Aus der Klinik für Anästhesiologie

(Univ.-Prof. Dr. med. K. Meissner)

der Medizinischen Fakultät der Universität Göttingen

The impact of expectation on patient-rated global effectiveness of total hip arthroplasty

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades

der Medizinischen Fakultät der

Georg-August-Universität zu Göttingen

vorgelegt von

Virginie Sien-pin Celine Ng Kuet Leong

aus

Kassel

Göttingen 2021

Dekan: Prof. Dr. med. W. Brück

Betreuungsausschuss

Betreuer/in: Prof. Dr. med. F. Petzke

Ko-Betreuer/in: Prof. Dr. rer. Biol. Hum. Dipl.- Psych. N. von Steinbüchel-Rheinwall

Prüfungskommission

Referent/in: Prof. Dr. med. F. Petzke

Ko-Referent/in: …..…….

Drittreferent/in: …..…….

Datum der mündlichen Prüfung: ……...

Hiermit erkläre ich, die Dissertation mit dem Titel "The impact of expectation on patient-rated global effectiveness of total hip arthroplasty" eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Göttingen, den …… ……

(Unterschrift)

Die Daten, auf denen die vorliegende Arbeit basiert, wurden teilweise publiziert:

Kästner A, **Ng Kuet Leong VSC**, Petzke F, Budde S, Przemeck M, Müller M, Erlenwein J (2021): The virtue of optimistic realism - expectation fulfillment predicts patient-rated global effectiveness of total hip arthroplasty. BMC Musculoskelet Disord 22, 180

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BMI	Body Mass Index
CPG	Chronic Pain Grade (von Korff)
CTS	Catastrophizing Thought Scale
DASS	Depression, Anxiety and Stress Scales
KPI	Kiel Pain Inventory
MPSS	Mainz Pain Staging System
NASS	North American Spine Society
NRS	Numeric Rating Scale
PHQ-15	Patient Health Questionnaire for Physical Symptoms
PROs	Patient Reported Outcomes
QUIPS	Qualitätsverbesserung in der postoperativen Schmerztherapie (Quality improvement in postoperative pain management)
SF-12	Short Form of the Health Survey Questionnaire
THA	Total Hip Arthroplasty
THS	Thought Helplessness Scale
TSK	Tampa Scale for Kinesiophobia
TSS	Thought Suppression Scale
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index
VAS	Visual Analogue Scale

1 Introduction

In the current era of medical practice, patient-based assessments, such as measures of patient satisfaction or patient's perception of effectiveness of surgery, are gaining in importance in the quality assessment of healthcare (Weldring and Smith 2013). A patient-centered assessment serves as a tool for quality evaluation of the health care provided (Mannion et al. 2009a). Additionally, it serves as a possible indicator of the likelihood that a patient, who is unsatisfied with the outcome, will take advantage of further medical resources resulting in greater costs for the healthcare system (Mannion et al. 2009a). It is, therefore, imperative to investigate patient-dependent variables, both biological and psychological, as well as patient-independent variables that can influence patient-reported outcomes after a medical intervention. Multiples studies show that factors, such as disturbances in physical and functional abilities, pain, mental illness, cognition and psychological distress, correlate with poor surgical outcomes after orthopedic surgeries (Rolfson et al. 2009; Ayers et al. 2013; Utrillas-Compaired et al. 2014; Khatib et al. 2015; Gil et al. 2018).

Expectations have an impact on outcome after a surgical intervention (Mahomed et al. 2002; Mannion et al. 2009a; Yoo et al. 2019). An increase in research concerning the impact of preoperative expectations on outcome after orthopedic surgeries has been observed. Reports of the exact nature of this relationship however remain a matter of controversy.

1.1 Expectations

The notion that expectation is a crucial determinant of behavior and experience has been thoroughly investigated in many psychological theories of learning (Peerdeman et al. 2016). The learning process of classical conditioning, which has been one of the most researched learning mechanisms, is nowadays attributed to outcome expectancies (Rescorla 1988; Pavlov 2010; Kirsch et al. 2014). Classical conditioning is the learning procedure of associating an initially neutral stimulus (such as a bell) to a biologically potent stimulus (for example food) and, over time, learning a conditioned response (for example salivating on hearing the bell) (Rescorla 1988).

The more modern terminology of outcome expectancy describes the perceived likelihood of a specific outcome occurring as a result of another stimulus (behavior or event) (Peerdeman et al. 2016). Outcome expectancy has been further divided into stimulus-outcome expectancy and behavior-outcome expectancy (Kirsch 1985). Stimulus-outcome expectancy pertains to the belief that a certain external event or stimulus provides a cue for the possible occurrence of other events (Kirsch 1985; Kirsch 1997). Stimulus-outcome expectancy can be further categorized into two types, depending on whether the outcome is an environmental event or a nonvolitional response (Kirsch 1999). Respectively, when focusing on the result, behavior-outcome expectancy can also be divided into outcome-response expectancy when the result is a nonvolitional response and outcome-behavior expectancy when the result is a volitional response (Bandura 1997). Behavior-outcome expectancy is the belief that a certain behavior will bring about a specific result or consequence (Kirsch 1999).

Rotter (1954) argued that the combination of the outcome expectancy and the importance assigned by the individual to the outcome was the true determinant of behavior, in accordance with the idea that cognitions are as important as outcome expectations in determining behavior. This was further developed by Bandura (1977), who theorized that behavior was not only determined by expectancies about the outcome, but also by beliefs in one's own capabilities of achieving the behavior and coined this effect self-efficacy expectancy. Self-efficacy expectancy does not only signal one's conviction in one's competence to act a specific behavioral outcome, but also encompasses beliefs about one's ability to establish the required "motivation, cognitive resources and courses of action" to achieve the expected goal (Bandura 1990).

In conclusion, there is no single definition of expectation in the literature, which has led to a significant amount of confusion in the terminology and the use of different models of expectancies in clinical settings. Although multiple models of expectancies have been presented in the literature, outcome expectancy has the largest empirical support and is the only form of expectancy addressed in this analysis (Peerdeman et al. 2016).

1.2 Total hip arthroplasty

Osteoarthritis is the leading joint disease in the world (Weltgesundheitsorganisation 2003; Fuchs et al. 2017). In 2014/2015 the Robert Koch Institute in Germany carried out a health survey which showed that 18% of adults, over the age of 18, suffered from osteoarthritis

in the last year (Fuchs et al. 2017). With an aging population in Germany, the prevalence of osteoarthritis is expected to increase substantially (Fuchs et al.2017).

The first published reports of Total Hip Arthroplasty (THA) date as far back as the 1940s, with Sir John Charnley introducing his concept of the low-friction THA in the 1960s (Shon et al. 2019). In those times the outcome of the surgical intervention was disappointing to say the least. Today, when reaching the limits of conservative therapy THA is an efficient surgical intervention for severe hip osteoarthritis, inflammatory arthritis, dysplasia and fracture (Higashi and Barendregt 2011; Jourdan et al. 2012). In 2017, about 256 000 hip replacement surgeries were performed in Germany alone (Statistik zu Hüftoperationen nach Ländern). Ultimately, total hip arthroplasty aims to reduce pain, increase hip function and flexibility and thereby enhance the quality of life for patients (Nilsdotter 2003; Ng et al. 2007). The majority of patients undergoing THA today benefit greatly from the intervention and experience the desired improvement in symptoms. Unfortunately, despite the progress in surgical techniques and prosthesis design, 7-15% of patients undergoing hip replacement, report being dissatisfied with the surgery (Jones et al. 2007; Anakwe et al. 2011) amounting to approximately 28 000 people per year in Germany.

Discrepancies have been observed between clinicians' and patients' opinions about the success of an orthopedic intervention (Baker et al. 2007). Patients may still be unsatisfied with the result of a surgery while the surgery has been, from a medical point of view, a success (Burton et al. 1979). In 2001 Bullens et al. found a poor correlation between patient-based satisfaction measured by a visual analog scale (VAS) and the objective surgeon-assessed knee score after knee replacement surgery. While patients put value on the knee's global functionality, the clinicians tended to concentrate on alignment, motion range and stability (Bullens et al. 2001). Anakwe et al. (2011) demonstrated that the restoration of function, pain relief, postoperative objective functional scores and expectation fulfilment correlated closely with patient satisfaction with THA.

Patients' assessment of the efficacy of already undergone treatment and their postoperative satisfaction with the respective outcome influences their desire to undergo further surgical interventions (Gonzalez Saenz de Tejada et al. 2014; Schaal et al. 2016). In the field of surgical methods and prosthetic implants, extensive technological advances have been made since the first THA, decreasing the likelihood of future significant breakthroughs in this domain (Palazzo et al. 2014). It has, therefore, become increasingly

important to understand and determine the factors affecting patient satisfaction in order to improve quality control in healthcare systems (Collins and Roos 2012). In research and clinical practice, patient-based assessment has thus become an important method of analyzing the success and effectiveness of a THA from the patient's position (Weldring and Smith 2013).

1.3 Expectations affecting outcome after total hip arthroplasty

Three different theoretical models have been suggested to describe the relationship between patient satisfaction and expectations (Mannion et al. 2009b). According to the first model, patients' expectation, i.e. the level of expectation, could predict satisfaction. Mahomed et al. (2002) demonstrated that patients' expectations determined their level of mobility and satisfaction following a total joint arthroplasty. Their findings also showed that patients who had more optimistic expectations, experienced less pain and disability six months after surgery (Mahomed et al. 2002). Indeed, some studies proclaim that different levels of expectations explain the variance found in outcomes in THA (Mahomed et al. 2002; Cross et al. 2009; Gandhi et al. 2009; Quintana et al. 2009). In orthopedic surgeries, unreasonable expectations among patients were observed with more than half of them describing higher expectations than their surgeon (Ghomrawi et al. 2011). In the pain literature, recent research has shown that persisting unrealistic expectations correlate with low degrees of satisfaction (Iles et al. 2009). Skatteboe et al. (2017) showed that unrealistic expectations, whether positive or negative, lead to worsening of patient outcome measures by patients with musculoskeletal disorders. Other studies, however, failed to demonstrate any link between expectations and outcome (Waljee et al. 2014).

The second model proposes that it is not the preoperative level of expectation that determines satisfaction, but rather the fulfillment of that particular expectation (Anakwe et al. 2011; Palazzo et al. 2014; Ng Kuet Leong et al. 2020). Accordingly, Palazzo et al's. (2014) findings show that patient satisfaction was determined by the level to which the result met the anticipated change. According to Mannion et al. (2009a) and other studies, a fulfilled expectation of a spinal surgery is a significant independent predictor of how effective it is (Haanstra et al. 2012; Tilbury et al. 2016).

Lastly, the third model suggests that the main determinants of patient satisfaction are postoperative clinical improvements in both symptomologies and functional outcomes,

independent of the level of expectations preoperatively or their achievement (Mannion et al. 2009b). Indeed, in a systematic review conducted in 2014, Waljee et al. demonstrated that while some investigations show that higher levels of expectations correlate with higher satisfaction, others have described a negative relationship between expectation and satisfaction with surgery. Also, the analysis found that expectations and perceived satisfaction following surgery were not significantly related in five of the analysed studies (Waljee et al. 2014). Investigations on the impact of expectation and the expectation fulfillment on patient-reported outcome measures, such as satisfaction, have thus been conflicting and incomplete.

1.4 Satisfaction with surgery and its global effectiveness

Clinical studies have increasingly reported on patient satisfaction as a measure of good outcome in orthopedic surgeries (Graham et al. 2014). Shirley and Sanders (2013) argued in their review that although satisfaction is a valuable tool in quality management, there is a substantial number of factors affecting patient satisfaction that are beyond the control of the surgeon. To better understand patient satisfaction, it is important to distinguish satisfaction with regard to the method of care from satisfaction with regard to the result of care (Donabedian 1988). For example, while the medical treatment might have been deemed as effective by both health care providers and patients, the process of care might have been unsatisfying due to high costs, physical and dietary restrictions or even due to an impolite environment in the hospital wards (Graham et al. 2014). It must be remembered that the concept of satisfaction varies between individuals, regardless of the tool used to measure patient satisfaction.

Many methods developed to assess satisfaction are unvalidated or rudimentary, as stated by Graham et al. (2014). While some only measure satisfaction with the treatment process and not with the result, such as the Consumer Assessment of Healthcare Providers and Systems (CAHPS), others often use a direct, one-domain measure such as a VAS with the question being “how satisfied are you with the results?” (Yoo et al. 2019). Certain validated outcome instruments, despite including questions about satisfaction, are not specifically meant to be used to assess patient satisfaction and are thus used incorrectly in literature (Judge et al. 2011a; Dawson et al. 2012). As patient satisfaction has not yet been well defined in orthopedic surgery surveys, the methods of assessment have yielded

irregular and inconclusive results. Therefore, the relevance of satisfaction in determining the effectiveness of a surgical intervention remains controversial.

Today, with emphasis laid on value-based care, the patient's opinion of the postoperative benefit gained has to be taken into account when contemplating the effectiveness of an intervention (Deyo et al. 1994; Deyo et al. 1998). When assessing elective surgical procedures where the therapeutic goal is to achieve functional improvement and pain relief, treatment effectiveness was particularly found to be of great importance (Stucki et al. 1996; Mannion et al. 2009a). Mannion et al. (2009a) also argued that, regardless of treatment satisfaction, the patient's perceived surgery effectiveness is what determines whether the patient will seek additional medical assistance for the same problem which would result in higher healthcare costs. Accordingly, the outcome measure used in this thesis is the patient's direct opinion of the intervention's global effectiveness. Thus, not patient satisfaction, which is highly susceptible to variable factors, but the extent to which the patient believed the surgery had helped him/her was investigated. Most studies to date, however, have investigated patient satisfaction as the outcome measure after orthopedic interventions. As the perceived benefit gained from the surgery influences patients' satisfaction, studies assessing patient satisfaction have been used to select the factors being analyzed in this thesis.

1.5 Thesis

Previous studies have identified several preoperative risk factors for bad outcome after THA such as age, gender, preoperative pain intensity, BMI (Body Mass Index), lower socio-economic status, educational level, preoperative physical function and co-morbidities (Rolfson et al. 2009; Hossain et al. 2011; Buirs et al. 2016). There has also been sufficient evidence demonstrating the impact of psychological distress and mental states and disease, such as anxiety, depressivity and depression, on patient self-reported satisfaction and functional outcome after joint replacement (Franklin et al. 2008; Rolfson et al. 2009; Ayers et al. 2013; Utrillas-Compaired et al. 2014). According to Sullivan et al. (2011) those who catastrophize their pain, have fear of movement because of pain and suffer from depression showed poorer recovery after total knee arthroplasty. Nevertheless, few studies have addressed these psychological factors in their analysis and have failed to assess whether the connection between expectations and outcome in THA is altered by these variables (Haanstra et al. 2012).

To investigate the impact of expectations and expectation fulfillments on patient-based global effectiveness of spinal surgery, Mannion et al. (2009a) used a hierarchical multiple regression analysis. In their study, patients were given two questionnaires; two months before and twelve months after the spinal surgery, in which actual improvements and changes in pain symptoms were measured (Mannion et al. 2009a). Results of their analysis showed that the discrepancy between preoperative expectation and postoperative status (or expectation-actuality-discrepancy) for the "worst symptom" (leg or back pain) was the only predictor of the global effectiveness of treatment. Mannion et al. (2009a) therefore concluded that it is not the expectations that are important determinants of outcome but rather the degree of expectation fulfillments for the "worst symptom" (leg or back pain).

The purpose of the present retrospective cohort analysis was to apply Mannion et al.'s (2009a) methodology to the area of total hip arthroplasty while allowing a more robust and wholesome analysis by including validated measures of psychological, functional and physical health. Twelve months following the intervention, patients' functional, physical and psychological health were assessed again in order to calculate the improvement in symptoms after surgery. Using a sequential multiple regression analysis, the objective of this thesis was to establish the relative relevance of three possible predictors of THA's overall efficacy as reported by patients: the expectations before surgery, their fulfillment and the changes in symptoms after THA.

