: The Institutional Review Boards of St. Josephs/Candler Health System, and Virginia Commonwealth University

Specific description of information: Potentially identifiable information about you will consist of demographic information, including age, gender, income level, and education level, learned from the demographic questionnaire. This information is being collected only for research purposes.

Your information will be identified by ID numbers, not names, and stored separately from research data in a locked research area. All personal identifying information will be kept in password protected files and these files will be deleted at the completion of the research study. The completed scripts will be kept in a locked file cabinet for three years after the study ends and will be destroyed at that time. Raw data, without any personal identifying information, will be kept indefinitely. Access to all data will be limited to study personnel. A data and safety monitoring plan is established. Confidentiality will be maintained to the extent possible by law.

This information is being disclosed for the following purposes: to conduct a study assessing how much change in pain and function is necessary to feel that the costs and risks of a Total Knee Replacement are justified.

I may revoke this authorization at any time by notifying the student principal investigator in writing to the following address: 10935 Abercorn Street, Savannah, GA, 31419. If I do revoke my authorization, any information previously disclosed cannot be withdrawn. Once information about me is disclosed in accordance with this authorization, the

recipient may redisclose it and the information may no longer be protected by federal privacy regulations.

I may refuse to sign this authorization form. If I choose not to sign this authorization form, my medical care will not be affected; however, I cannot participate in the research study.

This authorization will expire the date the research study ends.

I will be given a copy of this authorization form.

CONFIDENTIALITY

Your medical records will be kept as confidential as possible within the limitations of state and federal law. Federal Privacy Regulations require that you authorize the release of any health information that may reveal your identity. The persons and entities that you are authorizing to use of disclose your individually identifiable health information may include the study staff, St. Joseph's/Candler Health System, and Virginia Commonwealth University.

In order to analyze the data collected during this research study, all of the health information generated or collected about you during the study may be inspected by St. Joseph's/Candler Health System staff involved with the study, the Federal Food and Drug Administration (FDA), the Department of Health and Human Services (DHHS) agencies, the St. Joseph's/Candler Health System Institutional Review Board, and Virginia Commonwealth University Institutional Review Board (the committee that reviews, approves, and monitors research on human subjects). Because of the need to release information to these parties, absolute confidentiality cannot be guaranteed. Once your personal health information is released, it may be redisclosed, at which point your health information will no longer be protected by federal privacy regulations.

The results of this research may be presented at meetings or in publications; however, your identity will not be disclosed in those presentations. By signing this informed consent form, you are authorizing such access to your medical records. This authorization will have no expiration.

Will being in this study cost me any money?

All study costs will be covered by the study staff. There will be no costs to you. You will be provided with a $15.00 gift card as compensation for your participation in this study.

COMPENSATION FOR INJURY

If you suffer an injury as a result of your participation in this study, St. Joseph's/Candler Health System will provide the necessary treatment for such injury, but you, your insurance company or a government program will be billed for the treatment. No other compensation will be offered by St. Joseph's/Candler Health System. You are not waiving any legal rights, however, nor are you releasing the hospital or study staff from liability for negligence unrelated to the nature and risk of the treatment.

COMPENSATION AND CONFLICT OF INTEREST CLAUSE

St. Joseph's/Candler Health System and your study staff do not have a direct financial interest with the sponsor or in the final results of the study. However, each study participant will receive a $15.00 gift card as compensation for their time.

QUESTIONS ABOUT THE STUDY

If you have any questions concerning your participation in this study, or if you feel you have experienced a research-related injury, you should contact the study staff, Dr. Dan Riddle or Nancy Henderson (Wofford) at 912-856-3788. This is a 24-hour number in case of research-related emergencies that may occur after normal business hours.

If you have any questions about your rights as a research subject, you may contact:

St. Joseph's/Candler Health System Institutional Review Board

Dr. Harold A. Black, Chair

912-819-8087

The Institutional Review Board is a committee that monitors the safety and welfare of research subjects at St. Joseph's/Candler Health System.

