Karolinska Institutet

Institutionen för Kvinnors och Barns hälsa

Development and evaluation of outcome measures in children with knee disorders

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Skandiasalen, Astrid Lindgrens barnsjukhus, Solna.

Fredagen den 28 mars 2014, kl 09.00

av **Maria Örtqvist,**
Leg. Sjukgymnast

Huvudhandledare:
Docent Eva Weidenhielm Broström
Karolinska Institutet
Institutionen för Kvinnors och Barns hälsa

Bihandledare:
Professor Ewa M Roos
University of Southern Denmark, Denmark
Institute of Sports Science and Clinical Biomechanics

Professor Maura D Iversen
Northeastern University, Dept. of Physical Therapy Movement & Rehabilitation Sciences/ Harvard Medical School, Boston, USA/ Karolinska Institutet, Institutionen för Kvinnors och Barns hälsa

Med Dr. Per-Mats Janarv
Karolinska Institutet,
Institutionen för Kvinnors och Barns hälsa

Professor Suzanne Werner
Karolinska Institutet,
Institutionen för molekylär medicin och kirurgi, Centrum för idrottsskadeforskning och utbildning

Fakultetsopponent:
Professor Ulla Svantesson
Göteborgs Universitet, Sahlgrenska Akademin, Institutionen för Neurovetenskap och Fysiologi, Sektionen för klinisk neurovetenskap och rehabilitering

Betygsnämnd:
Professor Karin Harms-Ringdahl
Karolinska Institutet, Institutionen för Neurobiologi, Vårdvetenskap och Samhälle, Sektionen för Sjukgymnastik

Docent Isam Atroshi
Lunds Universitet
Institutionen för Kliniska Vetenskaper, Ortopediska kliniken
Hässleholm-Kristianstad

Docent Aina Danielsson
Göteborgs Universitet, Sahlgrenska Akademin, Institutionen för Kliniska Vetenskaper, Avdelningen för Ortopedi

Stockholm 2014

From the DEPARTMENT OF WOMEN'S AND CHILDREN'S HEALTH
Karolinska Institutet, Stockholm, Sweden

DEVELOPMENT AND EVALUATION OF OUTCOME MEASURES IN CHILDREN WITH KNEE DISORDERS

Maria Örtqvist

Stockholm 2014

Cover and KOOS-Child illustrations by Mathias Karlsson

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet. Printed by Åtta.45 Tryckeri AB, Sundbyberg.

To my family

*"What degree of difficulty have you experienced during the last week **when kneeling**?"*

(In Swedish: "... ligga på knä") (KOOS-item SP5)

ABSTRACT

Background and aim: The knee joint is one of the most common sites for injury in children. Severe knee injuries are often associated with long-term symptoms, functional impairment and disability. Today there is a lack of appropriate clinical outcome measures to evaluate knee function in children with knee disorder. As a result, instruments developed for adults are often used. Unfortunately, this may lead to inaccurate evaluation and selection of treatment. The overall aim of this thesis was to develop and evaluate outcome measures for use in children with knee disorders. In *Study I*, the aim was to evaluate the reliability of knee muscle strength measurements in healthy subjects using the Strength Measuring Chair (SMC) and to evaluate the agreement between the SMC and an Isokinetic Dynamometer (ID). The aims of *Study II,* were to evaluate the Single-limb mini squat test and the Quadriceps-angle (Q-angle), as discriminative tests of medio-lateral knee position, with respect to reliability and reference values. In *Studies III and IV* the aim was to evaluate the comprehensibility of the Knee Injury and Osteoarthritis Outcome Score (KOOS) when used in children, to suggest modification for a pediatric version (KOOS-Child), as well as to evaluate the psychometric properties of the KOOS-Child when used in children with knee disorders.

Patients and Methods: In *Studies I and II*, healthy children and adults were recruited and a test-retest design was used. In *Study I*, muscle strength tests were performed in 20 children and 23 adults during three sessions; two in the SMC and one in the ID. In *Study II*, 246 children were included and dynamic and static medio-lateral knee position was assessed by the Single-limb mini squat test and by the Q-angle respectively. In *Study III and IV*, children with various knee disorders were recruited. In *Study III,* cognitive interviews were conducted with 34 Swedish children to evaluate the comprehensibility of the KOOS when used in children. According to the findings the KOOS was modified and the KOOS-Child was developed. In *Study IV*, 115 children participated in three sessions to evaluate the psychometric properties of the KOOS-Child.