2 Methods

2.1 Patients and protocol

Patients undergoing elective hip replacement surgery due to osteoarthritis at the Orthopedic University Hospital of the Medical School Hannover from July to November 2012 were invited to partake in the analysis. One day before surgery, patients were admitted to the hospital. After receiving all needed information and signing a written consent form, patients filled out the questionnaires. In the afternoon of the day of admission, the functional status and pressure pain thresholds were measured. Concerning the analgesic therapy, all patients received the standardized pain management protocol developed and accredited for several years at the institution. A follow-up was carried out after twelve months to determine the result of the surgery using structured telephone interviews.

The protocol was approved by the Ethics Committee of the University Hospital of Goettingen (No. 5/4/12) and the Ethics Committee of the Hannover Medical School (No. 1483-2012) and was carried out according to the recommendations of the Declaration of Helsinki.

2.2 Inclusion and exclusion criteria

Patients were recruited by one of the physicians of the Department of Pain Medicine of the University Hospital Göttingen, the Department of Anesthetics and Intensive Care of the Annastift, or the Orthopedic Clinic of the Hannover Medical School.

The inclusion requirements consisted of being older than 19 years, having a surgical indication of unilateral total hip arthroplasty, being fluent in written and spoken German, being mentally and legally able to comprehend the medical information and the meaning/consequence of their participation and to provide written consent, and finally approving the inclusion in the 12-month follow-up review and the provision of contact information. Exclusion requirements consisted of having a planned spinal anesthesia, present drug addiction and dementia.

Twelve months after the intervention, patients with immediate peri- and postoperative complications (for example postoperative delirium, femoral head necrosis, prosthetic joint infection etc.) impacting the postoperative assessments were removed.

2.3 Analysis design

2.3.1 Preoperative procedure

On the preoperative day, patients were admitted on the inpatient ward and were asked to give their written consent. Subsequently, one of the physicians conducted a standardized medical assessment and a physical examination. In addition, to demographic data for example age, gender and education level, prior medication and previous illnesses were also noted. Patients were then asked to undergo functional tests, i.e. the standardized hip range of motion tests and the ability to walk stairs. An extensive characterization of the pain-related symptoms, pain sensitivity and chronicity were done. The intensity of hip pain and loss of mobility due to hip pain were noted, as well as the consumption of analgesics and past pain-related treatments. Expectations, psychosocial factors (anxiety, depression, stress and movement-related anxiety), patients' preoperative quality of life and cognitive pain processing were also recorded (see below).

2.3.2 Postoperative procedure

The patients' psychological status, quality of life, functional capacities, intensity of pain and pain characteristics were evaluated twelve months after surgery via standardized telephone interviews. The global effectiveness of THA, the changes in symptoms and the fulfillment of expectations were also assessed using a protocol with a specific selection of answers.

2.3.3 Patients' reasons to undergo surgery

Patients were requested to reveal the three most relevant reasons that prompted their decision to undergo the surgical intervention. Answer possibilities ranged from "other therapies had not helped, something must be done", "fear of the situation worsening", "to remain independent", "to improve everyday functioning", "to improve walking abilities", "to reduce pain" and "because my physician recommended the surgery". If none of the answers appealed to the patients, they were also given the opportunity to write down their personal answer.

2.3.4 Patients' most important factors for surgical success

To deem the surgical intervention as successful, the most significant aspect that needed to be improved was asked from the patients. Here the choices were improvement in hip pain, back pain, walking ability, independence, physical exercise, general function, social interactions or mental well-being.

2.3.5 Patients' most important postoperative change

Twelve months after the intervention, patients were requested to choose which factor had been the most crucial postoperative change. With the answer possibilities being again: hip pain, back pain, walking ability, independence, physical exercise, general function, social interactions or mental well-being.

2.3.6 Patient-rated global effectiveness of surgery

Taking example on Mannion et al. (2009a), the global effectiveness of the surgery was examined by use of the question, "How did the surgery help you overall?". Patients were given the choice between the following statements: "helped a lot, helped, helped a bit, did not help or made it worse". For the statistical analysis we ranged the answers from a score of 0 for "made it worse" to a score of 4 for "helped a lot".

2.3.7 Expectations

To determine patient expectations of THA an adapted form of the "Expectation Scale" from the North American Spine Society (NASS) Lumbar Spine Questionnaire was applied (Daltroy et al. 1996).

Eight domains including hip pain, walking ability, independence in everyday life, physical exercise, general function, social interactions and mental well-being were assessed.

Patients disclosed on a 6-point scale their expectations for each domain, with the options being "I don't know", "worse", "unchanged", "somewhat better", "better" and "much better". This questionnaire was translated into German for this analysis, as no formally validated German version was available.

Patients were also asked to note the level of fear they had with regard to the surgery itself as well as to the expected postoperative pain on a 10-point scale (1 = no fear and 10 = worst fear imaginable).

2.3.8 Fulfillment of expectations

Twelve months after surgery, patients reported their actual change in each of the previously described eight domains (2.3.7) on a 5-likert scale where a value of 1 was attributed to the answer "worse", a score of 2 to "same", a score of 3 to "somewhat better", a score of 4 to "better" and a score of 5 to "much better". Two methods of evaluating the expectation fulfillment, i.e. the level of correspondence between the anticipated change and the result were used. First, patients were asked the direct question "Have your expectations regarding this domain been fulfilled?" with the answers ranging from "yes" to "partly" and "no". As previously stated, Mannion et al.'s (2009a) analysis demonstrated that the discrepancy between expectation and postoperative actuality for the worst symptom was the sole predictor of global effectiveness of treatment. Therefore, as a second measure of expectation fulfillment (calculated expectation fulfillment), the actuality-expectation discrepancy was calculated by subtracting the expected change for each domain from the respective actual change twelve months after THA. The calculated actuality-expectation discrepancy ranged from a score of -5 to 5 points. A negative difference in actuality-expectation demonstrated less change than anticipated and was called "unmet expectations". A score of zero indicated a result as anticipated and was therefore named "as expected". A positive difference between actuality and expectation reflected a higher change than anticipated and was deemed "exceeded expectations".

2.3.9 Pain characteristics

In addition to the pain assessment during the routine check-up, the physicians further differentiated the aspects of pre-existing pain by means of an adapted form of the German Pain Questionnaire from the German Pain Society (Casser et al. 2012; Petzke et al. 2020). The multidimensional facets of pain were assessed: pain intensity, pain character, pain duration, pain frequency and past pain-related therapies.

2.3.9.1 Stages of pain chronicity

Based on the information of the standardized case history and pain assessment, patients were divided into three stages of pain chronicity (I- III) based on the Mainz Pain Staging

System (MPSS) from Gerbershagen (Gerbershagen et al. 2008). Four aspects of pain chronicity (three questions concerning temporal character, one question asking about spatial character, two questions addressing past medication intake and four questions regarding the career of the patient) were analyzed.

To each question, the answers are given up to three points. Each of the four aspects of pain chronicity is given an overall score from 1 to 3, based on the total sum of points for their respective questions. The sum of all four scores is then used to classify patients into the MPSS stages I to III, where III is the highest grade.

2.3.9.2 Severity of chronic pain

To describe the level of chronic pain and pain-related dysfunction the German version of the Chronic Pain Grade (CPG) according to Korff was used. Patients were asked seven standardized questions, which were then evaluated and assigned to one of four chronic pain stages (Klasen et al. 2004; Dixon et al. 2007). The evaluation consisted of three sections concerning the intensity of pain, three sections regarding disability in everyday life due to the pain and one question concerning the number of days the patients couldn't perform their daily activities due to the pain in the last three months. Here, an 11-point NRS was applied with the worst imaginable pain or no activities anymore represented by a score of 10 and no pain or no limitation represented by a score of 0. The mean of the present, maximum and average pain intensities during the previous three months was used to classify the patients as having either "high" pain intensity (NRS value higher or equal 5) or "low" pain intensity (NRS value lower than 5). Patients were then grouped according to their pain-related limitations in everyday life into four categories. Patients were asked how strongly, over the last three months, the pain limited their ability to participate in everyday tasks (washing, eating, buying groceries etc.), social and family events and lastly work (including household work). The categories were defined as follows: 0-2.9; 3.0-4.9;5.0-6.9;7.0+. These were then scored from of 0 to 3.

The disability-score was obtained by adding the degree of limitations and the frequency of limitations. That is the disability score is the sum of how often the patient was unable to participate in everyday tasks due to the pain in the last three months. Using the disability-score and pain intensity score, patients were assigned to one of these five chronic pain stages: Grade 1 = low impairment (< 3 NRS) and low pain intensity (< 5 NRS), Grade 2 = low impairment (< 3 NRS) and high pain intensity (≥ 5 NRS), Grade 3

= high impairment, moderately limiting (3 – 4 disability score) and Grade 4 = high impairment, severely limiting (5 – 6 disability score).

2.3.9.3 Overall pain sensitivity

Through the use of an electronic pressure algometer multiple pressure pain thresholds (PTT) were assessed in order to determine global pain sensitivity, (Somedic Production, Stockholm, Sweden). The algometers' probe (1cm^2) was placed bilaterally over five sites: the thumbnail, the lateral epicondyle, the trapezius muscle (midway between neck and acromion), the quadriceps femoris (midway between patella and groin) and the tibialis anterior muscles (10 cm distal to the tibial plateau). During pressure stimulation with a steadily rising intensity (50kPa/s), patients were told to signify the moment of first pain experience. The pressure application was halted either when experience of pain was recorded by patients or when the maximal pressure level (which was set at 1000kPa) was achieved. Overall pain sensitivity analyses were obtained by the average threshold (kPa) of all ten test sites.

2.3.10 Functional capacity

2.3.10.1 Hip function

The German form of the „Western Ontario and McMaster Universities Osteoarthritis Index" (WOMAC) was used as a validated assessment of physical function, pain and joint stiffness (Stucki et al. 1996a). The WOMAC has a high sensitivity in capturing improvements in patients with osteoarthritis and subsequent joint replacement (Bellamy et al. 1988; Bellamy et al. 2010).

The questionnaire includes three subscales with five questions concerning pain intensity, two questions about stiffness in the joint and 17 questions about limitations in everyday activities. The sum of all questions in the WOMAC can be a maximum of 240 points, where higher total scores represent a lower physical function due to higher levels of pain, stiffness and inactivity. The scores for the pain subscale ranged from 0 to 50 points, for the stiffness subscale 0-20 points and for the everyday activities 0-170 points. The scores are, furthermore, normalized to allow comparisons to be made for research purposes. This is done by multiplying the total scores of all three subscales by 100 and then dividing the obtained value through the maximum point on the scale. For example, if a patient gets 20

points from the 50 points of the pain subscale, normalization is done by calculating 20x 100 = 2000 and then 2000/ 50 = 40 (Stucki et al. 1996a).

2.3.10.2 Mobility assessment

To measure patients' level of mobility, the Timed up and Go Test was applied (Podsiadlo and Richardson 1991). The Timed Up and Go Test is a well-analyzed psychometric examination in which one times the act of getting up from a seated position, walking three meters, turning around, and going back to the original position. Patients were allocated to five degrees of mobility according to the time needed to complete the Timed Up and Go Test. Level 1 described patients who were "independently mobile" and took less than 10 seconds to fulfill the task. Level 2 included patients who took less than 20 seconds and were termed "mostly independent mobility". Level 3 "variable mobility" comprised of patients who needed 20 - 29 seconds for task completion. Level 4 termed "impaired mobility" included patients who spent longer than 30 seconds for the task and, finally, Level 5 where patients were unable to walk or to finish the test.

2.3.11 Health-related quality of life

To record the quality of life based on health, the Short-Form of the Health Survey Questionnaire (SF-12) translated in German was used (Jenkinson et al. 1997; Wirtz et al. 2018; Drixler et al. 2020). The SF-12 has demonstrated high reliability and validity also for people having chronic pain, so that the Health Survey Questionnaire has established itself as the standard instrument for measuring subjective life quality (Jenkinson et al. 1997). The SF-12 comprises of two essential dimensions of quality of life: the physiological condition and the psychological condition. In order to evaluate the questions, nominal answers are first given a value and the questions are then weighted accordingly. The weighting is based on a mental component summary or a physical component summary. Finally, the weighted questions are added to represent each dimension accordingly. With increasing value, the subjective life-quality increases and vice versa, the sums are normalized in relation to a German sample (Gandek et al. 1998; Drixler et al. 2020).

2.3.12 Psychological characteristics

2.3.12.1 Psychological distress

To assess psychological distress, the "Depression, Anxiety, Stress Scales" (DASS) was applied. The DASS was developed in 1995 (Lovibond and Lovibond 1995). The DASS is a validated test with high internal consistency and constitutes of three scales that are each made up of seven domains. In this analysis, the translated German form was utilized, which has been validated with a German sample of pain patients (Nilges and Essau 2015). Each domain can be assessed on a 4-point scale where greater psychological discomfort is represented by higher scores. Each domain has a different cut off value that is considered to be of clinical relevance (depression >10, anxiety >6 and stress >10).

2.3.12.2 Movement-related fear ("kinesiophobia")

The translated into German "Tampa Scale for Kinesiophobia" (TSK) was utilized to measure the fear of movement, which has been proven to be a crucial indicator for the persistence of impairment and pain (Roelofs et al. 2004). The seventeen questions that are related with mobility and pain can be rated from 1 to 4 points (where 1 = not at all and 4 = totally). For certain questions (4, 8, 12, 16) the amount had to be reversed. Finally, the sum of all scores was calculated. The sum of >37 is the cut off for a clinically relevant fear of movement (Houben et al. 2005; Haugen et al. 2008).

2.3.12.3 Somatization

To assess somatization among the patients, the 15-items scale Patient Health Questionnaire for Physical Symptoms (PHQ-15) derived from the full Patient Health-Questionnaire was used. Each item on the PHQ-15 is rated from 0 to 2 with the answers being accordingly "not bothered at all", "bothered a little" and "bothered a lot". According to the sum the fifteen statements, patients are divided into three levels of somatic symptom severity; level "low" (score of 5 to 9), level "medium" (score of 10 to14) and level "high" (score of 15 to 30) (Kroenke et al. 2002). The German form of PHQ-15 has been validated and has shown high reliability for clinical and occupational health care use (Kroenke et al. 2002).

2.3.12.4 Cognitive appraisal of pain

Through the Kiel Pain Inventory, the cognitive evaluation of pain, i.e., the patients' coping mechanisms and pain-related cognitions such as catastrophizing, was measured

(Hasenbring 1994). The questionnaire used was made up of three subscales from the KPI and contained the Catastrophizing Thought Scale (CTS), the Thoughts of Helplessness (THS) scale and the Thought Suppression Scale (TSS) (Hasenbring 1994). Each scale has as a headline, “when I am in pain, I think about…” and the patients must rate a set of statements on a 7-point scale with the answer possibilities being: never, nearly never, rarely, sometimes, often, most of the time and every time. The CTS is made up of five items that focus on the threatening and scary quality of pain (e.g. “What if the pain gets worse?”), the THS constitutes of nine items describing the lack of hope and positivity (e.g. “The pain will not get better” and the TSS contains four statements that demonstrates the attempts to cope with the pain (e.g. “Pull yourself together!”, “Do not make a scene”).

For each scale the average is calculated. The higher the average for each scale the higher is the respective degree of catastrophizing, thoughts of helplessness or suppression of thoughts and emotions.

2.4 Statistical Analysis

For the statistics and the descriptive assessments, the program SPSS (IBM SPSS Statistics for Windows, version 24.0; IBM Corp, Armonk, NY) was used. Significance level was set at $p < 0.05$ and all p - values recorded are two-sided. Via mean and standard deviation (mean (sdv)), continuous variables were defined while categorical variables were described as frequencies or medians and quartiles (median (1^{st} quartile ; 3^{rd} quartile)). Through the use of visual analysis of histograms and of the Kolmogorov-Smirnov-Test, the distribution of data was checked for normality.