PARTICIPATION/AUTHORIZATION

Your participation in this research study is voluntary. You have the right to decline participation or to withdraw from this study at any time. This will in no way affect your current or future medical care. If you decide to withdraw from this research study, you must inform the study staff.

You may also decide to take away your permission to use or disclose personal information about your health. If you choose to withdraw your permission, you must notify the study staff in writing. The study staff's mailing address is 10935 Abercorn Street, Savannah, GA, 31419, 912-856-3788. The study staff will still be able to use the information collected about you prior to your withdrawal from the study. Information that has already been sent to the study sponsor cannot be withdrawn.

Your participation in this study may be stopped at any time by the study staff without your consent. The reasons might include:

- the study staff thinks it necessary for your health or safety;
- you have not followed study instructions;
- administrative reasons require your withdrawal.

COSTS

There are no costs associated with your participation in this study.

CONFLICT OF INTEREST

Conflict of interest means a situation in which a member of the local research team for this study, including the study staff, and study coordinator(s), has a significant financial interest or other personal involvement that may compromise, or have the appearance of compromising, his or her professional judgment or integrity in conducting this study. No member of the local research team has a conflict of interest for this study**.**

STATEMENT OF CONSENT

I have read the above description of this research study. I have been informed of the risks and benefits involved, and all my questions have been answered to my satisfaction. By signing this form, I voluntarily consent to participate in the research study.

Unless I authorize the use and disclosure of my personal health information, I cannot participate in this research study. If I refuse to give my authorization, my medical care will not be affected.

I have received a copy of this consent form for my records.

Printed Name of Subject	*Signature of Subject*	*Date*
Printed Name of Person Obtaining Consent	*Signature of Person Obtaining Consent*	*Date*
Printed Name of Witness, if required	*Signature of Witness*	*Date*
Printed Name of Subject's Legally Authorized Representative, if required	*Signature of Subject's Legally Authorized Representative*	*Date*

APPENDIX K

Subject #_____

Date:_____

Patient Acceptable Symptom State (PASS) Questionnaire

Please circle one of the two responses below when answering the following question:

Taking into account all the activities you have during your daily life, your level of pain, and also your functional impairment, do you consider that your current state is satisfactory?

Yes

No

APPENDIX L

Subject #_____

Date:_____

COMPLICATIONS SURVEY

1. For how long did you remain in the hospital after your total knee replacement?
2. Did you have to be re-hospitalized after your initial discharge (yes or no)?
    a. If yes, for what reason were you hospitalized after your initial discharge?
    b. For how long did you remain in the hospital?

Curriculum Vitae

Nancy Henderson (Wofford), PT, DPT, OCS, CMTPT, Cert MDT

Georgia Southern University (Armstrong State University)