Results: In *Study I*, the SMC was found to reliably measure knee muscle strength in children and adults; however, a large disagreement was found between the instruments. In *Study II*, the reliability of the Single-limb mini squat test was determined moderate and a fair to moderate reliability of the Q-angle measurements was found. Q-angle reference values varied with age and sex, however the difference may not be clinically relevant. Findings from *Studies III and IV*, showed that the KOOS was not well understood by children, thus the KOOS-Child was developed. The KOOS-Child demonstrated good psychometric properties, i.e. it is valid, reliable and responsive to clinical change when used in children with knee disorders.

Conclusion: Outcome measures for the evaluation of different aspects of knee function, specifically developed for a pediatric population is very important and necessary. In the present thesis, existing measures originally designed for adults were evaluated for use in children and new outcome measures were developed. The studies have highlighted the importance of using instruments that are specifically designated for the study population when measuring knee muscle strength, and emphasized the difficulties encountered when comparing results from different strength measuring devices. Evaluation of medio-lateral knee position showed that the Single-limb mini squat test can be used in a pediatric population however the Q-angle needs further investigation before its use can be justified. A new patient-reported outcome measure, the KOOS-Child, was also developed to measure knee function and knee-related quality of life in children with various knee disorders. KOOS-Child is recommended to be used whenever studies intend to evaluate patient reported outcomes in children with knee disorders.

LIST OF PUBLICATIONS

The thesis is based on the following original articles and manuscript. Every paper will be referred to in the text by their Roman numerals.

I. **Örtqvist M**, Bartonek A, Gutierrez-Farewik E, Broström EW. Knee muscle strength – a challenge to measure. European Journal of Physiotherapy. 2013 December 19, Epub ahead of print.

II. **Örtqvist M**, Moström EB, Roos EM, Lundell P, Janarv PM, Werner S, Broström EW. Reliability and reference values of two clinical measurements of dynamic and static knee position in healthy children. Knee Surgery, Sports Traumatology, Arthroscopy. 2011; 19(12): 2060-6.

III. **Örtqvist M**, Roos EM, Broström EW, Janarv PM, Iversen MD. Development of the Knee Injury and Osteoarthritis Outcome Score for Children (KOOS-Child): comprehensibility and content validity. Acta Orthopedica 2012; 83(6): 666-73.

IV. **Örtqvist M**, Iversen MD, Broström EW, Janarv PM, Roos EM. Psychometric properties of the Knee injury and Osteoarthritis Outcome Score for children (KOOS-Child) in children with knee disorders. Submitted.

TABLE OF CONTENTS

ABBREVIATIONS

ACL	Anterior Cruciate Ligament
ADL	Activities of Daily Living
CFA	Confirmatory Factor Analysis
CHAQ	Child Health Assessment Questionnaire
CI	Confidence Interval
COSMIN	Consensus-based Standards for the Selection of Health Measurement INstruments
CTT	Classical Test Theory
EFA	Exploratory Factor Analysis
EQ-5D-Y	EQ-5D-Youth version
GPE score	Global Perceived Effect score
ICC	Intraclass Correlation Coefficient
ICF	International Classification of Functioning, Disability and Health
ID	Isokinetic Dynamometer
IKDC	International Knee Documentation Committee Subjective Knee form
Pedi-IKDC	Modified IKDC Subjective Knee form for children and adolescents with knee disorders
IRT	Item Response Theory
KOOS	Knee injury and Osteoarthritis Outcome Score
KOOS-Child LK1.0	Knee Injury and Osteoarthritis Outcome Score for Children (preliminary version)
KOOS-Child LK2.0	Knee Injury and Osteoarthritis Outcome Score for Children (final version)
MeSH	Medical Subject Headings
MIC	Minimal Important Change
Nm	Newton-meter
PRO	Patient-Reported Outcome
Q-angle	Quadriceps-angle
QoL	Quality of Life
ROC	Receiver Operating Curves
SD	Standard Deviation
SDC	Smallest Detectable Change
SEM	Standard Error of Measurement
SMC	Strength Measuring Chair
Sport/Rec	Sport and Recreational
VAS	Visual Analogue Scale