Correlation analyses were done to evaluate the connections between baseline features, preoperative expectations and their fulfillments, improvement in symptoms and the global effectiveness of THA according to the patient (Table 6 and Table 7). In the analysis the changes in symptoms for hip status and overall patient well-being after THA were included. These comprised of changes in hip pain over the previous three months (NRS), hip function and mobility (WOMAC), severity of chronic pain (CPG), quality of life based on physical and mental health (SF-12), psychological distress (DASS) and fear of movement (TSK). The aforementioned changes in symptoms were obtained from the quotient of assessment scores twelve months after surgery and the baseline values. Bonferroni Correction compensated for multiple testing. Cohen's Kappa (κ) was applied to establish the matching of identical categorizations (such as the expectations before surgery and change after surgery). Kendall Tau (τ) non-parametric rank test was used to obtain all correlations of ordinal data because it was defined as more stable and marginally more effective than the Spearman rank-order correlation coefficient (Bonett and Wright 2000; Croux and Dehon 2010). While for continuous data, such as age, Pearson correlation was used, the phi correlation coefficient of the Pearson's Chi-squared test was applied for dichotomous variables (such as gender).

To determine the relative significance of preoperative expectations, fulfillment of expectations and symptom changes scores in predicting the dependent variable patient-rated global effectiveness of surgery a sequential multiple regression analysis was chosen (Supplementary Table 1). Only variables that correlated significantly with the patient-rated global effectiveness of surgery, after Bonferroni-correction for multiple testing (Table 6 and Table 7), were stepwise added to the model. The additive gain in explained variance (R^2) is shown (Supplementary Table 1).

First, baseline characteristics were added to the sequential multiple regression analysis: Degree of school education, walking ability measured with the Timed up and go score, hip pain in the previous three months before THA, chronic pain (CPG) and hip function and mobility (WOMAC) (Supplementary Table 1). Second, preoperative hip pain expectation was entered. Third, change in symptom scores for the following variables were added: hip pain, quality of life based on Physical Component Summary (SF-12) and hip function (WOMAC). Fourth, actuality-expectation assessments for independence, walking ability, general function, physical exercise and social interactions were included.

Unfortunately, because of outstanding data for the predictor variables, with an increasing number of variables entered into the analysis, the sample size declined (Supplementary Table 1) which may lead to reduced statistical power and potentially biased results. The sequential multiple regression analysis was therefore replicated after performing multiple imputation of missing scores on patients having < 3 missing scores out of fourteen factors (N = 85) (Table 8). After analyzing the missing data, the missing values were determined to be randomly distributed (Supplementary Figure 1). For the imputation of ordinal and categorical variables, logistic regression models were applied while linear regression models of the SPSS package (automatic imputation algorithm of SPSS; IBM SPSS Statistics for Windows, version 24.0; IBM Corp, Armonk, NY) dependent on the MAR assumption (missing at random) was used for continuous variables.

The degree of explained variance (R^2) over the ten imputed value sets and the significant predictors' pooled beta regression coefficient for each step of the regression are shown in Table 8. The missing data analysis demonstrated that one value was lacking by 93% of the predictors entered at the final step. Furthermore, 40% (N = 36) of patients had a minimum of one outstanding value, which showed 6% of the 1260 scores missing (factors x variables) and therefore had to be imputed. Following White et al.'s (2010) suggestion that the number of imputations should be higher than the percentage of outstanding data (in this analysis approximately 6%), ten imputations were used. Nevertheless, according to Cohen the results showed a relatively high goodness-of-fit as they were replicated even after imputation with the ten value sets (1977).

The author of this thesis would like to underline that she used data that was collected from a previous project. The author conceived the presented idea, grouped the data, carried out the statistical work and developed the theoretical formalism.

3 Results

3.1 Baseline data of patient collective

During the said timeframe 172 patients admitted for unilateral total hip arthroplasty due to osteoarthritis were screened for the analysis. 47 patients did not fulfill the requirements for inclusion (one patient did not speak German, one was younger than 18 years, three patients were already taking part in other studies, one patient had dementia, one patient had an active drug abuse, three patients were recruited outside of the clinical setting and could thus not fill out the preoperative questionnaires, nine patients had spinal anesthesia, the surgery was postponed by 15 patients and 13 patients refused to participate in the analysis).

At the 12-month follow up, 95 patients were successfully contacted. Of these, 90 patients completed the whole survey and were evaluated in the final model (N = 49 female, N = 41 male).

All the patients in the data analysis underwent total hip replacement surgery, with 46 having the surgery on the right hip and 44 on the left hip. 49% of patients were retired, with the mean age being 63 $\pm$ 12.9 years and the BMI 27.6 $\pm$ 4.7. Most patients lived with their families and had finished School while 18% lived alone and just 1% had no school education. Primary osteoarthritis of the hip was the most common sign for surgery with 77%, 15% underwent surgery because of hip dysplasia, 5% because of necrosis of the femoral head and 3% due to posttraumatic osteoarthritis. Further sociodemographic and clinical information of included patients can be found in Table 1.

Table 1. Sociodemographic variables at baseline

Gender, No.	49 Women; 41 Men
Age at examination, years, mean ± SD	63 ± 12.9
Body Mass Index (BMI), kg/m^2, mean ± SD	27.6 ± 4.7
Degree of education, No. (%)	
No graduation	1 (1%)
Haupt-/Volkschule[a]	36 (40%)
Realschule/Mittlere Reife[b]	27 (30%)
Fachhochschule, Abitur, allg. Hochschule[c]	26 (29%)
Living situation, No. (%)	
With family members	72 (80%)
Alone	18 (20%)
Employment, No. (%)	
Employed	30 (33%)
Unemployed	11 (12%)
Retired	49 (55%)
Operated hip, No. (%)	
Right	46 (51%)
Left	44 (49%)
Surgery indication, No. (%)	
Primary Coxarthrosis	69 (77%)
Hip Dysplasia	14 (15%)
Necrosis	4 (5%)
Posttraumatic	3 (3%)

Descriptive statistics are based on N = 76 - 90 subjects due to varying numbers of missing data per variable; [a]*Hauptschule* in Germany refers to the final examination at grade 9; [b]*Realschule* finishes after grade 10 with the degree *mittlere Reife*; [c]*Gymnasium* finishes with the final examination called *Abitur* after grade 13. Table from Kästner et al. (2021).

3.2 Patients' reasons to undergo surgery

The three most crucial reasons for undergoing THA were: to improve walking abilities (70% of patients chose this answer), minimizing the pain (59%) and to become more independent (52%).

3.3 Patients' most important factors for surgical success

Hip pain was the most determining factor that would need to be changed in order to regard the surgery as successful for 47% of patients. While 30% found walking ability and 14% found general function to be the most significant change that had to happen. For independence in everyday life, back pain and physical exercise a total of only 15% viewed these factors to be important in determining surgery success.

3.4 Patients' most important postoperative change

After surgery the most significant change to have had occurred was for 59% hip pain, for 24% walking ability and for 7% general function (6% for independence in everyday life, 3% for back pain and 1% for physical exercise). The patients' answers to the question what was the most important change that had occurred after the surgery at the twelve months follow-up correlated slightly ($p = 0.006$, $\kappa = 0.188$) with the answers to the preoperative question what change would make the surgery a success (3.3) (Landis and Koch 1977).

3.5 Patient-rated global effectiveness of surgery

Nearly all patients (98%) reported that the surgery had helped them overall. While no patient declared a worsening in their condition after surgery (Figure 1).

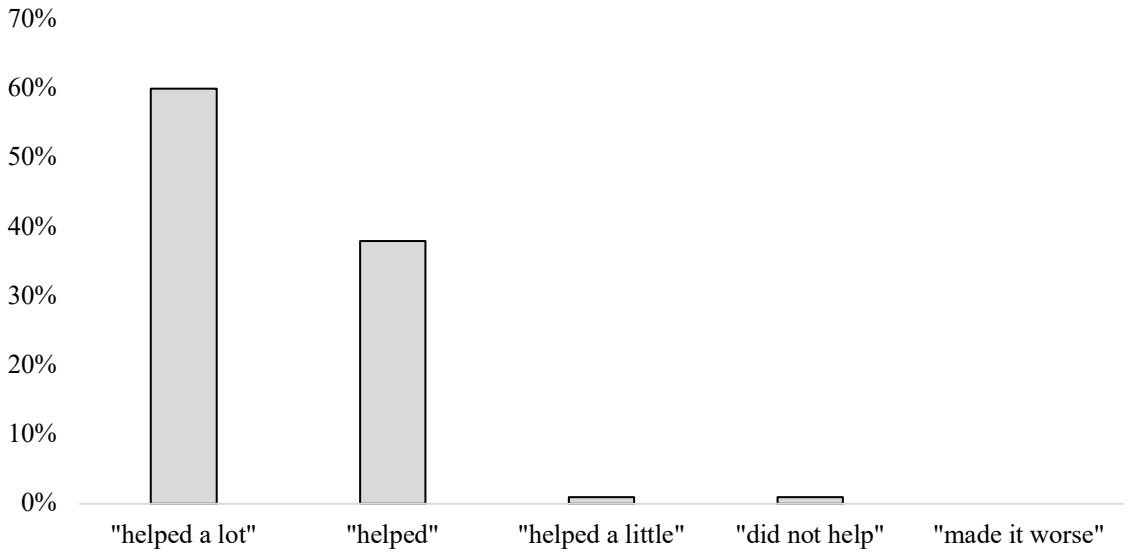


Figure 1. Patient-rated global effectiveness of surgery (N = 90) twelve months after THA

3.6 Expectations

Patients were found to have the highest expectations for progress in hip pain, walking ability and independence, which was consistent with the patient-rated main reasons for undergoing THA. Patients had overall rather high expectations for improvement in all 8 domains. In particular, patients with worse preoperative hip function and lower life quality showed the highest expectations for postoperative general function.

Asked about the actual outcome twelve months after surgery, 98% of participants reported progress (somewhat better, better or much better) in hip pain, walking ability and independence in everyday life. 92% showed an improvement in general function. The largest discrepancies between preoperative expectations and perceived postoperative outcome were seen in the domains of back pain and physical exercise. Only 62% of patients reported a benefit in social interactions. The expectations before THA and corresponding outcomes twelve months following THA are demonstrated in Table 2.

Table 2. Comparison of patients' expectations at baseline and respective outcomes twelve months after total hip arthroplasty (THA)

Domains	Percentage of patients in each category at baseline (pre) and twelve months (12 m) after THA											
	Much better		Better		Somewhat better		Unchanged		Worse		Uncertain	
	pre	12 m	pre	12 m	pre	12 m	pre	12 m	pre	12 m	pre	12 m
Hip pain	79	66	21	30	-	2	0	2	-	-	-	-
Back pain	37	6	34	30	7	30	9	33	-	1	13	-
Walking ability	83	48	15	36	1	14	0	1	-	1	1	-
Independence	67	33	25	46	2	19	2	2	-	-	4	-
Physical exercise	44	15	35	28	9	27	7	30	-	-	5	-
General function	59	21	37	42	1	29	3	8	-	-	-	-
Social interactions	32	2	30	20	5	40	27	38	-	-	6	-
Mental well-being	39	5	36	31	2	40	21	23	-	1	2	-

Descriptive statistics are based on N = 79 - 90 subjects due to varying numbers of missing data per variable. Table from Kästner et al. (2021).

3.7 Fulfillment of expectations

To note, when questioned at the twelve-month follow-up if their expectations before surgery were fulfilled, patients had a more optimistic view on the degree of expectation fulfillment in relation to the actuality-expectation discrepancy scores (compare Figure 2 and Figure 3). Only 3% stated that their expectations had not been met for hip pain, walking ability and independence in everyday life. For the domains of back pain, social interactions, physical exercise and mental well-being patients reported lower levels of expectation fulfillment (Figure 2). Thus, more hip specific outcomes, that are naturally most likely to be improved by the hip arthroplasty are almost completely fulfilled, while less clearly related outcomes fair more poorly.

The calculated expectation fulfillment shown in Figure 3 demonstrates likewise that apart from the two domains hip pain and walking ability, 52 to 65% of patients did not meet their expected outcome in all other domains. Thus, mirroring the self-reported expectation fulfillment, the actuality-expectation discrepancy scores showed that only for the hip-specific domains (hip pain and walking ability) most patients reached their expected improvement. Figure 4 illustrates the association between both methods of assessing expectation fulfillment. A high correspondence between the two aforementioned methods for all eight domains was shown (for all domains $p \leq 0.001$, $\tau_B = 0.355$ - 0.538).

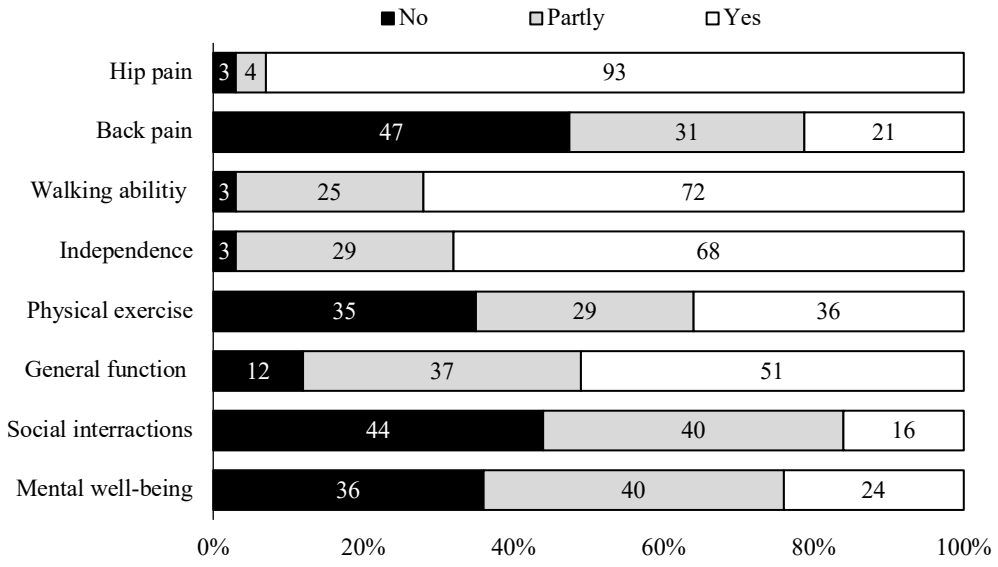


Figure 2. Patients' self-reports of expectation fulfillment twelve months after surgery. Sample size varies N = 88 - 90 due to missing data. The graph shows in black the percentages of patients that

responded “No”, in gray that responded “Partly” and in white that responded “Yes” to the question “Have your expectations regarding this domain been fulfilled?” during the twelve-month follow-up.

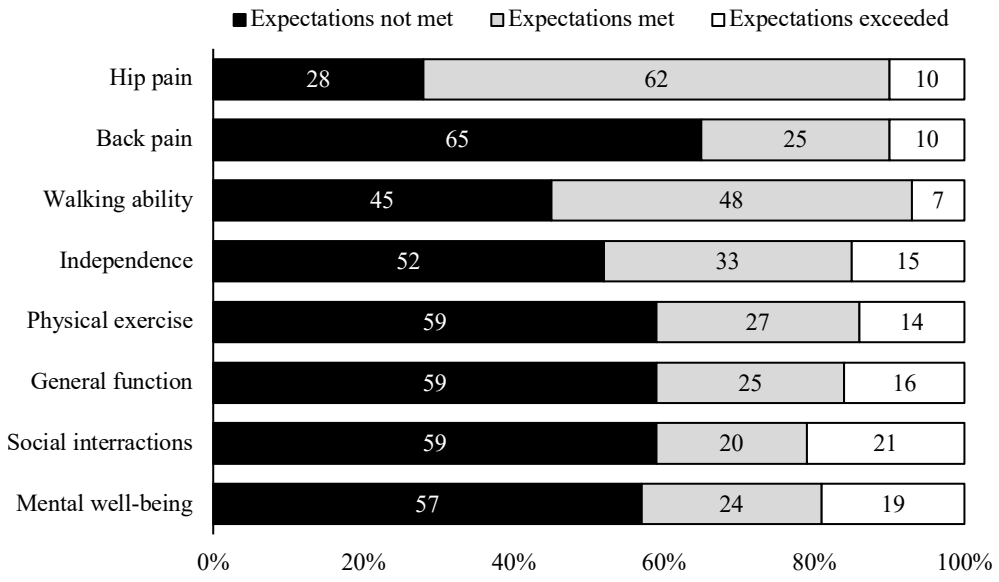


Figure 3. Calculated expectation fulfillment. Sample size varies N = 79 - 90 due to missing data. The actuality-expectation discrepancy was obtained by subtracting the expectation value from the actual value at twelve months. The obtained scores were then divided into three groups. The “unmet expectations” group, shown here in black, demonstrates the percentage of patients, who had a worse outcome than expected. The “expectations met” group shown here in gray shows the percentage of patients, who had an outcome as expected. Lastly, the “expectations exceeded” group shown in white represents the percentage of patients, whose outcome was better than expected. Figure from Kästner et al. (2021).