EDUCATION

Virginia Commonwealth University, Richmond, VA

PhD in Health Related Sciences (concentration physical therapy) **In progress**

Dissertation adviser: Dr. Dan Riddle

Elon University, Elon, NC **June 2004**

Transitional Doctorate of Physical Therapy

Elon University, Elon, NC **May 2002**

Master of Physical Therapy

Longwood University, Farmville, VA **May 1998**

B.A., Modern Languages, concentration in Spanish

LICENSURE AND CERTIFICATIONS

Physical Therapy Licensure Georgia (PT008582) **2006–Present**

Physical Therapy Licensure North Carolina (PT008582) **2002–2006**

Board Certified Clinical Specialist **2008–Present**

Orthopedic Physical Therapy

American Board of Physical Therapy Specialists

Certified Myofascial Trigger Point Therapist; **2015-Present**

Myopain Institute

Certified Specialist in Mechanical Diagnosis and Therapy; **2005–Present**

McKenzie Institute

TEACHING EXPERIENCE

Georgia Southern University, Savannah, GA **August 2010–Present**

Assistant Professor, Doctor of Physical Therapy Program

- **PHTH 7262: Physical Therapy Practice Issues II (Fall)- 2014-present**
  Develop and manage all aspects of course, which covers research methods, biostatistics, and introduction to multivariate statistics
- **PHTH 7232: Foundations of Examination, Evaluation, and Intervention (Fall)**
  Co-teach class; serve as primary instructor for the evaluation and treatment intervention for cervical, thoracic, lumbar, sacroiliac, and hip dysfunction.
- **PHTH 7212: Introduction to Pathophysiology II (Fall)**
  Teach the cervical and lumbar surgeries portion of the course.
- **PHTH 9901: Physical Therapy Project I (Fall)**
  Serve as Research Committee Chair for student thesis group.
- **PHTH 9902: Physical Therapy Project I (Spring)**
  Serve as Research Committee Chair for student thesis group.
- **PHTH 9904: Physical Therapy Project IV (Spring)**
  Serve as Research Committee Chair for student thesis group.
- **SMED-3005: Applied Anatomy and Kinesiology (Fall, spring, summer)-2010-spring 2015**
  Develop and manage all aspects of course, which covers bony and soft tissue anatomy for the musculoskeletal system. Lecture on the basics of exercise and development of exercise programs.
- **PHTH 7390: Case Management (Spring)**
  Managed a simulation clinic for first-year DPT students. Students learned all the different PT clinic roles and managed a full case load of mock patients.
- **PHTH 7632: Adv. Foundations of Examination, Evaluation and Intervention II (Spring)**
  Assisted with instruction of advanced clinical techniques, including examination and treatment of spinal dysfunction.

Armstrong Atlantic State University, Savannah, GA **August 2009–August 2010**

Full-time Instructor, Doctor of Physical Therapy Program

- **PHTH 5232: Foundations of Examination, Evaluation, and Intervention (Fall)**

Co-teach class; serve as primary instructor for the evaluation and treatment intervention for cervical, thoracic, lumbar, sacroiliac, and hip dysfunction.

- **PHTH 5212: Introduction to Pathophysiology II (Fall)**
  Taught the cervical and lumbar surgeries portion of the course.

- **SMED-5005: Musculoskeletal Basis of Exercise (Fall and Spring)**
  Developed course, which covered the bony and soft tissue anatomy of the musculoskeletal system. Lectured on the basics of exercise and development of exercise programs.

- **SMED-5055: Pathophysiology of Sports Medicine (Fall)**
  Developed course, which covered the pathophysiology of the body systems. Lectured on the influence of exercise on each system.

- **PHTH 7390: Case Management (Spring)**
  Managed a simulation clinic for first-year DPT students. Students learned all the different PT clinic roles and managed a full case load of mock patients.

- **PHTH 7632: Adv. Foundations of Examination, Evaluation and Intervention II (Spring)**
  Assisted primary instructor with the instruction of advanced clinical techniques, including examination and treatment of spinal dysfunction.

- **RHAB 4111: Pathophysiology (Spring)**
  Co-taught class; taught pathophysiology of the musculoskeletal system.

Armstrong Atlantic State University, Savannah, GA August 2006– August 2009

Adjunct Instructor, Doctor of Physical Therapy Program

- **PHTH 5232: Foundations of Examination, Evaluation, and Intervention (Fall)**
  Co-teach class; serve as primary instructor for the evaluation and treatment intervention for cervical, thoracic, lumbar, sacroiliac, and hip dysfunction.

- **SMED-5005: Musculoskeletal Basis of Exercise (Fall)**
  Developed course, which covered the bony and soft tissue anatomy for the musculoskeletal system. Lectured on the basics of exercise and development of exercise programs.

PROFESSIONAL EXPERIENCE

U.S. Government; Winn Army Community Hospital, Hinesville, GA July 2010–Present

PRN Physical Therapist

- Evaluate and treat active-duty soldiers in a military outpatient orthopedic setting
- Have the ability to order medications and diagnostic imaging as needed.

West Rehab, Hinesville, GA Jan. 2010–2015

PRN Physical Therapist

- Treat patients in a private practice outpatient orthopedic setting.

Tricare Humana Services; Winn Army Community Hospital Oct. 2007–July 2009

Hinesville, GA

Physical Therapist

- Had the ability to order medications and diagnostic imaging as needed.

Effingham Rehab Services, Rincon, GA Feb. 2006–Oct. 2007

Clinic Director/Physical Therapist

- Managed a private practice orthopedic clinic of both clinical and administrative staff.
- Served as clinical coordinator of education and clinical instructor.