THESIS AT A GLANCE

STUDY	AIM	METHODS/DESIGN	RESULTS	CONCLUSIONS
I	To evaluate the reliability of strength measurements in the Strength Measuring Chair (SMC) when used in healthy children and adults and to assess agreement between the SMC and an ID (Biodex).	Muscle-strength tests were performed in a test-retest design during three sessions: two in the SMC and one in the ID in 20 healthy children (5-13 years) and 23 adults (23-60 years).	Excellent intra-subject reliability (ICC 0.93-0.99) in both instruments and an excellent test-retest reliability (ICC 0.87-0.93) of measurements in the SMC were found. Disagreements between instruments were found in both groups, but were more pronounced in the children's group.	The SMC, reliably measured knee muscle strength in children and adults. However, the large disagreement found between the instruments warrants care in standardizing measuring positions for body size and demands caution in comparing one's muscle strength measured using different methods.
II	To evaluate the reliability of two tests, one of dynamic and one of static medio-lateral knee position when used in healthy children, to present pediatric reference values, and to evaluate the association between the two tests.	246 healthy children (9-16 years) were included in a test-retest design. Dynamic and static medio-lateral knee position was assessed by the Single-limb mini squat test and by the Q-angle respectively.	Single-limb mini squat test: moderate reliability (kappa 0.48–0.57, 76–79% agreement). Q-angle: fair to moderate reliability (ICC 0.35-0.42, SDC 4-5°). Q-angle reference values varied with age and sex (mean 13.5°– 15.3°). No association was found between the two tests.	The difference in Q-angle between age and sex was <5° and may not be clinically relevant. The results indicate that two different concepts were measured. We suggest that the Single-limb mini squat test can be used, but the Q-angle needs further investigation to justify its use.
III	To examine the comprehensibility of the KOOS in children with knee disorders and to recommend modifications for a pediatric version (KOOS-Child)	A qualitative approach was used, and cognitive interviews were conducted with 34 Swedish children (10-16 years) with knee disorders to evaluate the comprehensibility of the KOOS.	Many children found the instructions confusing and had difficulty tracking items based on time frame as well as difficulty understanding terms. Mapping errors resulted from misinterpretation of items and from design issues. Several items were found irrelevant.	The KOOS is not well understood by children, especially younger children. Modifications were made to the KOOS based on qualitative feedback from the children, resulting in the pediatric version, the KOOS-Child.
IV	To refine the preliminary KOOS-Child through the process of deleting redundant items and to test the final KOOS-Child for construct validity, reliability, responsiveness, and interpretability in children with knee disorders.	115 children (7-16 years) with knee disorder were evaluated in three sessions (baseline, 1-3 weeks, and 3 months). Several PROs were completed (KOOS-Child, CHAQ, EQ-5D-Y, GPE scores, and VAS scores of knee function). An anchor-based approach was used to evaluate responsiveness and interpretability.	Excellent test-retest reliability (ICC 0.78-0.91) and high internal consistency (α=0.80-0.90) except for symptom subscale (α=0.59) was found. Construct validity was confirmed. Greater Effect Sizes were found in those reporting improvement. MIC changes were greater than the SDC for patients reporting to be better and much better.	The KOOS-Child demonstrates good psychometric properties when used in children with knee disorders. It can be used in monitoring individuals and groups and is responsive to change. We recommend the use of the KOOS-Child when evaluating children with knee disorders.

1 INTRODUCTION

Childhood participation in sports is widespread and provides major psychosocial, physical, and health-related benefits. Unfortunately, sports participation also increases the participants' risk of overuse and acute knee injuries [1,2]. Severe knee injuries, such as ligament injuries, are associated with long-term symptoms and sequelae, such as knee osteoarthritis, during the second or third decade of life [3-6]. This is a major concern for the injured children, clinicians, and researchers. Unlike literature on evaluation methods, treatments, and factors that influence outcomes in adults with knee injuries, literature regarding knee injuries in children is sparse. Due to lack of appropriate clinical outcome measures for children with knee injuries, instruments developed for adults are often used in clinical settings and research. The use of adult measures in children may lead to inaccurate evaluation and selection of treatments [7]. Thus, the purpose of this thesis was to develop and evaluate outcome measures designed specifically for children with knee disorders in order to accurately evaluate treatments and long-term health outcomes in these children.