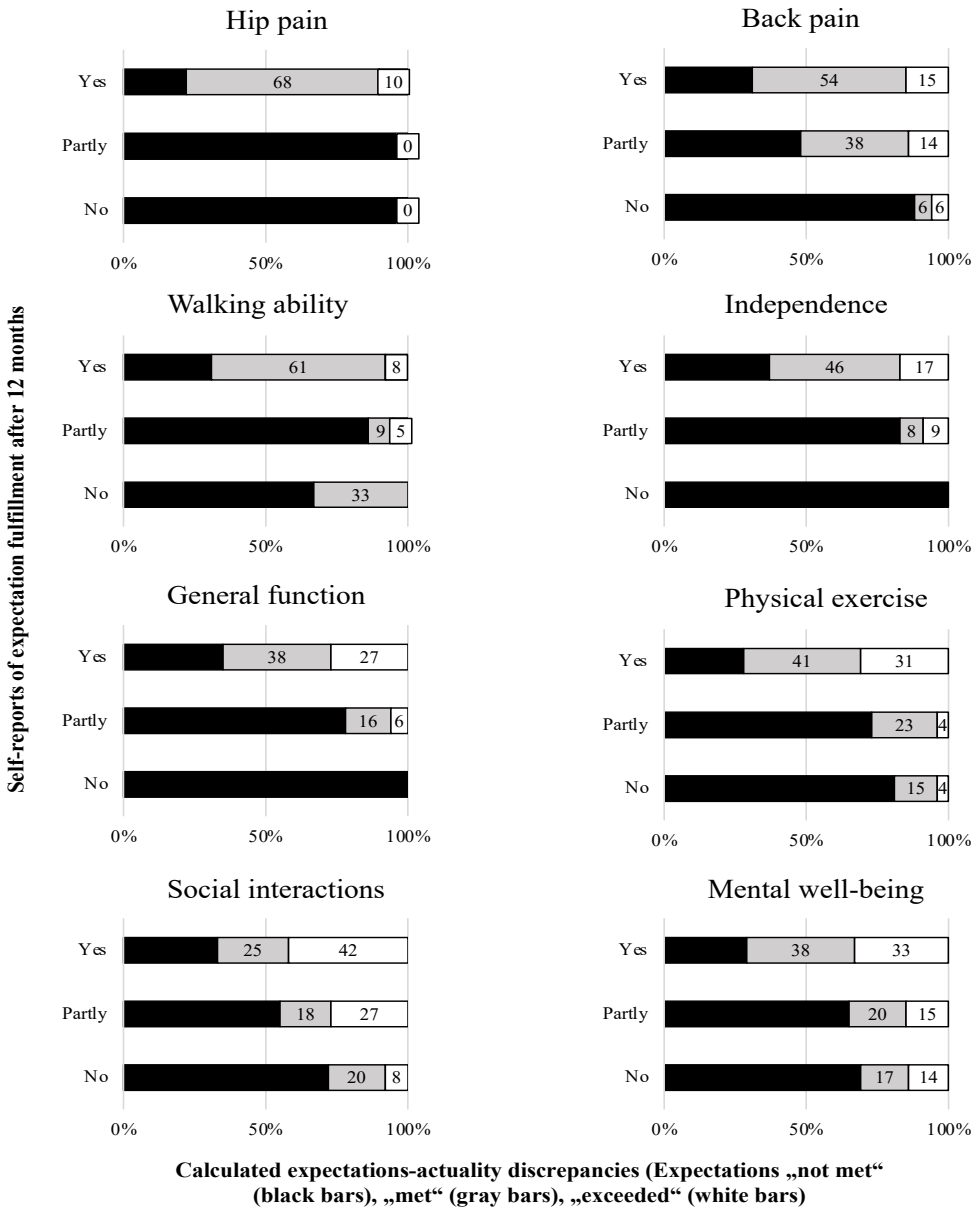


Figure 4. Comparison of the self-reports of expectation fulfillment and the calculated actuality-expectation discrepancies. Sample size varies N = 79 - 90 due to missing data. The calculated measure is the difference between the expectation and the reported change twelve months after surgery. For the self-reported expectation fulfillment, patients were asked “Were your expectations for this domain fulfilled?” with the answer possibilities being “yes”, “partly” and “no” (y axis). The x-axis shows the percentages of patients in each self-reported group (“yes”, “partly” and “no”). The calculated actuality-expectation discrepancy shows that expectations not met in black (result was worse than expected), met expectations (result was as expected) in gray and exceeded expectations (result was better than expected) in white.

3.7.1 Association between preoperative expectations and postoperative outcome

Apart from the expectations concerning the domain social interactions, preoperative expectations did not correlate with the respective changes reported twelve months after THA as shown in Table 3. Where the preoperative expectations ranging on a five-point scale from “helped a lot” to “made it worse” and the corresponding postoperative improvement also ranging on a scale from “helped a lot” to “made it worse”. Scores for general function and mental well-being at least showed a trend.

Table 3. Association between preoperatively expected and reported outcome twelve months after THA

Expectations	Corresponding outcome at twelve-month follow-up Correlation coefficient (p - value)
Hip pain	$T_B = 0.162$ ($p = 0.127$)
Back pain	$T_B = 0.021$ ($p = 0.826$)
Walking ability	$T_B = 0.112$ ($p = 0.268$)
Independence in everyday life	$T_B = 0.096$ ($p = 0.327$)
Physical exercise	$T_B = 0.101$ ($p = 0.286$)
General function	$T_B = -0.187$ ($p = 0.056$)
Social interactions	$T_B = -0.250$ ($p = 0.008$)
Mental well-being	$T_B = -0.179$ ($p = 0.058$)

Correlation was obtained through Kendal Tau. Nominally significant p - values are set in italics., significance level p = 0.05, Sample size varies N = 79 - 90 due to missing data. Significant data is shown in italic.

3.8 Pain characteristics

3.8.1 Baseline

On average, patients reported moderate hip pain in the three months prior to the THA 6 (4 ; 7). Only about 20% of the participants had a high degree of pain chronicity (MPSS) while most patients showed low to medium pain chronicity (see Table 4).

Nearly a quarter of the patients suffered from hip pain for more than 5 years, while a third of them reported having had hip pain for the last 2-5 years. Half of the patients reported having had hip pain for two years or less.

Only 17% of patients were found to have low levels of pain and no impairment in everyday life, respectively grade 1 CPG. About 60% of patients had a high degree of impairment (CPG 3 and 4) and 23% high levels pain with only low to moderate impairment. (see Table 4).

3.8.2 Twelve-months follow-up

In contrast at twelve months after surgery, patients reported low to no pain over the last three months 0 (0 ; 2.5). Moreover, only 15% of the patients reported being highly impaired due to chronic pain (CPG 3 and 4), while the majority of patients reported having low impairment and low pain intensities.

3.9 Functional capacity

3.9.1 Hip function

3.9.1.1 Baseline

Patients had an average total WOMAC score of 53.1 (SD ± 20.8) points, showing moderate to high degrees of low physical functioning. The pain subscale had a mean of 50.5 (SD ± 22.3), the stiffness subscale a mean of 54.1 (SD ± 24.9) and the subscale of inactivity a mean of 54.0 (SD ± 21.6) (see Table 4).

3.9.1.2 Twelve-months follow-up

Twelve months after surgery, patients overall showed improvement in hip function and mobility with a mean decreased to 4.7 (SD ± 8.4) for the total WOMAC score.

3.9.2 Mobility assessment

All patients were able to complete the test, with only 15% of patients showing decreased to impaired mobility.

Table 4. Pain-related and functional variables at baseline

Pain characteristics		
Average hip pain in last 3 months before surgery, NRS (0-10), median (1st; 3rd quartile)		6 (4 ; 7)
Pain chronicity, MPSS[a], No. (%)		
	Stage I (low)	36 (40%)
	Stage II (medium)	37 (41%)
	Stage III (high)	17 (19%)
Overall severity of chronic pain, Chronic Pain Grade (von Korff), No. (%)		
	Grade 1	15 (17)
	Grade 2	20 (23)
	Grade 3	17 (18)
	Grade 4	38 (42)
Duration of hip pain, time intervals, No. (%)		
	1 to 12 months	18 (20.1%)
	12 to 24 months	22 (24.4%)
	2 to 5 years	31 (34.4%)
	More than 5 years	19 (21.1%)
Pressure pain threshold (PPT), kPa, mean ± SD		391.6 ± 179.4
Functional capacity		
Hip function and mobility, WOMAC[b] score, mean ± SD		53.1 ± 20.8
Walking ability, Timed up and go test score, No. (%)		
	Level 1	21 (24%)
	Level 2	54 (61%)
	Level 3	11 (13%)
	Level 4	2 (2%)
	Level 5	0

Descriptive statistics are based on N = 76 - 90 subjects due to varying numbers of missing data per variable; [a]MPSS = Mainz Pain Staging System (Gerbershagen et al., 2008); [b]WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index (Stucki et al., 1996a). Table from Kästner et al. (2021).

3.10 Quality of life based on health

3.10.1 Baseline

Our patient collective showed overall low physical health with almost normal mental health with a mean of 29.9 ± 7.9 for physical component summary of SF-12 and a mean of 49.2 ± 12.4 (Median 51.2 (40.7 ; 59.5)) for mental component summary of SF-12 see Table 5.

3.10.2 Twelve-months follow-up

Twelve months after THA patients showed better quality of life than before the surgery. The physical component summary of SF-12 mean score increased by 18 points to 47.8 ± 9, while the mental component summary of SF-12 increased by 5.7 points to 54.9 ± 6.8.

3.11 Psychological characteristics

3.11.1 Baseline

Patients showed rather low scores on the depression, anxiety and stress scales (Table 5). With only about 10% having scores over the cut-offs of clinical relevance. Specifically, only 7% of patients had a clinically relevant depression score, 8% a clinically relevant anxiety score and only 10% a clinically relevant stress score.

Only a third of patients showed a clinically relevant fear of movement (Table 5).

The majority of patients showed low to no relevant somatization, 32% demonstrated mild levels and only on subject had a high level of somatization (Table 5). Overall, patients demonstrated low levels of catastrophizing, thoughts of helplessness and thought-suppression mechanisms (Table 5).

3.11.2 Twelve-months follow-up

Although the preoperative levels of depression, anxiety and stress were already low in the present patient population, twelve months after the intervention there was not only an improvement in all three scales but also a decrease in the number of clinically relevant depression, anxiety and stress levels. The depression and anxiety scale both demonstrated a median of 0 (0 ; 1) while the stress scale showed a median of 2 (0 ; 5). Here, 3% patients

were above the cut-off for depression, 2% were higher than the cut-off on the anxiety scale and 6% patients had a score on the stress scale that was clinically relevant.

Concerning kinesiophobia, an improvement was assessed twelve months after surgery where only 20% of patients reported having a clinically relevant fear of movement. The TSK median at twelve months after surgery decreased to 30 points (27 ; 35.3) where higher scores represent higher fear of movement and 37 points being the cut-off for clinical relevance.

Table 5. Quality of life and psychological characteristics at baseline

Quality of life		
Health-related quality of life, SF-12[a], mean ± SD		
	SF-12 Physical Component Summary	29.9 ± 7.9
	SF-12 Mental Component Summary	49.2 ± 12.4
Psychological characteristics		
Psychological distress, DASS[b], median (1st; 3rd quartile)		
	DASS Depression	3 (1 ; 5)
	DASS Anxiety	1 (0 ; 3)
	DASS Stress	5 (2 ; 8)
Kinesiophobia, TSK[c], median (1st; 3rd quartile)		36 (31 ; 41)
Somatization, PHQ-15[d], median (1st; 3rd quartile)		5 (4 ; 8)
Cognitive appraisal of pain, KPI[e], median (1st; 3rd quartile)		
	Catastrophizing thought scale	2 (0.5 ; 3.1)
	Helplessness scale	0.4 (0 ; 1.2)
	Thought suppression scale	2.8 (1.3 ; 3.8)

Descriptive statistics are based on N = 76 - 90 subjects due to varying numbers of missing data per variable; [a]SF-12 = short form of the Health Survey Questionnaire (Jenkinson et al., 1997); [b]DASS = Depression, Anxiety, Stress Scales (Nilges and Essau, 2015); [c]TSK = Tampa Scale for Kinesiophobia (Roelofs et al. 2004); [d]PHQ-15 = Patient Health Questionnaire (Kroenke et al. 2002); [e]KPI = Kiel Pain Inventory (Hasenbring, 1994). Table from Kästner et al. (2021).

3.12 Predictors of the patient-rated global effectiveness of the surgery

To determine the predictors of the patient-rated global effectiveness of the surgery preoperative risk factors for bad outcome after THA age, gender, preoperative pain intensity, BMI, educational level, pain characteristics, functional capacity, Quality of life and psychological characteristics were first correlated with the outcome patient-rated

global effectiveness of the surgery (Table 6). Here the educational level, the average preoperative pain, the overall severity of chronic pain assessed with the CPG von Korff, walking ability and hip function (WOMAC) were found to be significantly associated (Table 6). Next, preoperative expectations, changes in symptoms and the calculated actuality-expectation discrepancy scores were correlated using Kendall Tau with the outcome (Table 7). Interestingly, hip pain which was the second most determining factor in deciding to undergo THA, was the only expectation that significantly correlated with the patient-rated global effectiveness of surgery (Table 7). The change in symptoms for average hip pain in the last 3 months, hip function and mobility (WOMAC) and the physical component summary from SF-12 were also found to significantly correlate with the outcome (Table 7). Finally, from the calculated actuality-expectation discrepancy scores walking ability, independence, physical exercise, general function and social interactions were all found to have a significant correlation with the patient-rated global effectiveness of THA.

In Table 8 the findings of the sequential multiple regression analysis following multiple imputation are presented. In the first step baseline variables that significantly correlated with the patient-rated global effectiveness of THA were included, failing to identify a significant predictor (Table 6). Next, hip pain expectation, which was the only preoperative expectation which had a significant correlation with the dependent variable (Table 7) was added. Here, hip pain expectation and walking ability before surgery (timed up and go test) predicted the dependent variable. At step 3 changes in symptoms (the quotient of outcome measures at twelve months and the respective preoperative values) that significantly correlated with the outcome (Table 7) were entered. No significant contribution was observed at this step. In the final step, calculated actuality-expectation scores that significantly correlated with patient-related effectiveness (walking ability, independence, physical exercise, general function and social interactions (Table 7)) were entered. The largest shift in the described variance in the model was brought on by the addition of the calculated actuality-expectation scores. Here, the final predictors of patient-rated global effectiveness of THA were established: hip pain expectation, walking ability before surgery, calculated expectation fulfillment in walking ability, independence in everyday life and general function. The pooled beta was also included in the analysis with Hip pain expectation having the largest impact on the outcome; for every 1 step change in the pain expectation assessment one can expect an increase in the global effectiveness of surgery of 4.605. (Table 8). Similarly, taking walking ability as an

example for every 1 step change in the walking ability one can expect a decrease of 3.103 in the outcome (Table 8). The final model consisting of the five aforementioned significant predictors accounted for 52% of the variance (R^2) in the patient-rated global effectiveness of surgery (Table 8).