Washington Harris Group, Womack Army Medical Center Oct. 2005–Feb. 2006

Fayetteville, NC

Physical Therapist

- Served as physical therapist for the Neurosurgery Telemedicine project with Walter Reed Medical Center.

Spectrum Healthcare Resources; Womack Army Medical Center March 2004–Sept. 2005

Fayetteville, NC

Physical Therapist

- Awarded Employee of the Month (August, 2005)

Cape Fear Valley Health System; Fayetteville, NC July 2002–March 2004

Physical Therapist

RESEARCH

Research interests

Injury Prevention • Musculoskeletal impact of physical training on active-duty soldiers • Running form • Assessing meaningful change

Peer Reviewed Publications

- Worst, H., **Henderson, N.**, Decarreau, R., & Davies, G. A Novel Test to Assess Change of Direction: Development, Reliability, and Rehabilitation. Considerations. *Int J Sports Phys Ther*. 2019;14(2):1-9.
- **Henderson, N**., Decarreau, R., Worst, H., & Davies, G. "Ultrasound Measurements of Humeral Translations during Shoulder Accessory Passive Motion Testing in Healthy Individuals." *Int J Sports Phys Ther.* 2016;11(5):746-56.
- Elliott, T., Marshall, K., Davies, G., & **Wofford, N**. "The Effect of Sitting on Stability Balls on Non-Specific Low Back Pain, Disability, and Core Endurance: A Randomized Controlled Crossover Study." *Spine*. 2016;41(18): E1074-80.
- Coffman N, Herlocker A, Mainer M, Pertain S, **Wofford N**, Campbell M, Mincer A, Lake D. "Determining the Best Method for Teaching Subjects to Activate their Transverse Abdominal Muscle: A Randomized Controlled Trial" *Orthopaedic Physical Therapy Practice Journal*. 2014: 26(1);28-35.
- Crowell M, **Wofford N.** Lumbopelvic Manipulation in Patients with Patellofemoral Pain Syndrome. *J Man Manip Ther. 2012:*20(3);113-120.
- Lake D, **Wofford N.** The Effect of Therapeutic Modalities on Patients with Patellofemoral Pain Syndrome: A Systematic Review. *Sports Health*. 2011: 3(2);182-189

Peer Reviewed Abstracts

- Crowell M, **Wofford N.** Lumbopelvic Manipulation in Patients with Patellofemoral Pain Syndrome (Abstract). *Journal of Orthopaedic and Sports Physical Therapy. 2012:*42(1);A41-59.

Book Chapters

- **Henderson N**, Worst H, Decarreau R, Cook, J. Home Study Continuing Education Monograph, American Physical Therapy Association: 2018.

Peer Reviewed Presentations

- **Henderson, N**, Decarreau R. "Body Fat Management in the Post-operative tactical athlete." Speaker Presentation, Combined Sections meeting, 2019.