1.1 ANATOMY AND INJURY PATTERN OF THE KNEE IN THE GROWING CHILD

The knee is a complex joint that sustains great pressure in weight bearing and consists of two joints working together as a hinge; the tibiofemoral joint and the patellofemoral joint. Multidirectional forces affect the knee joint. Motion and stability are balanced by static and dynamic structures working together. The skeleton possesses underlying stability. Main additional static stabilizers include the Anterior Cruciate Ligament (ACL), the Posterior Cruciate Ligament, the Medial and Lateral Collateral Ligaments, the Medial patellofemoral ligament and the Medial and Lateral menisci, complemented by dynamic stability of the neuromuscular system [8,9].

The injury pattern in the musculoskeletal system is different in children and in adults due to children's skeletal immaturity and differences in biomechanical properties. The weakest points are the growth plates and insertion sites of muscles, tendons, and ligaments [10,11]. Skeletally immature individuals will have fractures involving the growth plates. Avulsion fractures in the attachments of muscles, tendons, and ligaments are more frequent compared with ruptures in the soft tissue structures [10]. The timing of puberty and subsequent skeletal maturity is highly individual including large variations, but on average, the peak of growth spurt occurs at the age of 11.5 and 13.5 years in girls and boys respectively [12].

1.2 KNEE INJURIES IN THE GROWING CHILD

The knee joint is one of the most common sites for injury in younger individuals and sports-related injuries are at the leading cause of injury in adolescents [1,2,11]. In Sweden, sports-related injuries account for 28 percent of all injuries in children aged 0 - 17 years presenting to an emergency unit for medical care. These injuries peak in children 13-15 years of age [13]. In a review of the prevalence of knee injuries in adolescents aged 13-19 years, Louw et al [14] reported that injury rates vary considerably among studies, values ranging from 10% to 25% (number of knee injuries/total number of injuries); with more recent studies reporting higher prevalence [14].

The risk of sustaining a knee injury is influenced by sex and type of activity and increases with activity level and age [1,11,14,15]. In general, knee injuries are more common in adolescent girls than in boys [11,14,15].

Knee injuries are often divided into overuse and acute knee injuries [1]. Two of the most common long-term overuse knee injuries in children and adolescents are the patellofemoral pain syndrome [16,17] and the Osgood Schlatter lesion [18]. Both diagnoses represent common knee problems in children and adolescents who participate in sports and usually present during the growth spurt [18,19]. Children with patellofemoral pain syndrome present with non-specific knee pain and/or a feeling of patellar instability [16]. Children diagnosed with Osgood Schlatter lesion often experience swelling and tenderness over the tuberositas tibiae (the insertion of the patellar tendon) and complain of exercise-induced pain [18]. For both these diagnoses, the symptoms usually disappear gradually or when skeletally maturity is reached [16].

Knee injuries that lead to knee joint hemarthrosis can indicate a serious acute knee injury [20]. Traumatic patellar dislocation is the most common acute knee injury in children and adolescents; its yearly incidence is 1/1000 in children aged 9-15 years [21]. A less common but also serious acute knee injury sustained during childhood is the Anterior Cruciate Ligament (ACL) injury [22]. Its true incidence in skeletally immature children is not known [23]; however, it has been suggested that the incidence is increasing. For both these diagnoses, most injuries occur during sports [11,15,21,24-26] and treatment involves either physiotherapy or surgery. However, these different treatment algorithms still have limited evidence in the literature [22,26,27].

1.3 PREVENTING KNEE INJURIES IN CHILDREN

Prevention programs to reduce the knee injury risk in children are on the rise. A significantly reduced incidence of knee injuries in adolescents and adults has been reported following the implementation of prevention strategies, such as neuromuscular warm-up programs [28-30]. In adolescent female football players, following the implementation of a 15-minute neuromuscular warm-up program twice a week, there was a reduction by almost two-thirds in the overall rate of ACL injury [30]. However, the use of prevention warm-up programs has only suggested limited effects in children younger than 12 years of age [31].