Table 6. Correlations[a] between baseline variables and patient-rated global effectiveness of THA

Baseline variables	Global effectiveness of THA
Sociodemographic variables and Body Mass Index	**Correlation coefficient (p - value)[b]**
Gender	Phi = 0.255 (p = 0.118)
Age at examination, years	Pearson's r = -0.47 (p = 0.658)
Degree of school education	Kendall Tau_B (T_B) = 0.301 (**p = 0.002**)
Body Mass Index, BMI, kg/m^2	Pearson's r = -0.247 (p = 0.019)
Pain characteristics	**Correlation coefficient (p - value)[c]**
Average hip pain in the last 3 months before surgery, Numeric Rating Scale (0-10)	T_B = -0.330 (**p < 0.001**)
Overall severity of chronic pain, CPG (von Korff)	T_B = -0.375 (**p < 0.001**)
Pain chronicity, MPSS[d]	T_B = -0.229 (p = 0.021)
Duration of hip pain	T_B = 0.109 (p = 0.256)
Pressure pain threshold (PPT)	R = 0.190 (p = 0.075)
Functional capacity	**Kendall Tau_B (p - value)[e]**
Walking ability, Timed up and go test score	T_B = -0.313 (**p = 0.002**)
Hip function and mobility, WOMAC[f]	T_B = -0.311 (**p = 0.001**)
Quality of life and psychological characteristics	**Kendall Tau_B (p - value)[g]**
Health-related quality of life, SF-12[h]	
SF-12 Physical Component Summary	T_B = -0.020 (p = 0.825)
SF-12 Mental Component Summary	T_B = 0.052 (p = 0.562)
Psychological distress, DASS[i]	
DASS Depression	T_B = -0.134 (p = 0.143)
DASS Anxiety	T_B = -0.085 (p = 0.373)
DASS Stress	T_B = -0.128 (p = 0.157)
Somatization, PHQ-15[j]	T_B = -0.011 (p = 0.911)
Kinesiophobia, TSK[k]	T_B = -0.074 (p = 0.433)
Cognitive appraisal of pain, KPI[l]	
Catastrophizing thought scale	T_B = -0.051 (p = 0.602)
Helplessness scale	T_B = -0.218 (p = 0.019)
Thought suppression scale	T_B = -0.230 (p = 0.014)

Multiple testing adjusted significances (applying Bonferroni to each variable cluster) are set in boldface. Nominally significant p - values are set in italics. Variables significantly associated with global effectiveness of THA were included as step 1 in the sequential multiple regression analysis (Table 3). [a]Correlation analyses are based on N = 76 - 90 subjects due to varying numbers of missing data per variable; [b]Bonferroni-adjusted significance level: p = 0.05/4; [c]Bonferroni-adjusted significance level: p = 0.05/3 [d]MPSS = Mainz Pain Staging System (Gerbershagen et al., 2008); [e]Bonferroni-adjusted significance level: p = 0.05/2; [f]WOMAC=Western Ontario and McMaster Universities Osteoarthritis Index (Stucki et al., 1996a); [g]Bonferroni-adjusted significance level: p = 0.05/12; [h]SF-12 = short form of the Health Survey Questionnaire (Jenkinson et al., 1997); [i]DASS = Depression, Anxiety, Stress Scales (Nilges and Essau, 2015); [j]PHQ-15 = Patient Health Questionnaire (Kroenke et al. 2002); [k]TSK = Tampa Scale for Kinesiophobia (Roelofs et al. 2004); [l]KPI = Kiel Pain Inventory (Hasenbring, 1994). Table modified from Kästner et al. (2021).

Table 7. Correlations between preoperative expectations, change in symptoms and calculated actuality-expectation discrepancy scores with the patient-rated global effectiveness of THA

Preoperative expectations[a]	**Global effectiveness of THA Kendall Tau_B (p - value)**[b]
Hip pain	T_B = 0.329 (**p = 0.002**)
Back pain	T_B = -0.101 (p = 0.325)
Walking ability	T_B = 0.163 (p = 0.122)
Independence	T_B = 0.064 (p = 0.533)
Physical exercise	T_B = 0.095 (p = 0.355)
General function	T_B = -0.045 (p = 0.664)
Social interactions	T_B = -0.196 (p = 0.047)
Mental well-being	T_B = -0.006 (p = 0.949)
Change in symptoms[c]	**Global effectiveness of THA Kendall Tau_B (p - value)**[d]
Average hip pain in the last 3 months	T_B = -0.322 (**p = 0.002**)
Overall severity of chronic pain condition, CPG (von Korff)	T_B = -0.146 (p = 0.122)
Hip function and mobility, WOMAC	T_B = -0.295 (**p = 0.002**)
Health-related quality of life, SF-12[h]	
SF-12 Physical Component Summary	T_B = 0.304 (**p = 0.001**)
SF-12 Mental Component Summary	T_B = 0.011 (p = 0.899)
Psychological distress, DASS[i]	
DASS Depression	T_B = -0.211 (p = 0.049)
DASS Anxiety	T_B = 0.064 (p = 0.600)
DASS Stress	T_B = -0.053 (p = 0.586)
Kinesiophobia, TSK[k]	T_B = -0.243 (p = 0.011)
Calculated actuality-expectation discrepancy scores[e]	**Global effectiveness of THA Kendall Tau_B (p - value)**[b]
Hip pain	T_B = 0.157 (p = 0.126)
Back pain	T_B =0.101 (p = 0.346)
Walking ability	T_B = 0.473 (**p < 0.0001**)
Independence	T_B = 0.471 (**p < 0.0001**)
Physical exercise	T_B =0.340 (**p = 0.001**)
General function	T_B =0.314 (**p = 0.001**)
Social interactions	T_B = 0.317(**p = 0.001**)
Mental well-being	T_B = 0.215 (p = 0.026)

Multiple testing adjusted significances (applying Bonferroni to each variable set) are set in boldface. Nominally significant p - values are set in italics. Variables significantly associated with global effectiveness of THA were included as steps 2 to 4 in the sequential multiple regression analysis (Table 3). [a]Correlation analyses are based on N = 79 - 90 subjects due to varying numbers of missing data per variable; [b]Bonferroni-adjusted significance level: p = 0.05/8; [c]Correlation analyses are based on N = 54 - 90 subjects due to varying numbers of missing data per variable; change scores were obtained by dividing the outcome measures at 12-month follow-up by the preoperative scores; [d]Bonferroni-adjusted significance level: p = 0.05/9; [e]Correlation analyses are based on N = 69 - 90 subjects due to varying numbers of missing data per variable. Table modified from Kästner et al. (2021).

Table 8. Results of the sequential multiple regression analysis (N = 85a): Variance explained in the global effectiveness of total hip arthroplasty at twelve-months follow-up by sociodemographic and medical variables, preoperative expectations, change in symptoms and the fulfillment of expectations (multiple imputation of missing values)

Steps	Variables included upon each step	Range of R^2 across the 10 imputed data sets	Pooled ß (p - value)	Only significant predictor variables
Step 1	Degree of school education, average hip pain in the last 3 months before surgery, overall severity of chronic pain (CPG), hip function and mobility (WOMAC), walking ability (Timed up and go score)	0.261 – 0.304	-	-
Step 2	Variables from step 1, **preoperative hip pain expectation**	0.372 – 0.416	-2.143 (p = 0.033)	**Walking ability** (Timed up and go score)
			3.745 (p < 0.001)	**Hip pain expectation**
Step 3	Variables from step 2, **change in symptoms for: Average hip pain in the last 3 months, hip function and mobility (WOMAC), health-related quality of life (SF-12 Physical Component Summary)**	0.419 – 0.493	-2.338 (p = 0.020)	**Walking ability** (timed up and go score)
			3.370 (p = 0.001)	**Hip pain expectation**
Step 4	Variables from step 3, **calculated actuality-expectation discrepancies for walking ability, independence, physical exercise, general function and social interactions**	0.613 – 0.689	-3.103 (p = 0.002)	**Walking ability** (timed up and go score)
			4.605 (p < 0.001)	**Hip pain expectation**
			2.601 (p = 0.009)	Calculated expectation fulfillment for **Walking ability**
			2.952 (p = 0.002)	Calculated expectation fulfillment for **Independence**
			-2.783 (p = 0.006)	Calculated expectation fulfillment for **General function**
Final model	Walking ability (timed up and go score), hip pain expectation, actuality-expectation discrepancy for walking ability, independence and general function	0.510 – 0.544	-	-

[a]Individuals with more than 3 out of 14 missing predictors were excluded from the analysis. The variables marked in bold are added to the model at each respective step. Predictor variables individually significantly associated with global effectiveness of THA (see Table 6 and Table 7) were entered in four steps. The significant predictors in the final model were: Walking ability at baseline, hip pain expectation and the calculated actuality - expectation discrepancy (calculated expectation fulfillment) of walking ability, independence in everyday life and general function. β in final model = β regression coefficient after all listed variables have been entered; R^2_{change} = Increase in explained variance by step; adjusted R^2, = R^2- (k - 1)/ (n - k) * (1 - R^2) where n = no. observations, k = no. independent variables. Level of significance was set to p < 0.05. Table modified from Kästner et al. (2021).

4 Discussion

The objective of the present thesis was to perceive the degree to which preoperative expectations, their fulfillment and the actual progress in symptoms after surgery were related to the outcome of patient-rated global effectiveness of THA. Essentially, the core results of Mannion et al. (2009a) have been reproduced, showing that meeting the anticipated change is what mostly determines patients' opinions of the success of total hip arthroplasty. To note, it was only for the more functional domains such as walking ability, independence in everyday life and general function that expectation fulfilment was the core predictor of the outcome. For the more tangible sensory-affective perception of pain associated to the hip, a positive expectation was shown to have more of an impact on the outcome irrespective of whether the expectation was fulfilled or not. This was made even clearer by the fact that the actuality-expectation discrepancy for hip pain showed no significant correlation to the outcome. The improvements in core symptoms concerning hip status and postoperative patient condition did not predict the outcome, although symptoms improvements were impressive after twelve months. Walking ability, assessed objectively with the timed up and go test, at baseline, however, predicted the outcome variable. Showing that patient mobility before surgery, had a considerable impact on the patients' subjective assessment of surgery effectiveness.

In contrast to previous studies, neither gender nor age significantly correlated with expectation. Similarly, although depression has been shown to be associated with poor satisfaction, the results of this paper did not demonstrate a connection between depression scores and postoperative outcome. The patients in this analysis, however, had very low depression scores of questionable clinical relevance. Indeed, 92% of the patients showed no depressive symptoms and the association between depression and expectation could thus be rendered statistically irrelevant in our analysis.

Patients had high expectations for improvement in all eight domains (hip pain, back pain, walking ability, independence in everyday life, physical exercise, general function, social interactions and mental well-being) after surgery. In accordance with the three main reasons for undergoing the intervention, patients had the highest expectations for hip pain, walking ability and independence. Only expectation for social interactions was found to correlate with the actual change, all other expectations did not correlate with the

corresponding change twelve months after surgery. The self-reported expectation fulfillment was higher in all eight domains than the calculated expectation fulfillment. The self-reported expectation fulfillment for the hip specific outcomes (hip pain, walking ability and independence), that are most likely improved by the hip arthroplasty were almost completely fulfilled, while less clearly related outcomes faired more poorly.

Twelve months after the total hip arthroplasty, patients reported better mental and physical health, better function and mobility and less to no pain. It will, therefore, come to no surprise that patients in this thesis claimed high levels of global effectiveness of surgery with only two patients reporting some to no improvement after the intervention.

4.1 Predictors of patient-rated global effectiveness of surgery

In this thesis, walking ability before surgery, hip pain expectation and the fulfillment of expectation concerning walking ability, independence in everyday life and general function predicted the patient-rated global effectiveness of THA.

A more optimistic hip pain expectation correlated with higher perceived surgery effectiveness according to the patient. Therefore, patients with expectations of low hip pain after surgery showed greater perceived global effectiveness of the THA. This concurs with the expectation-outcome model, where a strong correlation between positive expectations and outcome is postulated (Haanstra et al. 2012). Judge et al. (2011b) findings demonstrated that higher preoperative aspirations correlated with greater function and agility twelve months after intervention. In a study on spinal surgery, conducted by Saban and Penckofer (2007), higher levels of optimism predicted better quality of life after surgery. Mannion et al. (2009b) found that expectations of both postoperative pain and functional recovery correlated with satisfaction after a total knee arthroplasty. In a multiple regression model, however, none of the two achieved significance in the explained variance of the model. In the current analysis, nevertheless, hip pain expectation was the only expectation to predict global treatment effectiveness in the multiple regression model, as well as the only expectation to correlate with the global effectiveness of THA during the univariate analysis.

The achievement of preoperative expectations concerning walking ability, independence in everyday life and general function predicted the patient-perceived effectiveness of THA. This supports the repeatedly shown notion that it is not the level of expectations themselves that determine therapeutic outcome, but rather the fulfillment of patients'

expectations (Kravitz 1996; Pager 2004; Saban and Penckofer 2007; Palazzo et al. 2014). Hamilton et al. also (2013) found patient satisfaction to be dependent on pain relief, satisfaction with the hospital staff, perceived pain, functional ability and on meeting the patient's expectations. Our findings mirror Mannion et al.'s results (2009a), where the degree of fulfillment of preoperative expectations predicted patient-rated global outcome of spinal surgery. Specifically, our findings demonstrate that the more the expectations of walking ability, independence and general function were exceeded, the higher were the patients' ratings of surgery effectiveness. In accordance, one could suggest that manipulating patients' (in this case functional) expectations, i.e. lowering expectations, could result in increasing the probability of the patient finding the surgery a success. This, nonetheless, is neither a realistic nor practical approach in medicine. Lowering patients' expectations would not only hinder patients from undergoing necessary treatment but would also bring about a lack of faith and trust in the health care system (Bandura 1978; Mannion et al. 2009a). Rather, realistic expectations should be conveyed so as to minimize disappointment due to unrealistically optimistic expectations.

Interestingly, the calculated expectation fulfillment of hip and back pain were the only domains of expectation fulfillments that did not correlate with the global effectiveness of THA. To comprehend the impact of these expectations on the result, it is salient to distinguish the nature of an emotional expectation, such as pain, from a functional expectation, such as walking ability, general function and independence in everyday life (Peerdeman et al. 2016). Pain is a multidimensional construct, heavily influenced by response expectancies, which are in turn attributed to the pain experience in itself (Peerdeman et al. 2016). Moreover, pain is also influenced by stimulus-outcome expectancies from external cues and outcome-response expectancies, as well as by self-efficacy expectancies, which pertain to beliefs of one's ability to endure pain (Vancleef and Peters 2011; Jackson et al. 2014). Pain catastrophizing, trust, worry, optimism and neuroticism haven all been closely related to pain (Quartana et al. 2009; Rawdin et al. 2013; Basten-Günther et al. 2019). Functional recovery expectations for example for walking ability pertain to a much broader and abstract notion that applies more closely to outcome-behavior expectancies and self-efficacy expectations (Sullivan et al. 2011). Functional recovery after surgery is attributed to much higher active involvement from the patient than pain experience after surgery (Sullivan et al. 2011). Patients have little direct impact on their pain experience after surgery, while their functional recovery is mainly dependent on their level of motivation, activity and interest in resuming household

chores, physical exercise and social interactions, which in turn, might positively affect the pain experience (Sullivan et al. 2011). Accordingly, former investigations have suggested that outcome-behavior expectancies are mediated by motivational factors, while the processes occurring during outcome-response expectancies are mainly direct and automatic (Bandura 1977; Kirsch 1985). This underlines the difficulty of disentangling different forms of expectations from one another. Although the present analysis focused on outcome expectancies, it might have failed to prevent patients from also involving their self-efficacy expectancies in assessing their expectations. Thus, depending on the domain being assessed, different types of expectancies may have the largest impact and in turn different results may be observed. Following this thread of thought, it came to no surprise that the achievement of expectations concerning walking ability, independence in everyday life and general function were strong predictors of the patients' ratings of surgery effectiveness, while pain expectation fulfillments were not. It seems logical that patients would base their assessment of surgery effectiveness on the fulfillment of expectations concerning volitional responses, that are under their direct control, instead of on the expectation fulfillment of a nonvolitional response such as pain. Showing similar results, Sullivan et al. (2011) showed that outcome-behavior expectancies were found to be stronger predictors of physical status and pain intensity one year after total knee arthroplasty than response expectancies.