- **Henderson, N**, Wilburn, K. "Novel Approaches to Education with Tactical Athletes." Speaker Presentation, Combined Sections meeting, 2019.
- **Henderson, N**, Shing, T, Kardouni J. "Association between MRI and Career Attrition in Soldiers." Platform Presentation, Combined Sections meeting, 2019.
- Westrick R, Wilburn K, **Henderson N**. Development of Military-Specific Outcomes Measures: The Military Upper Quarter Functional Scale. Poster Presentation, Combined Sections, 2019
- Spaid, B, Langevin, M, Sullivan, L, **Henderson, N**. "The Effect of Fatigue on Footstrike in Recreational Runners." Poster Presentation, Combined Sections meeting, 2018
- Westrick R, Wilburn K, **Henderson N**. Development of Military-Specific Outcomes Measures: The Military Lower Quarter Functional Scale. Poster Presentation, Combined Sections meeting, 2018.
- Worst H, **Henderson N**, Decarreau R, Davies G. "A novel running test; the dynamic, agility, lateral performance test." Poster Presentation, Combined Sections, 2018.
- Spaid, B, Langevin, M, Sullivan, L, **Henderson, N**. "The Effect of Fatigue on Footstrike in Recreational Runners." Speaker presentation, Physical Therapy Association of Georgia Conference, 2018.
- **Henderson, N**, Dummar, M. "Soldier Athlete Human Performance Optimization Program." Speaker Presentation, Tactical Strength and Conditioning Conference, 2018.
- **Henderson N,** Decarreau R "The Effect of Body Fat on Military Performance." Speaker Presentation, Tactical Strength and Conditioning Conference, 2017
- **Henderson, N.,** Decarreau, R., Worst, H., & Davies, G. "Ultrasound Measurement of Humeral Glide during Shoulder Joint Mobilizations in Healthy Individuals." Poster Presentation, Combined Sections Meeting, 2016.
- Burdette, A., Cattanach, E., Wilkins, R., **Wofford, N**., Decarreau, R., & Davies, D. "Accuracy of Self-Perception of Footstrike while Running." Poster Presentation, Combined Sections Meeting, 2016.
- Bebe, A., Brackett, G, Davis, J., Tran, L., Davies, G., & Wofford, N. "The effectiveness of low volume versus very low volume upper extremity plyometric exercises on shoulder performance" Poster Presentation, Combined Sections, Meeting, 2016.
- Brett, A., Hayes, A., Headrick, J., Huggins, T., Maher, C., Sikes, B., Thompson, A., **Wofford, N**., Mincer, A., & Schaeffer, K. "Comparison of Conservative Treatments for Plantar Heel Pain." Poster Presentation, Combined Sections Meeting, 2016.
- Henry, A., Motes, M. Roberts, S., Davies, G., **Wofford, N**., & Reinmann, B. "The Scientific Basis Underlying Exercise of the Shoulder Internal Rotators: Concentric, Eccentric, Ballistic, and Plyometrics." Poster Presentation, Combined Sections Meeting, 2016.
- Angus N, Schweizer K, Burkhalter J, **Wofford N**., Fletcher J, Lake D. "Comparing Functional Movement Screen Scores at a Brigade Level Across Military Occupation Specialties in an Active-Duty Population." Poster Presentation, Combined Sections Meeting, 2015.
- Elliott T, Marshall K, Davies G, **Wofford N.** "The Effect of Sitting on Stability Balls on Nonspecific Low Back Pain and Core Endurance: A Randomized, Controlled, Prospective Study." Poster Presentation, Combined Sections Meeting, 2015.
- Burkett D, Halby G, Davies G, **Wofford N.** "A Comparison of Isolated Total Leg Strengthening Versus Functional Training on Strength, Power, Balance, and Agility in Healthy Individuals." Poster Presentation, Combined Sections Meeting, 2015.

- Hodgdon K, Davis S, Davies G, **Wofford N.** "The Effects of Hip Abductor and External Rotator Fatigue in Patients with Patellofemoral Pain Syndrome Compared to Health Individuals." Poster Presentation, Combined Sections Meeting, 2015.
- Holzwarth A, McMahan D, Keen L, Davies G, **Wofford N.** "Prospective Randomized Training Study Evaluating Low- Versus Moderate-Volume Plyometric Training on Functional Outcome Measures of the Shoulder." Poster Presentation, Combined Sections Meeting, 2014.
- Farthing D, Jackson A, Smith M, **Wofford N.,** Mincer A. "Determining the Most Effective Feedback Method in Teaching the Abdominal Drawing-in Maneuver in Subjects with Low Back Pain: A Randomized Clinical Trial." Poster Presentation, Combined Sections Meeting, 2014.
- Fitzpatrick K, Anderson A, Bond J, Morris D, Davies G, **Wofford N.** "A Prospective Randomized Clinical Trial Comparing the Effects of Open-Chain, Closed-Chain, and Combined Open- and Closed-Chain Exercises on Subjects with Patellofemoral Pain Syndrome." Poster Presentation, Combined Sections Meeting, 2014.
- Coffman N, Herlocker A, Manior M, Pertain S, **Wofford N**. "A Comparison of Various Instructional Methods for Contraction of the Transverse Abdominis Using Real-Time Ultrasound: A Randomized Trial." Poster Presentation, Combined Sections Meeting, 2013.
- Crowell M, **Wofford N.** "Lumbopelvic Manipulation in Patients with Patellofemoral Pain Syndrome." Platform presentation, Combined Sections Meeting, 2012.
- Dale D, Fulp K, Harrell A, Posey A, Thompson AW, Mincer A, **Wofford N. "**Management of Plantar Fasciitis with Conservative Therapies: A Randomized Trial." Poster Presentation, Combined Sections Meeting, 2012.
- **Wofford N.** "Classification Approach to Low Back Pain." Physical Therapy Association of Georgia. Spring Conference, Atlanta GA, April 2010.