1.4 THE CONCERN OF EARLY OSTEOARTHRITIS AFTER KNEE INJURY

Acute knee injuries in children are often associated with long-term symptoms, functional impairment and disability [2,15,32]. The evidence regarding treatments and long-term prognosis and health outcomes after childhood knee injuries is limited [33]. Serious knee injuries may lead to reduced future involvement in physical activity and to less optimal health later in life. Another concern is the possible early development of knee osteoarthritis, a degenerative joint disease that affects the articular cartilage, bone and synovium [5,32-34]. Osteoarthritis requires long-term management and may lead to disability and an inability to participate in sports and daily activities. Osteoarthritis is most common in adults aged 50 years or more [35]. However, ACL and meniscal injuries sustained by young adults significantly increase the risk of early osteoarthritis [3,4,6,32,34,36]. Previous studies demonstrate that the development of osteoarthritis after knee injuries may occur as early as at age of 30 [3,4,6,33,36]. Data on whether children and adolescents develop osteoarthritis owing to sports-related acute knee injuries are limited [5,37]. However, in a review evaluating osteoarthritis as an outcome of pediatric sports by Caine et al [5], they concluded that a link exists between youth sports injury - especially acute knee injuries and osteoarthritis [5].

1.5 EVALUATION AND OUTCOME MEASURES

When evaluating knee function in children and adolescents with a knee injury, several clinical outcome measures are often applied. In the present thesis, different outcome measures of muscle strength, static and dynamic measures of knee position, and patient-reported outcomes (PROs) are presented.

Clinical examination often consists of radiological measures, joint laxity tests [38], and measures of knee position [39,40]. Clinical outcome measures of physical function are usually based on muscle strength assessments [41] and on different functional performance tests, such as hop tests [42-44]. PRO measures are often represented by different questionnaires that add information regarding the patients' own perception of knee function [45,46]. A PRO measure can express clinical symptoms, activity limitations, and participation restrictions that the patient may experience in daily life.

Many functional performance tests have shown good validity and reliability when used in adults [43,44]. But until today, none of the available functional tests has been psychometrically assessed for use in a pediatric population with knee disorders. Furthermore, there is no reliable and valid self-reported Swedish questionnaire aimed at measuring knee outcome in children suffering from a knee disorders.

Earlier studies demonstrate that measures such as static joint laxity and radiographic changes, which are frequently reported outcome after ACL-injury in adults and in studies of progression of osteoarthritis, poorly correlate with patient's self-reported symptoms and function [6,47]. In addition, low to moderate correlations between maximal muscle strength measured in an ID and functional performance tests are also reported in the literature (r=-0.31-0.78) [48-50]. These data highlight the importance of including several outcome measures in order to focus on the whole individual and on functioning in daily life.

1.5.1 Clinical outcome measures

Measures of muscle strength

Muscle strength measurements and different functional performance tests are often used in subjects with knee disorders [41-44,51], as muscle function may influence the risk of sustaining knee injuries and outcomes following knee injuries [52,53].

Muscle strength measurements after treatment and as an explanatory factor for muscle weakness are often used in research and clinical settings [41,54-56]. Different strength measuring techniques are used, such as the Hand-Held Dynamometry and Isokinetic Dynamometry (ID) [41]. The Hand-Held dynamometer is easy to use in clinical settings and commonly used in children with muscular dystrophy [54] and cerebral palsy [55]. The Hand-held dynamometer has been shown to measure isometric strength with a good reliability in different populations [57,58]. However, there are disadvantages to this method, such as influence of the tester's strength and difficulties stabilizing the subject's testing position, especially when used in stronger subjects and muscle groups [59,60].

The ID is widely used when evaluating muscle strength in rehabilitation and sports science and is regarded as the most valid strength assessment method in adults [41,56]. However, there are disadvantages to this method, including cost and the difficulty of modifying the instrument to test children [41,61]. The reliability of the ID when measuring knee muscle strength in adults is good (Intraclass Correlation Coefficient (ICC) 0.86-0.99) [62,63], but varying results have been reported in children [41,61,64]. Holm et al. [49] have presented reference values of isokinetic knee muscle strength using an ID in children 7-12 years old. The results show, that isokinetic knee muscle strength increases with age, but no sex differences were present up to age 11. There was also large variability within each age group, indicating that a normative sample of muscle strength measurements includes a wide range of values for each age group. Children below age 7 were not included in the analyses because of difficulties in testing procedure in the ID and the need to modify the testing positions [49].

The Strength Measuring Chair (SMC) is a stationary dynamometer, that was developed by the Motion Analysis Laboratory at Astrid Lindgren's Children's Hospital to achieve a more standardized assessment of isometric muscle strength in the lower extremities [65]. This method has proved useful and reliable (ICC 0.84-0.87) for measuring isometric plantarflexor strength in healthy children and adults [65].