4.2 Expectations

Patients were shown to have high expectations across all included domains (hip pain, back pain, walking ability, independence, physical exercise, general function, social interactions and mental well-being). Consistent with the three main reported reasons for undergoing THA, patients were found to have the highest expectations for improvement of hip pain, walking ability and independence in everyday life. It is not known whether the high expectations observed by patients in the analysis can be attributed to dispositional optimism or to expectations that have been established based on information or experiences acquired prior to the surgery or indeed a combination of both (Mannion et al. 2009a). Dispositional optimism is the expectation that when faced with a problem across important life domains, good outcomes will generally occur (Scheier and Carver 1987). A study conducted by Saban and Penckofer (2007) demonstrated that three months after spinal surgery, patients with high levels of optimism (evaluated with the use of the

Life Orientation Test-Revised (Scheier et al. 1994)) reported better quality of life even when entered into a multivariate model with fulfilled expectations. This effect was, unfortunately, not further assessed in this thesis and future studies about expectations are advised to include optimism in their analysis.

Interestingly, patients with lower quality of life based on health and worse preoperative hip function showed the highest expectations concerning postoperative general function. Mancuso et al. (2003) also observed similar results showing that the older the patient was and the worse the preoperative hip specific and general physical status was, the higher the expectations of THA were. It is crucial that patients understand that the medical intervention, here the THA, is being carried out for the specific illness, i.e. hip-problems and does not serve as a remedy for all ongoing difficulties or illnesses (Mannion et al. 2009a). If this matter is not explicitly addressed, then possible inappropriate expectations would remain unchecked and patients would undergo THA with unrealistic expectancies, ultimately leading to disappointment with the procedure (Mannion et al. 2009a). Indeed, 37% of patients had strong hopes of progress in back pain after THA, despite the fact that the main goal of THA is to alleviate hip pain and increase hip joint mobility. Consequently, the domain back pain had the largest amount of self-reported (47%) and calculated (65%) degree of unfulfilled expectation ratings. These results suggest that patients require a better understanding about the process and goal of THA. Although some patients reported progress in back pain after the surgical intervention, the latter should not be a determining factor in assessing the global effectiveness of THA.

Kirsch et al. (2020) findings showed that 31% of the patients reported unrealistic expectations for recovery when questioned about pain expectations concerning nerve surgery. A systematic analysis showed that unrealistic aspirations, whether high or low, have contributed to a decline in patient reported outcomes (Iles et al. 2009). While some studies report expectations having a negative impact on intervention outcome, others fail to report a correlation between the two (Haanstra et al. 2012; Swarup et al. 2019). These contradicting results might be due to the lack of standardized methods used to evaluate expectations, which might hide or even exaggerate the effect of expectations on outcome. In Haanstra et al. ‘s systematic review, (2012) inconsistencies in not only the terminology and definition of expectation were observed, but also in the categorization of patient expectations. They argued that this could be because of the lack of strict definition and consensus within scientific literature. More recently, Swarup et al. (2019) found that

although an improvement in the methodology of assessing expectations has been assessed, non-validated methods are still being used, which notably diminishes the comparison and interpretation of expectation research. For example, Koenen et al. (2014), Suda et al. (2010) and Gandhi et al. (2009) measured patient expectations via the use of three to six question surveys, which encompass patient outcome measures but have not been validated. Alternatively, Cross et al. (2009) made use of the Hospital for Special Surgery (HSS) Hip and Knee Replacement Expectations Surveys, which was then further refined and validated by Mancuso et al. (Mancuso et al. 2001; Mancuso et al. 2003). Balck et al. (2016) translated the aforementioned survey into the German language and culturally adapted it, in order to provide a validated and reliable German HSS-expectation survey for hip joint replacement, which unfortunately was not yet available at the time of the analysis. To achieve more reliable, comparable and precise data, only validated expectation surveys should be used when assessing expectations in prospective inquiries.

The complexity in assessing expectations does not only entail having a validated method of measuring expectations, but also encompasses the difficulty in defining the construct of expectation and classifying the expectation being assessed. This matter has already been addressed in this thesis, where the difference between the effect of response expectancies and behavior expectancies on the intervention outcome was discussed. Furthermore, as previously discussed one could question the patient about a specific type of expectancy, the patient, however, might give a response, while being influenced by a different type of expectancy. Thus, while multiple theoretical models have attempted to explain some of the complexity, future studies are required to define the interplay among the different expectation types and to refine existing expectation theories, to allow a consensus about what should be measured and how it should be evaluated.

4.3 Fulfillment of expectations

Patients' self-reported fulfillment of expectations was more optimistic in all domains than the actual calculated expectation fulfillment was. Interestingly, the most prominent differences between the two methods of assessing expectation fulfillment were observed in the first four domains (hip pain, walking ability, independence and general function) that patients chose as being the most important change that had to occur after the surgery to define the intervention as a success. For the domains hip pain, walking ability and independence, only 3% of the patients were observed to have not achieved their

expectations when asked directly, while the calculated expectation fulfillments showed that 28-52% of the patients did not meet the expected improvement. Similarly, while only 12% of patients claimed having unfulfilled expectations concerning general function, 59% did not actually meet the expected improvement in the calculated expectation. One could argue with the answer possibilities available being “Yes”, “No” and “Partly”, patients would have the tendency to reply “partly” instead of “no” to the question “Have your expectations concerning this domain been met?”. However, even when grouping the answer possibilities “No” and “Partly” together, the self-reported expectation fulfillments were still considerably more optimistic than the calculated expectation fulfillments. This overly optimistic attitude could be attributed to the social desirability bias, which describes the inclination of a subject to report the effect the subject believes is viewed as socially favorable (Stöber 1999). This could be addressed by using a method of evaluating social desirability, such as the Social Desirability Scale (SDS-17) and evaluating the analysis with the use of patients scores in SDS-17 (Stöber 1999).

According to a systematic review, Waljee et al. (2014) postulated that studies in the field of consumer marketing could shed light onto the underlying mechanism, through which expectations are associated with intervention outcomes. The expectancy - discrepancy theory hypothesized that people use their expectations as a point of reference in order to evaluate an occurrence (Oliver 1980). An individual is satisfied with the outcome, when the latter meets or exceeds prior expectations (Oliver 1980). Interestingly, this effect has been observed in this thesis, as the more the patients seemed to meet or exceed their expectations, the higher was their rating of global surgery effectiveness. It is further postulated that an individual’s assessments of an event can be influenced by previously made expectations, this is known as the assimilation-contrast theory (Ross et al. 1987). According to this theory, when an individual’s appraisal of an occurrence does not match their expectations, the person would adjust their evaluation to match the expectation (Anderson 1973). This could explain the discrepancy between the self-reported and calculated expectation fulfillment, while patients do not remember the reported preoperative expectations twelve months prior the follow-up (and can therefore not match the reported actual change to the expected change), they can report higher rates of expectation fulfillment than actually experienced. According to the cognitive dissonance theory, in order to decrease the mental discomfort that the patient would have if the postoperative change did not meet the expected change, the patient would still report having his/her expectations fulfilled (Anderson 1973). Although these theories have been

used in research for consumer assessment, they could also be relevant when analyzing patient expectations for therapeutic interventions.

4.4 Limitations

The results of the present thesis must be contemplated in light of several constraints, despite the salient characteristics such as the prospective nature, the use of standardized measures of assessment for psychological and functional domains and the detailed clinical characterization of the patient cohort.

A key limitation in this analysis, like in many others, was that no validated measure of assessing expectation was used. No standardized or validated questionnaire to evaluate expectations and their fulfillment adapted to the German language and culture, existed during the time when the analysis was planned, and the data collected. For the analysis, an adapted version of the validated expectation scale of the NASS Lumbar Spine Questionnaire for spinal surgery was used. The NASS survey entails a set of multiple-domain, multiple-answers questions where patients must report their expectations, using a Likert scale with 5 answer possibilities indicating how likely a certain outcome is (Daltroy et al. 1996; Schochat et al. 2000). The domains used in the NASS are either absolute outcomes, such as “moderate pain relief” and “complete pain relief”, or they are more vague outcomes, such as “to be able to do more everyday activities” or “to be able to do more physical exercise”. The domains were again assessed during the follow-up, where moderate and complete pain relief were grouped into one domain “pain relief”, in order to evaluate the expectation fulfillment with the response possibilities going from “not at all” to “completely”. Mannion et al. (2009a) made use of the NASS expectations, but made the test simpler and easier for patients to comprehend, while also adapting the domains to be more applicable to the group of patients suffering from nerve root compression. Mannion et al. (2009a) argued that simply asking patients about the expected improvement for a specific domain before surgery and the actual improvement during the follow-up, would be more prospective (as it wouldn’t rely on remembering preoperative expectations) and the expectation fulfillment could then be objectively calculated.

The use of the NASS in the context of hip surgery with a not yet validated german translation reduces the validity of our findings. Nonetheless, all eight factors (hip pain, back pain, walking ability, independence, physical exercise, general function, social

interaction and mental well-being) assessed were applicable to the peri-operative and postoperative treatment and healing procedure after THA. The prevalence of lower back pain amongst patients undergoing THA ranges from 21 – 60% while most studies report at least half of THA candidates reporting some level of back pain complaints (Parvizi et al. 2010; Stupar et al. 2010; Hsieh et al. 2012; Staibano et al. 2014). Following Mannion et al.'s (2009a) example, expectation fulfillment was also assessed for each domain individually rather than grouping all responses to a single score, as done in the NASS survey. This could prevent effects from being diluted or hidden by domains of lesser relevance to the analysis (Mannion et al. 2009a). This is supported by the findings of this thesis, which showed that hip pain expectation and the fulfillment of walking ability, independence in everyday life and general function were the only significant predictors of the outcome, when compared to the other types of expectations and their fulfillments. Future research should, as already mentioned, make use of the now established expectation evaluations with the aim of achieving more accurate, refined and comparable results.

Similarly, for the assessment of the dependent variable the patient-rated global effectiveness of surgery a non-validated question was applied thus also reducing the validity of the analysis. Furthermore, the actual outcome was extremely positive limiting the relevance of these findings, as they pertain to very small differences in the range of perceived effectiveness of surgery.

Another limitation is that the generalizability of the findings could potentially be diminished due to the single center design in Germany (thus culturally and demographically limited) and of a potentially selective patient cohort. Also, neither the surgery approach nor which surgeon performed the THA was controlled. Both could have an impact on outcome and could thus affect the level of expectation fulfillment.

Despite efforts to prevent drop-outs (through telephone calls and reminders) 8% of the sample cohort did not participate in the concluding analysis for unknown reasons. The drop-out rate in this analysis, however, was very small compared to other prospective cohorts (Gonzalez Saenz de Tejada et al. 2014). Further, the possibility of attribution bias cannot be excluded. The optimistic patient-rated global effectiveness of surgery could have been an overestimation as a result of the specific exclusion of non-responders.

Lastly, another central limitation of the analysis is the varying number of outstanding data for the various predictors assessed in the sequential regression model. With increasing

number of predictors, a decline in the number of values was observed, which reduces the external validity, possibly leading to bias and problems with power. The sequential analysis was therefore replicated after multiple imputation of the outstanding data.

4.5 Clinical implications

In the present thesis, the walking ability before surgery, hip pain expectation, fulfillment of the expectation of walking ability, independence in everyday life and general function were found to predict patient-perceived global effectiveness of total hip arthroplasty. Specifically, the analysis shows that more optimistic expectations concerning hip pain after THA is favorable for patients while for functional domains, a more realistic approach seems to render the best outcome.

One of the central challenges that the medical world must face today is the explosion of (mis)information that has come about with the progress of technology and the modernization of society. Despite the current burst of healthcare and medical knowledge readily accessible through technological advancements, real patient comprehension of their health and medical treatment possibilities remains consistently low (O'Connor et al. 2004; Fagerlin et al. 2006). With more than half of the patients undergoing orthopedic surgeries showing unrealistically high expectations in comparison to their surgeons (Ghomrawi et al. 2011). There is a high need for better communication of information between health care providers and patients. This interaction should not only provide the patient with educational information about the risks and benefits of the intervention, the postoperative healing process and rehabilitation, but should also address the pre-determined expectations patients might have and thereby lead to shared decision-making and a truly informed consent.

It is of great importance to consider ways in which patients' expectations can be readjusted to better outcome and patient satisfaction (Swarup et al. 2019). Mancuso et al. (2008) carried out a randomized controlled analysis to measure the impact of preoperative educational courses on patient expectations of total hip and knee arthroplasty. Their findings demonstrated that through these educational courses that addressed expectations, patients' perception of satisfaction and surgery outcome can in essence be altered and enhanced (Mancuso et al. 2008). Accordingly, Thomas and Sethares (2008) showed that the provision of a multidisciplinary course for patients contributed to a better comprehension of the process and anticipated outcome of a surgical procedure.

Interestingly, a recent analysis by Auer et al. (2017) showed that better patient education does not only enhance the patient experience of the intervention but can also translate to more objective benefits for the health care system. They showed that only two brief psychological preoperative interventions optimizing outcome expectations significantly shortened the length of the intensive care stay.

Concerning expectations and their impact on outcome, it is fundamental for the research community to reach consensus of what is important to be measured (what type of expectation) and how it should be operationalized (Hafkamp et al. 2020). This can be achieved by refining current theoretical models of expectations, while also taking the interactions among the different types of expectations into consideration. Future studies should test whether each individual expectation has an influence on the outcome, when compared to the other kinds of expectations, and whether it is even possible to fully disentangle the different types of expectations from each other.

To conclude, while further investigation is required to fully comprehend the impact of expectations, patient expectations and their fulfillment clearly play a pivotal role in predicting patients' opinion of the effectiveness of a medical intervention. The findings of this thesis show that the walking ability before surgery, the level of hip pain expectation and the fulfillment of functional expectations for walking ability, independence and general function predicted the patient-perceived global effectiveness of total hip arthroplasty. Concerning sensory domains such as hip pain, overly positive expectations with regards to status after THA do not have to be dampened (Hafkamp et al. 2020). Although standardized assessments of mobility and function, pain intensity, demographic data and mental well-being were entered into the analysis, expectations and the fulfillment of expectations were still the strongest predictors of good postoperative outcome. Given the large discrepancies between expectations and actual outcome for back pain, physical exercise and associated functions, a need for the adjustment of unrealistic expectations is underlined. Importantly, in implementing intermodal and self-effective strategies into the post-THA patient-care plan, self-efficacy expectancies can be promoted to facilitate functional recovery and the overall rehabilitation process (Bandura 1977). Therefore, it is essential for medical practitioners to offer not only comprehensive information concerning the procedure, outcome and possible complications of a therapeutical intervention, but also to directly address patient expectations. This

seemingly small but crucial factor, while costing just a little extra time, can bring about a substantial benefit to the overall treatment outcome for the patient.

5 Summary

Total Hip Arthroplasty (THA) is a standard elective surgery for hip osteoarthritis when reaching the limits of conservative medicine. Patient-based assessments are increasingly used for quality judgment after medical interventions. Identifying the biological and psychological characteristics that influence patients' opinion is critical, in order to improve the efficiency and outcome of medical treatments. Emerging investigations underlines the role of expectation in predicting patient-based assessment after orthopedic interventions. The nature of the impact of expectations on patients' perceived success of THA, however, remain conflicting. Using a sequential multiple regression model, the goal of this thesis was to determine the relative importance of preoperative expectations, expectation fulfillment and symptom improvements in predicting the patient-rated global effectiveness of THA.

Ninety patients (49 female, 41 male ; mean age 63 ± 12.9 years) took part in the analysis and completed a detailed preoperative questionnaire assessing sociodemographic, clinical, functional and psychological characteristics. Expectations for eight domains hip pain, back pain, walking ability, independence, physical exercise, general function, social interactions and mental well-being) were measured using a 5-point scale. Twelve months following THA, patients reported progress changes and expectation fulfillments in each of these 8 domains. The outcome variable, i.e. the patient-rated global effectiveness of THA was evaluated on a 5-point scale. Finally, a sequential multiple regression analysis was conducted with the dependent variable being the patient-rated global effectiveness of THA.