Non-Peer Reviewed Presentations

- **Henderson N**. "International Healthcare." Interprofessional Journal Club session, Interprofessional Journal Club. (February 2019)
- **Henderson N.** Running Course, as part of the Marne Warrior Athlete PRT Academy. Fort Stewart .(July 2017-2018).
- **Wofford N.** "Giving Back: An International Volunteer Experience." Alpha Eta Allied Health Honor Society, Armstrong Atlantic State University. (March 2011)
- **Wofford N.** "The Effect of Therapeutic Modalities on Patients with Patellofemoral Pain Syndrome: A Systematic Review using the Bizzini and Pedro methods." Local orthopaedic journal club session, Savannah GA (February 2011)
- **Wofford N.** "Classification Approach to Low Back Pain" International Volunteer Experience, Lima Peru, May-June 2010.
- **Wofford N.** "Foot Orthoses and Their Effects on Lower Extremity Biomechanics." Local orthopaedic journal club session, Savannah GA. (March 2009).
- **Wofford N. "**Update on Spinal Clinical Prediction Rules." Local orthopaedic journal club session, Savannah GA (October 2009).

Graduate students

Serve as Doctoral Physical Therapy Capstone Project Committee Member or Chair for >75 students in Physical Therapy department at Georgia Southern University

- "The Effect of Instrumented Assisted Soft Tissue Mobilization on Hamstrings length." (Committee chair) 2018-present
- "The Effect of Eccentric Training on Hamstrings length." (Committee chair) 2018-present
- "The Effect of Fatigue on Footstrike in Recreational Runners" (Committee chair) 2017-2019
- "Ability of Runners to Change Foot Strike from Rear foot to Mid foot or Forefoot with the use of a Metronome and POSE Running Techniques" (Committee chair) 2016-present
- "The Association between muscle length and footstrike in recreational runners" (Committee Member) 2017-present
- "Scientific Basis Underlying Shoulder Exercises: Concentric, Eccentric, Ballistic, and Plyometrics." (Committee Member) 2014-2016
- "Comparing Self-Perception of Foot Strike to Actual Foot Strike." (Faculty Chair) 2014-present
- "The Effects of low versus very low volume plyometric training on functional outcome measures." (Committee Member) 2014-2016
- "Management of Plantar Fascitis with Conservative Therapies: Modalities or Manual Therapy." (Committee Member) 2014-2016
- "CrossFit: How much do they really know?" (Committee Member) 2014-present
- "Effects of Isolate versus Multi-Joint Training on the Shoulder Complex." (Committee Member) 2014-2016
- "Comparing Functional Movement Screen Scores at a Brigade Level Across Military Occupation Specialties in an Active-Duty Population." (Faculty Chair) 2013-2015
- "The Effect of Sitting on Stability Balls on Nonspecific Low Back Pain and Core Endurance: A Randomized, Controlled, Prospective Study." (Committee Member) 2013-2015
- "A Comparison of Isolated Total Leg Strengthening Versus Functional Training on Strength, Power, Balance, and Agility in Healthy Individuals." (Committee Member) 2013-2015
- "The Effects of Hip Abductor and External Rotator Fatigue in Patients with Patellofemoral Pain Syndrome Compared to Health Individuals." (Committee Member) 2013-2015
- "Prospective Randomized Training Study Evaluating Low- Versus Moderate-Volume Plyometric Training on Functional Outcome Measures of the Shoulder." (Committee Member) 2012-2014
- "Determining the Most Effective Feedback Method in Teaching the Abdominal Drawing-in Maneuver in Subjects with Low Back Pain: A Randomized Clinical Trial." (Faculty Chair) 2012-2014
- "A Prospective Randomized Clinical Trial Comparing the Effects of Open-Chain, Closed-Chain, and Combined Open- and Closed-Chain Exercises on Subjects with Patellofemoral Pain Syndrome." (Committee Member) 2012-2014
- "A Comparison of Various Instructional Methods for Contraction of the Transverse Abdominis Using Real-Time Ultrasound: A Randomized Trial" (Faculty Chair) 2011-2013
- "Management of Plantar Fascistis with Conservative Therapies: A Randomized Trial" (Committee Member) 2009–2011
- "A Systematic Review of Lumbar Stabilization Exercises" (Committee Member) 2008–2009