Measures of static and dynamic knee position

Knee malalignment is associated with knee function, risk of future knee injuries, and the progression of degenerative processes of the knee joint [32,53,66-69]. Knee malalignment is also related to worse patient-reported function in adults after knee injury [70]. Several authors have suggested that the prevention of a "knee medial to foot" position can reduce the risk of knee injuries [42,71]. Based on this knowledge, it is important that appropriate measures be used to evaluate the dynamics of knee alignment in children. The dynamics of knee malalignment depend on both static and dynamic measures of knee position [66]. In the present thesis, two measures were used in order to evaluate medio-lateral knee position: the Q-angle (static measure) [40] and the Single-limb mini squat test (dynamic measure) [39,51].

The Q-angle is described as an index of the vector for the combined pull of the extensor mechanism and the patellar tendon [40]. It is formed by the intersection of two lines: one from the anterior superior iliac spine to the center of the patella, and the other from the center of the patella to the tibial tubercle. The Q-angle is a debated method [72], although it is still used in clinical practice, especially by orthopedic surgeons, to assess the risk of patellar dislocation. A Q-angle of more than 20 degrees has been suggested as one of many risk factors for patellar dislocation [73]. Few studies have evaluated the Q-angle in children, and no consensus in terms of reference values exists [68,74,75]. Shultz et al. [75] described differences in Q-angle values in children between sexes and maturational status groups, where girls from the group with a median age of 13 years (range 12-18) showed higher Q-angles than the respective group of boys with a median age of 14.5 years (range 12-18). However, no differences where found in younger ages (median age 11, range 9-14) [75].

The Single-limb mini squat test is a functional test, easy to administer, and developed to resemble activities of daily living, such as descending stairs [39,51,76]. The Single-limb mini squat test defines knee function as the maximum numbers of knee bendings per 30 seconds in order to evaluate the subject's ability to perform fast changes between eccentric and concentric muscle force over the joint [51,76]. The test is also a dynamic measure of medio-lateral knee position [39]. The Single-limb mini squat test demonstrates good reliability and validity when used in adult subjects [39,51], but until now, it has not been evaluated when used in children.

Patient-reported outcome measures

Patient-reported outcome (PRO) measures are an important supplement to other clinical outcome measures because they add the patient's perspective. PROs can provide information regarding activity limitations and participation restrictions experienced after a knee injury. There is a general consensus that PRO measures should serve as the gold standard in assessment of musculoskeletal conditions, in which the patient's perspective and health-related quality of life (QoL) are of primary interest [45,46]. Today, PRO measures are often used as the primary outcome measure in clinical trials of knee injury and can be provided through different scores. A PRO measure can be either generic or disease-specific. Generic scores allow comparison across diagnoses, but are less responsive to changes in specific conditions [77]. Examples of generic scores are the EQ-5D (adult version) and EQ-5D-Youth version (EQ-5D-Y), which measures health-related QoL [78,79]. Examples of disease-specific PRO measures developed for adult subjects with knee disorders include the Knee injury Osteoarthritis Outcome Score (KOOS) [80], the International Knee Documentation Committee Subjective Knee form (IKDC) [81], the Tegner activity scale [82], and the Lysholm knee scoring scale [83].

The KOOS is a widely used self-administered disease-specific measure designed to assess patient-reported knee outcomes among adults with joint injuries or degenerative diseases. The KOOS addresses five domains: pain, symptoms, activities of daily living (ADL), sports/recreation (Sport/Rec), and knee-related Quality of Life (QoL) [80,84,85]. It is distinguished from other knee-specific measures by the inclusion of separate subscales for evaluating the aforementioned domains and by the presentation of separate subscale scores as a profile. The KOOS has been shown to be highly reliable, valid, and responsive to change in adult subjects with different knee disorders, as well as for different interventions – among them ACL injury, meniscectomy, cartilage repair, and total knee replacement [84,86,87]. However, until today, its applicability in children had not been established.

Data indicate that the use of generic adult PRO QoL measures for evaluating QoL (SF-36 and EQ-5D) in pediatric populations with orthopedic conditions are inappropriate due to high ceiling effects [7]. Iversen et al. [88] has also showed the lack of comprehensibility when the adult version of the IKDC is used in children with knee disorders [88]. These findings highlight the importance of specifically designed outcome measures that reflect the unique needs of the pediatric population.