Walking ability before THA, hip pain expectation, expectation fulfillment in walking ability, independence and general function were found to predict the patient-rated global effectiveness of THA. Patients showed high expectations for improvement in all eight domains. Patients had the highest expectations for hip pain, walking ability and independence. For all eight domains, self-reported expectation fulfillment was higher than calculated expectation fulfillment. Compared to actual improvement at twelve months after surgery, 28% of patients were overoptimistic for hip pain, around 45% for walking ability and about 60% for back pain, physical exercise, general function, social interactions and mental well-being. Patients showed better physical, functional and psychological health after surgery. Indeed, patients showed high levels of surgery effectiveness ratings with only two patients stating only some to no improvement.

In line with current investigations concerning the impact of expectations on postoperative outcome, the present thesis demonstrates the value of addressing explicitly patient expectations during the interaction between health-care provider and patient before a medical intervention. While to obtain postoperative success optimistic expectations appear to be helpful for emotional expectations such as hip pain, self-efficacy expectations and a more rational perspective should be encouraged for complex functional expectations such as for walking ability, independence in everyday life and general function to increase the likelihood of their fulfillment.

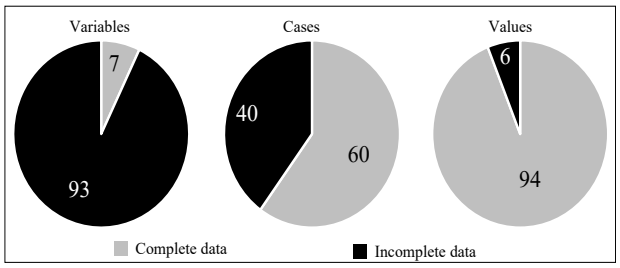
6 Attachment

A

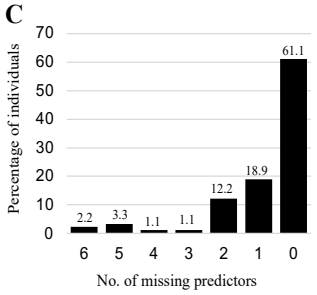
	N	Mean	Std. Deviation	Missing Count	Missing Percent	No. of imputed values (10 imputations)	Type of model used for multiple imputation[a]
Outcome measure (dependent variable): Global effectiveness of total hip arthroplasty	90	3.6	0.6	0.0	0.0	-	
Predictors: Degree of school education	90			0.0	0.0	-	
Average hip pain in the last 3 months before surgery	89	5.7	1.9	1.0	1.1	10	Logistic regression
Overall severity of chronic pain (CPG)	89	2.9	1.1	1.0	1.1	10	Logistic regression
Hip function and mobility (WOMAC)[b]	76	53.1	20.8	14.0	15.6	140	Linear regression
Walking ability (Timed up and go score)	88	1.9	0.7	2.0	2.2	20	Logistic regression
Preoperative hip pain expectation	86	4.8	0.4	4.0	4.4	40	Logistic regression
Symptom change scores:							
Average hip pain in the last 3 months	89	0.1	0.1	1.0	1.1	10	Linear regression
Hip function and mobility (WOMAC)	79	0.1	0.1	11.0	12.2	110	Linear regression
Health-related quality of life (SF-12 physical)	82	1.7	0.6	8.0	8.9	80	Linear regression
Calculated expectations-actuality discrepancy scores:							
Walking ability	86	-0.6	0.9	4.0	4.4	40	Logistic regression
Independence	84	-0.5	0.9	6.0	6.7	60	Logistic regression
Physical exercise	78	-0.9	1.3	12.0	13.3	120	Logistic regression
General function	87	-0.7	1.2	3.0	3.3	30	Logistic regression
Social interactions	81	-0.9	1.6	9.0	10.0	90	Logistic regression

a. Each model uses all other variables (including the outcome measure) as main effects.
b. Variables more than 10% missing values are marked in grey

B



C



Supplementary Figure 1. Missing data analysis of the primary outcome measure and the 14 predictor variables included sequentially into the regression model. A. The table displays the descriptives (mean and standard deviation), the percentage of missing values per variable, the number of imputed values (missing values x number of imputations) and the type of model used for multiple imputation. For the dependent variable global effectiveness of total hip arthroplasty the data set was complete. More than 10% missing values were identified for the rating instrument measuring hip function and mobility (WOMAC), its symptom change score (quotient of postoperative and preoperative WOMAC score) and the calculated expectations-actuality discrepancy scores (fulfillment of expectations) of physical exercise and social interactions. **B.** The variable chart shows that for 93 % of the predictors included at step 4 of the regression model at least 1 value is missing. The cases chart shows that 40% (N = 36) of the participants has at least one missing value on a variable. The values chart shows that 6% of the 1260 values (cases × variables) are missing. **C.** The bar graph displays the percentage of participants having none to six missing predictors. Subjects with more than three missing predictors were excluded from the analysis on the imputed data set. Figure from Kästner et al. (2021).

Supplementary Table 1. Results of the sequential multiple regression analysis. Variance explained in the global effectiveness of THA at twelve-months follow-up by sociodemographic and medical variables, change in symptoms, expectations and actuality-expectation discrepancy

Steps and variables entered upon each step	R^2	R^2_{change} (p - value)	β in final model (only significant predictor variables shown)	p-value of final β
Step 1 (N = 72): Degree of school education, average hip pain in the last 3 months before surgery, overall severity of chronic pain (CPG), hip function and mobility (WOMAC), walking ability (Timed up and go score)	0.332	-	-0.363 (Timed up and go score)	0.002
Step 2 (N = 68): Preoperative hip pain expectation	0.417	0.085 (< 0.001)	-0.363 (Timed up and go score) 0.276 (Hip pain expectation)	0.002 0.006
Step 3 (N = 62): Symptom change scores for: Average hip pain in the last 3 months, hip function and mobility (WOMAC), health-related quality of life (SF-12 Physical Component Summary)	0.473	0.056 (< 0.0001)	-0.308 (Timed up and go score) 0.327 (Hip pain expectation)	0.010 0.003
Step 4 (N = 54): Calculated actuality-expectation discrepancies for walking ability, independence, physical exercise, general function and social interactions	0.733	0.260 (< 0.0001)	0.504 (Hip pain expectation) 0.407 (Expectation fulfillment walking ability) 0.407 (Expectation fulfillment independence)	<0.0001 0.004 0.025
Final model (N = 82; hip pain expectation, actuality-expectation discrepancy for walking ability and independence)	0.402	-	-	-

Due to varying numbers of missing data per variable, sample sizes vary. Predictor variables individually significantly associated with global effectiveness of THA (see Table 6 and Table 7) were entered in four steps. The significant predictors in the final model were: Hip pain expectation and the actuality-expectation discrepancy (calculated expectation fulfillments) of walking ability and independence in everyday life. β in final model = β regression coefficient after all listed variables have been entered; R^2_{change} = Increase in explained variance by step; adjusted $R^2 = R^2 - (k - 1) / (n - k) * (1 - R^2)$ where n = no. observations, k = no. independent variables. Level of significance was set to $p < 0.05$. Table modified from Kästner et al. (2021).

7 References

Anakwe RE, Jenkins PJ, Moran M (2011): Predicting Dissatisfaction After Total Hip Arthroplasty: A Study of 850 Patients. J Arthroplasty 26, 209–213

Anderson RE (1973): Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance. J Mark Res 10, 38

Auer CJ, Laferton JAC, Shedden-Mora MC, Salzmann S, Moosdorf R, Rief W (2017): Optimizing preoperative expectations leads to a shorter length of hospital stay in CABG patients: Further results of the randomized controlled PSY-HEART trial. J Psychosom Res 97, 82–89

Ayers DC, Franklin PD, Ring DC (2013): The Role of Emotional Health in Functional Outcomes After Orthopaedic Surgery: Extending the Biopsychosocial Model to Orthopaedics: AOA Critical Issues. J Bone Jt Surg-Am Vol 95, e165-1–7

Baker PN, van der Meulen JH, Lewsey J, Gregg PJ (2007): The role of pain and function in determining patient satisfaction after total knee replacement: Data from the national joint registry for England and Wales. J Bone Joint Surg Br 89, 893–900

Balck F, Kirschner S, Jeszenszky C, Lippmann M, Günther K-P (2016): Validität und Reliabilität der deutschen Version des HSS-Erwartungsfragebogens zum Hüftgelenkersatz. Z Für Orthop Unfallchirurgie 154, 606–611

Bandura A (1977): Self-efficacy: Toward a unifying theory of behavioral change. Psychol Rev 84, 191–215

Bandura A (1978): Reflections on self-efficacy. Adv Behav Res Ther 1, 237–269

Bandura A (1990): Some Reflections on Reflections. Psychol Inq 1, 101–105

Bandura A: Self-efficacy: The exercise of control. W.H. Freeman and Company, New York 1997

Basten-Günther J, Peters M, Lautenbacher S (2019): Optimism and the Experience of Pain: A Systematic Review. Behav Med Washington DC 45, 323–339

Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW (1988): Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip

or knee. J Rheumatol 15, 1833–1840

Bellamy N, Wilson C, Hendrikz J (2010): Population-based normative values for the Western Ontario and McMaster (WOMAC) osteoarthritis index and the Australian/Canadian (AUSCAN) hand osteoarthritis index functional subscales. Inflammopharmacology 18, 1–8

Bonett DG, Wright TA (2000): Sample size requirements for estimating pearson, kendall and spearman correlations. Psychometrika 65, 23–28

Buirs LD, Van Beers LWAH, Scholtes VAB, Pastoors T, Sprague S, Poolman RW (2016): Predictors of physical functioning after total hip arthroplasty: a systematic review. BMJ Open 6, e010725

Bullens PH, van Loon CJ, de Waal Malefijt MC, Laan RF, Veth RP (2001): Patient satisfaction after total knee arthroplasty: a comparison between subjective and objective outcome assessments. J Arthroplasty 16, 740–747

Burton KE, Wright V, Richards J (1979): Patients' expectations in relation to outcome of total hip replacment surgery. Ann Rheum Dis 38, 471–474

Casser HR, Hüppe M, Kohlmann T, Korb J, Lindena G, Maier C, Nagel B, Pfingsten M, Thoma R (2012): [German pain questionnaire and standardised documentation with the KEDOQ-Schmerz. A way for quality management in pain therapy]. Schmerz Berl Ger 26, 168–175

Cohen J: Statistical Power Analysis for the Behavioral Sciences. Academic Press, New York 1977

Collins NJ, Roos EM (2012): Patient-Reported Outcomes for Total Hip and Knee Arthroplasty: Commonly Used Instruments and Attributes of a "Good" Measure. Clin Geriatr Med 28, 367–394

Cross M, Lapsley H, Barcenilla A, Parker D, Coolican M, March L (2009): Patient expectations of hip and knee joint replacement surgery and postoperative health status. The Patient 2, 51–60

Croux C, Dehon C (2010): Influence functions of the Spearman and Kendall correlation measures. Stat Methods Appl 19, 497–515

Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH (1996): The North American spine society lumbar spine outcome assessment Instrument: reliability and validity tests.

Spine 21, 741–749

Dawson J, Boller I, Doll H, Lavis G, Sharp RJ, Cooke P, Jenkinson C (2012): Factors associated with patient satisfaction with foot and ankle surgery in a large prospective study. Foot Edinburgh Scotland 22, 211–218

Deyo RA, Andersson G, Bombardier C, Cherkin DC, Keller RB, Lee CK, Liang MH, Lipscomb B, Shekelle P, Spratt KF (1994): Outcome measures for studying patients with low back pain. Spine 19, 2032S-2036S

Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G (1998): Outcome measures for low back pain research. A proposal for standardized use. Spine 23, 2003–2013

Dixon D, Pollard B, Johnston M (2007): What does the chronic pain grade questionnaire measure? Pain 130, 249–253

Donabedian A (1988): The quality of care. How can it be assessed? JAMA 260, 1743–1748

Drixler K, Morfeld M, Glaesmer H, Brähler E, Wirtz MA (2020): Validierung der Messung gesundheitsbezogener Lebensqualität mittels des Short-Form-Health-Survey-12 (SF-12 Version 2.0) in einer deutschen Normstichprobe. Z Für Psychosom Med Psychother 66, 272–286

Fagerlin A, Lakhani I, Lantz PM, Janz NK, Morrow M, Schwartz K, Deapen D, Salem B, Liu L, Katz SJ (2006): An informed decision? Breast cancer patients and their knowledge about treatment. Patient Educ Couns 64, 303–312

Franklin PD, Li W, Ayers DC (2008): The Chitranjan Ranawat Award: functional outcome after total knee replacement varies with patient attributes. Clin Orthop 466, 2597–2604

Fuchs J, Kuhnert R, Scheidt-Nave C (2017): 12-month prevalence of osteoarthritis in Germany. Journal of Health Monitoring 2, 51–56

Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, Bullinger M, Kaasa S, Leplege A, Prieto L, Sullivan M (1998): Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiologie 51, 1171–1178

Gandhi R, Davey JR, Mahomed N (2009): Patient Expectations Predict Greater Pain

Relief with Joint Arthroplasty. J Arthroplasty 24, 716–721

Gerbershagen HJ, Ozgür E, Straub K, Dagtekin O, Gerbershagen K, Petzke F, Heidenreich A, Lehmann KA, Sabatowski R (2008): Prevalence, severity, and chronicity of pain and general health-related quality of life in patients with localized prostate cancer. Eur J Pain London England 12, 339–350

Ghomrawi HMK, Franco Ferrando N, Mandl LA, Do H, Noor N, Gonzalez Della Valle A (2011): How Often are Patient and Surgeon Recovery Expectations for Total Joint Arthroplasty Aligned? Results of a Pilot Study. HSS J ® 7, 229–234

Gil JA, Goodman AD, Mulcahey MK (2018): Psychological Factors Affecting Outcomes After Elective Shoulder Surgery: J Am Acad Orthop Surg 26, e98–e104

Gonzalez Saenz de Tejada M, Escobar A, Bilbao A, Herrera-Espiñeira C, García-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 Musculoskelet Disord 15, 248

Graham B, Green A, James M, Katz J, Swiontkowski M (2014): Measuring Patient Satisfaction in Orthopaedic Surgery: J Bone Jt Surg-Am Vol 97, 80–84

Haanstra TM, van den Berg T, Ostelo RW, Poolman RW, Jansma IP, Cuijpers P, de Vet HC (2012): Systematic review: Do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? Health Qual Life Outcomes 10, 152

Hafkamp FJ, Gosens T, de Vries J, den Oudsten BL (2020): Do dissatisfied patients have unrealistic expectations? A systematic review and best-evidence synthesis in knee and hip arthroplasty patients. EFORT Open Rev 5, 226–240

Hamilton DF, Lane JV, Gaston P, Patton JT, MacDonald D, Simpson AHRW, Howie CR (2013): What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. BMJ Open 3, e002525

Hasenbring M: The Kiel Pain Inventory Manual. Three questionnaire scales for assessment of pain-related cognitions, emotions and copying strategies. Bern: Huber 1994

Haugen AJ, Grøvle L, Keller A, Grotle M (2008): Cross-cultural adaptation and validation of the Norwegian version of the Tampa scale for kinesiophobia. Spine 33, E595-601

Higashi H, Barendregt JJ (2011): Cost-Effectiveness of Total Hip and Knee

Replacements for the Australian Population with Osteoarthritis: Discrete-Event Simulation Model. PLoS ONE 6, e25403

Hossain M, Parfitt DJ, Beard DJ, Darrah C, Nolan J, Murray DW, Andrew G (2011): Does pre-operative psychological distress affect patient satisfaction after primary total hip arthroplasty? BMC Musculoskelet Disord 12, 122