Research projects in progress

- **Henderson, N**, Shing, T, Kardouni J. "Association between MRI and Career Attrition in Soldiers." Manuscript preparation.
- Westrick R, Wilburn K, **Henderson N**. "Development of Military-Specific Outcomes Measures: The Military Lower Quarter Functional Scale." Manuscript preparation.
- Westrick R, Wilburn K, **Henderson N**. "Development of Military-Specific Outcomes Measures: The Military Upper Quarter Functional Scale." Data collection.

GRANTS

External Grants

- Associate investigator, "Rapid Access to Comprehensive Expertise to improve musculoskeletal injury-related outcomes (RACE to improve MSKI-related Outcomes)." Accepted, Joint Program Committee 5, 2019, 3.4 million dollars.
- Primary Investigator, "Smallest Worthwhile Effect Values for a Total Knee Arthroplasty." Accepted, Physical Therapy Association of Georgia, 2015, $500.00

Internal Grants

- Co-Investigator, "Reducing Back Pain Symptoms in Migrant Farmworkers: A Pilot Study." Accepted, Georgia Southern University, 2012, $10,032

ACADEMIC SERVICE

Professional

Journals

- Reviewer*: Tactical Strength and Conditioning;* 2017-present
- Reviewer: *Archives of Physical Medicine and Rehab;* 2014-present
- Reviewer: *Sports Health;* 2014-present
- Reviewer: *Archives of Gerontology and Geriatrics;* 2011-present

Books

- Reviewer: *100 Orthopaedic Cases.* Pearson Education.
- Reviewer: Cook C, Hegedus E. *Orthopedic Physical Examination Tests*. 2nd ed. Upper Saddle River, NJ: Pearson Education, Inc; 2013.

University

- Development and implementation of the Soldier Athlete Human Performance Optimization program (January 2017-present)
- Member, Graduate Faculty Status Committee (2014-present)
- Reviewer, SPARC II study abroad grant (December 2014)
- Reviewer, SPARC II study abroad grant (March 2014)

College

- Organize and lead yearly Department of Rehabilitation Sciences Service Trip to Costa Rica (2017-Present)
- Member, Curriculum Committee (2014-2015)
- Member, Interdisciplinary Healthcare Task Force, Armstrong Atlantic State University (2011-2012)
- Member, Search Committee for Rehabilitation Sciences department chair, Armstrong Atlantic State University (2012-present)
- Member, Search Committee for CSDS program director, Armstrong Atlantic State University (2011)
- Member, Advisory Committee to the Dean, College of Health Professions, Armstrong Atlantic State University (2010–December 2011)

Department

- Organize and lead yearly Department of Rehabilitation Sciences Service Trip to Costa Rica (2017-Present)
- Rho Tau Faculty Advisor, Armstrong Atlantic State University (2010–2015)

Program

- Chair, Nontenure Expectations Committee, 2018-Present
- Member, White Coat Ceremony Committee, 2018-Present
- Member, Search Committee for Director of Clinical Education, 2019-Present
- Development and implementation of the Soldier Athlete Human Performance Optimization project (January 2017-present)
- Organize and lead yearly Department of Rehabilitation Sciences Service Trip to Costa Rica (2017-Present)
- Member, Faculty Search Committee (2015-2016)
- Member, Faculty Search Committee (2014-2015)
- Member, Faculty Search Committee (2013-2014)
- Member, Department Head Search Committee (2012)
- Assisted in revising departmental accreditation outcomes for The Commission on Accreditation in Physical Therapy Education (CAPTE; 2010–2013)

- Member, Curriculum Committee, Department of Physical Therapy, Armstrong Atlantic State University (2010-present)
- Member and Admissions Interviewer, DPT Admissions Committee, Armstrong Atlantic State University (2009–Present)

PROFESSIONAL ORGANIZATION SERVICE

International

- International volunteer for Healthcare Volunteers Overseas in Lima, Peru. Evaluated patients with spinal pain, mentored physical therapists and physical therapy students, and lectured daily on various topics related to low back pain. Lectures were given in both English and Spanish. Presented recap of volunteer experience to AASU PT students and faculty upon return. (May–June 2010)

National

- Member, Education Task Force, Imaging Special Interest Group for the American Physical Therapy Association (2012-present)

State

- Active Member; Physical Therapy Association of Georgia (PTAG; 2006–Present)
- Member of the PTAG Nominating Committee (2010–2012)
- District 2 Director for PTAG (2008–2012)
- PTAG Board Member (2008–2012)

Local

- Development and implementation of the Soldier Athlete Human Performance Optimization project (January 2017- present)
- Serve as the Georgia Southern representative to the Community Health Promotions Council with the 3rd Infantry Division (January 2017- present)
- Serve as the Georgia Southern representative to the Provider Resiliency Working Group at Winn Army Hospital (2016-present)
- Organize monthly local orthopaedic journal club sessions (2009–2018)
- Provided Functional Movement Screens to the local community, alongside DPT students, at Georgia GameChangers (fall, 2015)
- Assist 1st brigade surgeon with injury prevention measures (2014-2015)
- Serve as Race Director for PT Month 5K run/walk to raise money for Armstrong State University's PT club (2009–2010)
- Provide pro bono services at the Good Samaritan clinic (2011-2016)
- Provide volunteer service at the Migrant Health Fair (Feb 2011)

HONORS AND AWARDS

- Employee of the Month at Spectrum Healthcare Resources; Womack Army Medical Center (August 2005)
- Clinical Excellence Award, Master of Physical Therapy Program, Elon University (May 2002)

PROFESSIONAL MEMBERSHIPS

American Physical Therapy Association (APTA; 2002–present) • Orthopaedic and Sports Chapter of APTA (2002–present) • Health Care Policy and Administration Chapter of APTA (2008–2009) • Physical Therapy Association of Georgia (PTAG; 2006–present) • PTAG District Director and Board Member (2008–present)

CONTINUING EDUCATION (PAST 5 YEARS)

- APTA Combined Sections Meeting (January 2019)
- Tactical Strength and Conditioning Conference (April 2018)
- APTA Combined Sections Meeting (February 2018)
- APTA Combined Sections Meeting (February 2017)
- Tactical Strength and Conditioning Conference (April 2017)
- APTA Combined Sections Meeting (February 2016)
- APTA Combined Sections Meeting (February 2015)
- Myopain Dry Needling Course IV (April 2015)
- Myopain Dry Needling Course III (March 2015)
- Myopain Dry Needling Course II (October 2014)
- Myopain Dry Needling Course I (September 2014)
- APTA Combined Sections Meeting (February 2014)
- APTA Combined Sections Meeting (February 2013)
- APTA Combined Sections Meeting (February 2012)
- APTA Combined Sections Meeting (February 2011)
- APTA-approved monthly local orthopedic journal club sessions (2006–present)

REFERENCES

- Dr. Jim Karnes: (716) 908-7243; jkarnes@georgiasouthern.edu

- Dr. Haley Worst: (706) 344-7647; hworst@georgiasouthern.edu

- Dr. Kristen Wilburn: (210) 416-0029; kristenzosel@gmail.com