Houben RMA, Leeuw M, Vlaeyen JWS, Goubert L, Picavet HSJ (2005): Fear of movement/injury in the general population: factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia. J Behav Med 28, 415–424

Hsieh P-H, Chang Y, Chen DW, Lee MS, Shih H-N, Ueng SWN (2012): Pain distribution and response to total hip arthroplasty: a prospective observational study in 113 patients with end-stage hip disease. J Orthop Sci Off J Jpn Orthop Assoc 17, 213–218

Iles RA, Davidson M, Taylor NF, O'Halloran P (2009): Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. J Occup Rehabil 19, 25–40

Jackson T, Wang Yalei, Wang Yang, Fan H (2014): Self-Efficacy and Chronic Pain Outcomes: A Meta-Analytic Review. J Pain 15, 800–814

Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, Stradling J (1997): A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? J Public Health Med 19, 179–186

Jones CA, Beaupre LA, Johnston DWC, Suarez-Almazor ME (2007): Total joint arthroplasties: current concepts of patient outcomes after surgery. Rheum Dis Clin North Am 33, 71–86

Jourdan C, Poiraudeau S, Descamps S, Nizard R, Hamadouche M, Anract P, Boisgard S, Galvin M, Ravaud P (2012): Comparison of Patient and Surgeon Expectations of Total Hip Arthroplasty. PLoS ONE 7, e30195

Judge A, Arden NK, Price A, Glyn-Jones S, Beard D, Carr AJ, Dawson J, Fitzpatrick R, Field RE (2011a): Assessing patients for joint replacement: can pre-operative Oxford hip and knee scores be used to predict patient satisfaction following joint replacement surgery and to guide patient selection? J Bone Joint Surg Br 93, 1660–1664

Judge A, Cooper C, Arden NK, Williams S, Hobbs N, Dixon D, Günther K-P, Dreinhoefer K, Dieppe PA (2011b): Pre-operative expectation predicts 12-month post-

operative outcome among patients undergoing primary total hip replacement in European orthopaedic centres. Osteoarthritis Cartilage 19, 659–667

Kästner A, Ng Kuet Leong VSC, Petzke F, Budde S, Przemeck M, Müller M, Erlenwein J (2021): The virtue of optimistic realism - expectation fulfillment predicts patient-rated global effectiveness of total hip arthroplasty. BMC Musculoskelet Disord 22, 180

Khatib Y, Madan A, Naylor JM, Harris IA (2015): Do Psychological Factors Predict Poor Outcome in Patients Undergoing TKA? A Systematic Review. Clin Orthop Relat Res 473, 2630–2638

Kirsch I (1985): Response expectancy as a determinant of experience and behavior. Am Psychol 40, 1189–1202

Kirsch I (1997): Response expectancy theory and application: A decennial review. Appl Prev Psychol 6, 69–79

Kirsch I: How expectancies shape experience. American Psychological Association, Washington 1999

Kirsch I, Kong J, Sadler P, Spaeth R, Cook A, Kaptchuk TJ, Gollub R (2014): Expectancy and conditioning in placebo analgesia: Separate or connected processes? Psychol Conscious Theory Res Pract 1, 51–59

Kirsch M, Brown S, Smith BW, Chang KWC, Koduri S, Yang LJS (2020): The Presence and Persistence of Unrealistic Expectations in Patients Undergoing Nerve Surgery. Neurosurgery 86, 778–782

Klasen BW, Hallner D, Schaub C, Willburger R, Hasenbring M (2004): Validation and reliability of the German version of the Chronic Pain Grade questionnaire in primary care back pain patients. Psycho-Soc Med 1, Doc07

Koenen P, Bäthis H, Schneider MM, Fröhlich M, Bouillon B, Shafizadeh S (2014): How do we face patients' expectations in joint arthroplasty? Arch Orthop Trauma Surg 134, 925–931

Kravitz RL (1996): Patients' expectations for medical care: an expanded formulation based on review of the literature. Med Care Res Rev MCRR 53, 3–27

Kroenke K, Spitzer RL, Williams JBW (2002): The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 64, 258–266

Landis JR, Koch GG (1977): The measurement of observer agreement for categorical

data. Biometrics 33, 159–174

Lovibond PF, Lovibond SH (1995): The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 33, 335–343

Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH, Katz JN (2002): The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. J Rheumatol 29, 1273–1279

Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF, Williams-Russo P (2001): Patients' expectations of knee surgery. J Bone Joint Surg Am 83, 1005–1012

Mancuso CA, Sculco TP, Salvati EA (2003): Patients with poor preoperative functional status have high expectations of total hip arthroplasty. J Arthroplasty 18, 872–878

Mancuso CA, Graziano S, Briskie LM, Peterson MGE, Pellicci PM, Salvati EA, Sculco TP (2008): Randomized Trials to Modify Patients' Preoperative Expectations of Hip and Knee Arthroplasties. Clin Orthop 466, 424–431

Mannion Anne F., Junge A, Elfering A, Dvorak J, Porchet F, Grob D (2009a): Great Expectations: Really the Novel Predictor of Outcome After Spinal Surgery? Spine 34, 1590–1599

Mannion Anne F, Kämpfen S, Munzinger U, Kramers-de Quervain I (2009b): The role of patient expectations in predicting outcome after total knee arthroplasty. Arthritis Res Ther 11, R139

Ng CY, Ballantyne JA, Brenkel IJ (2007): Quality of life and functional outcome after primary total hip replacement: A FIVE-YEAR FOLLOW-UP. J Bone Joint Surg Br 89-B, 868–873

Ng Kuet Leong VS, Kästner A, Petzke F, Przemeck M, Erlenwein J (2020): The influence of pain expectation on pain experience after orthopedic surgery: an observational cohort study. Minerva Anestesiol 86, 1019–1030

Nilges P, Essau C (2015): [Depression, anxiety and stress scales: DASS--A screening procedure not only for pain patients]. Schmerz Berl Ger 29, 649–657

Nilsdotter A-K (2003): Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. Ann Rheum Dis 62, 923–930

O'Connor AM, Llewellyn-Thomas HA, Flood AB (2004): Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Aff Proj Hope Suppl Variation, VAR63-72

Oliver RL (1980): A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. J Mark Res 17, 460

Pager CK (2004): Expectations and outcomes in cataract surgery: a prospective test of 2 models of satisfaction. Arch Ophthalmol Chic Ill 1960 122, 1788–1792

Palazzo C, Jourdan C, Descamps S, Nizard R, Hamadouche M, Anract P, Boisgard S, Galvin M, Ravaud P, Poiraudeau S (2014): Determinants of satisfaction 1 year after total hip arthroplasty: the role of expectations fulfilment. BMC Musculoskelet Disord 15, 53

Parvizi J, Pour AE, Hillibrand A, Goldberg G, Sharkey PF, Rothman RH (2010): Back Pain and Total Hip Arthroplasty: A Prospective Natural History Study. Clin Orthop 468, 1325–1330

Pavlov PI (2010): Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. Ann Neurosci 17, 136

Peerdeman KJ, van Laarhoven AIM, Peters ML, Evers AWM (2016): An Integrative Review of the Influence of Expectancies on Pain. Front Psychol 7, 1270

Petzke F, Hüppe M, Kohlmann T, Kükenshöner S, Lindena G, Pfingsten M, Nagel N (2020): Handbuch Deutscher Schmerz-Fragebogen.

Podsiadlo D, Richardson S (1991): The timed „Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 39, 142–148

Quartana PJ, Campbell CM, Edwards RR (2009): Pain catastrophizing: a critical review. Expert Rev Neurother 9, 745–758

Quintana JM, Escobar A, Aguirre U, Lafuente I, Arenaza JC (2009): Predictors of health-related quality-of-life change after total hip arthroplasty. Clin Orthop 467, 2886–2894

Rawdin B, Evans C, Rabow MW (2013): The relationships among hope, pain, psychological distress, and spiritual well-being in oncology outpatients. J Palliat Med 16, 167–172

Rescorla RA (1988): Pavlovian conditioning. It's Not What You Think It Is. Am Psychol 43, 151–160

Roelofs J, Goubert L, Peters ML, Vlaeyen JWS, Crombez G (2004): The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. Eur J Pain London England 8, 495–502

Rolfson O, Dahlberg LE, Nilsson J-å., Malchau H, Garellick G (2009): Variables determining outcome in total hip replacement surgery. J Bone Joint Surg Br 91-B, 157–161

Ross CK, Frommelt G, Hazelwood L, Chang RW (1987): The role of expectations in patient satisfaction with medical care. J Health Care Mark 7, 16–26

Rotter JB: Social learning and clinical psychology. New York: Prentice-Hall 1954

Saban KL, Penckofer SM (2007): Patient expectations of quality of life following lumbar spinal surgery. J Neurosci Nurs J Am Assoc Neurosci Nurses 39, 180–189

Schaal T, Schoenfelder T, Klewer J, Kugler J (2016): Determinants of patient satisfaction and their willingness to return after primary total hip replacement: a cross-sectional study. BMC Musculoskelet Disord 17

Scheier MF, Carver CS (1987): Dispositional optimism and physical well-being: the influence of generalized outcome expectancies on health. J Pers 55, 169–210

Scheier MF, Carver CS, Bridges MW (1994): Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. J Pers Soc Psychol 67, 1063–1078

Schochat T, Rehberg W, von Kempis J, Stucki G, Jäckel WH (2000): [The North American Spine Society Lumbar Spine Outcome Assessment Instrument: translation and psychometric analysis of the German version in rehabilitation patients with chronic back pain]. Z Rheumatol 59, 303–313

Shirley ED, Sanders JO (2013): Patient satisfaction: Implications and predictors of success. J Bone Joint Surg Am 95, e69

Shon WY, Park B-Y, R RN, Park PS, Im JT, Yun HH (2019): Total Hip Arthroplasty: Past, Present, and Future. What Has Been Achieved? Hip Pelvis 31, 179

Skatteboe S, Røe C, Fagerland MW, Granan L-P (2017): Expectations of pain and functioning in patients with musculoskeletal disorders: a cross-sectional study. BMC Musculoskelet Disord 18, 48

Staibano P, Winemaker M, Petruccelli D, de Beer J (2014): Total joint arthroplasty and

preoperative low back pain. J Arthroplasty 29, 867–871

Stöber J (1999): Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität. Diagnostica 45, 173–177

Stucki G, Meier D, Stucki S, Michel BA, Tyndall AG, Dick W, Theiler R (1996a): [Evaluation of a German version of WOMAC (Western Ontario and McMaster Universities) Arthrosis Index]. Z Rheumatol 55, 40–49

Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN (1996b): Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine 21, 796–803

Stupar M, Côté P, French MR, Hawker GA (2010): The association between low back pain and osteoarthritis of the hip and knee: a population-based cohort study. J Manipulative Physiol Ther 33, 349–354

Suda AJ, Seeger JB, Bitsch RG, Krueger M, Clarius M (2010): Are patients' expectations of hip and knee arthroplasty fulfilled? A prospective study of 130 patients. Orthopedics 33, 76–80

Sullivan M, Tanzer M, Reardon G, Amirault D, Dunbar M, Stanish W (2011): The role of presurgical expectancies in predicting pain and function one year following total knee arthroplasty: Pain 152, 2287–2293

Swarup I, Henn CM, Gulotta LV, Henn RF (2019): Patient expectations and satisfaction in orthopaedic surgery: A review of the literature. J Clin Orthop Trauma 10, 755–760

Thomas KM, Sethares KA (2008): An investigation of the effects of preoperative interdisciplinary patient education on understanding postoperative expectations following a total joint arthroplasty. Orthop Nurs 27, 374–381

Tilbury C, Haanstra TM, Leichtenberg CS, Verdegaal SHM, Ostelo RW, de Vet HCW, Nelissen RGHH, Vliet Vlieland TPM (2016): Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. J Arthroplasty 31, 2139–2145

Utrillas-Compaired A, De la Torre-Escuredo BJ, Tebar-Martínez AJ, Asúnsolo-Del Barco Á (2014): Does preoperative psychologic distress influence pain, function, and quality of life after TKA? Clin Orthop 472, 2457–2465

Vancleef LMG, Peters ML (2011): The influence of perceived control and self-efficacy

on the sensory evaluation of experimentally induced pain. J Behav Ther Exp Psychiatry 42, 511–517

Waljee J, McGlinn EP, Sears ED, Chung KC (2014): Patient expectations and patient-reported outcomes in surgery: A systematic review. Surgery 155, 799–808

Weldring T, Smith SMS (2013): Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). Health Serv Insights 6, 61–68

Weltgesundheitsorganisation (Eds.): The burden of musculoskeletal conditions at the start of the new millennium: report of a WHO scientific group ; [a WHO Scientific Group on the Burden of Musculoskeletal Conditions at the Start of the New Millennium met in Geneva from 13 to 15 January 2000] (WHO Technical Report Series 919). World Health Organization, Geneva 2003

White IR, Daniel R, Royston P (2010): Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. Comput Stat Data Anal 54, 2267–2275

Wirtz MA, Morfeld M, Brähler E, Hinz A, Glaesmer H (2018): Association of physical morbidity and health-related quality of life in a representative sample of older German people. Eur J Health Psychol 25, 140–151

Yoo JS, Patel DV, Mayo BC, Massel DH, Karmarkar SS, Lamoutte EH, Singh K (2019): Postoperative satisfaction following lumbar spinal fusion surgery: patient expectation versus actuality. J Neurosurg Spine 1–7

Statistik zu Hüftoperationen nach Ländern. https://de.statista.com/statistik/daten/studie/182669/umfrage/hueftgelenksoperationen-in-ausgewaehlten-oecd-laendern/; Accessed on 30.07.2020

Acknowledgments

I would like to thank Prof. Dr. med. Frank Petzke for giving me the opportunity to complete my thesis in his department. I am extremely grateful for his continuous support, leadership and kind encouragement throughout these years.

An immense thank you goes to my supervisors, Priv. Doz. Dr. med. Joachim Erlenwein and Dr. Dipl. Psych. Anne Kästner, without whom I would not have been able to complete this thesis. Joachim, I am incredibly grateful for your invaluable encouragement and support, as well as your constant good mood and positive energy. Anne, I will always think fondly of your indispensable guidance and constructive criticism of my work, as well as our long talks and your encouraging words. Thank you for all that you have done.

Last but not least, I would like to say a big thank you to Prof. Petzke, Joachim and Anne for giving me the opportunity to publish two articles in peer-reviewed journals within the last year.

Curriculum Vitae

My name is Virginie Sien-Pin Celine Ng Kuet Leong, I was born on the 15th of July 1990 in Kassel and grew up in Mauritius.

In November 2008, I passed the High School Certificate at the Loretto College of Quatre-Bornes, Mauritius. In March 2009, I moved to China and spent a year in Beijing to deepen my knowledge of my Chinese heritage and culture. In July 2010, I moved to Canada and started a Bachelor of Sciences with a double major in Psychology and Neuroscience and a minor in Physiology at the University of Toronto. After completing the Bachelor in July 2014, I started my medical studies at the University of Göttingen.

As practical experience, I completed two internships in psychiatry, one at the University of Göttingen and the other at Vivantes Klinikum Neukölln in Berlin. Furthermore, I gained practical experience in gynaecology and obstetrics, urology, ophthalmology, general medicine and pain medicine. During my practical year, I spent four months in Switzerland in the Department of Internal Medicine at the Spital Aarberg. I then spent four months at the Department of Ophthalmology in Vivantes Neukölln Klinikum in Berlin. Finally, I completed my surgical trimester at the Evangelisches Krankenhaus Göttingen-Weende.

During my medical studies, I worked part-time, from 2016 to 2019, at the Evangelisches Krankenhaus Göttingen-Weende, as a nursing assistant at the weaning ward of the Pulmonary Department in Lenglern. Moreover, from November 2015 to April 2016, I volunteered at Die Johanniter, in a refugee center in Rosdorf and Hann. Münden.

In June 2021 I completed my medical studies and have, since then, been working at the Sankt Gertrauden Krankenhaus in Berlin in the field of Ophthalmology.

Concerning my experience in research, I started my doctoral thesis in November 2017 and have published two articles (2020 and 2021) in peer-reviewed journals as co-first author.