As mentioned earlier, Swedish-language PRO measures available for children with knee disorders are lacking. There is one questionnaire recently developed by Iversen et al. [88], the Modified IKDC Subjective Knee form for children (Pedi-IKDC), that evaluates knee function after knee injury in children [88,89]. Unfortunately, this questionnaire is only available in English, uses an aggregated total score, and does not include dimensions for higher physical function or knee-related QoL.

The process for developing pediatric PRO measures is similar to the process used to develop adult measures. However, some additional issues need to be considered, such as readability, age-related vocabulary, language comprehension, comprehension of the construct being measured, and recall issues. Further, the response scales, wording, and format may need modification to take into account children's language and cognitive skills [88,90]. It has been shown that children as young as 5 can reliably and validly self-report their health-related QoL if an age-appropriate PRO measure is used [91].

1.6 PSYCHOMETRIC EVALUATION

Before using any evaluation method in a healthcare or research setting, it is essential to evaluate the quality of the measurement properties, i.e. the psychometric properties [77,92-94]. Important psychometric properties and characteristics of an outcome instrument are reliability, validity, responsiveness and interpretability [77,94,95]. These properties are closely related to the population and testing situation, which means they refer to the results obtained from a measurement, not to the instrument itself [77].

There is an ongoing debate regarding the terminology of psychometric properties [77,96]. A consensus regarding the taxonomy, terminology and definitions of measurement properties for health-related instruments was recently reached by an international consensus group called the COSMIN (COnsensus-based Standards for the Selection of health Measurement Instruments; Fig. 1) [95,96]. All definitions of psychometric properties used in the last two studies in the present thesis follow these recommendations and are introduced below.

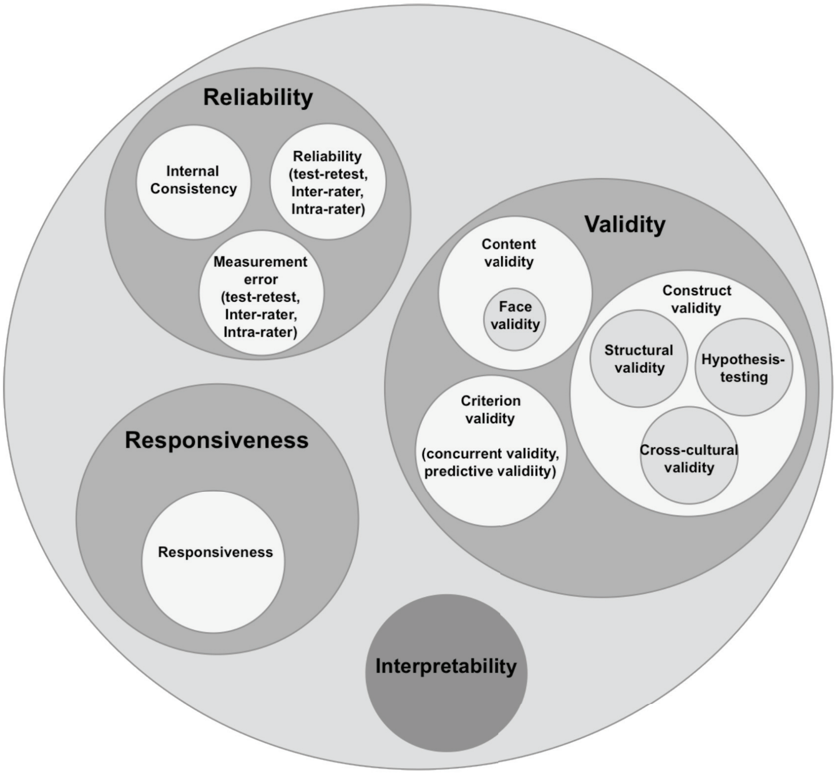


Figure 1. COSMIN taxonomy of relationship of measurement properties (Mokkink et al., Qual Life Res (2010) [95] (Modified and printed with permission of Springer)

1.6.1 Reliability

For an evaluative outcome measure, it is important that repeated measurements for stable patients remain constant over time to ensure that the error involved in measurement is small enough to detect actual changes in the construct being measured [77]. Many terms for reliability have been used over the years in the literature, such as; *reproducibility, repeatability, consistency* and *agreement* [94,96]. Despite this incongruence in terminology, reliability in general refers to the degree to which the measurement is free from measurement error for repeated measurements under various conditions [77].

The COSMIN definition of *reliability* is "the extent to which scores for patient who have not changed are the same for repeated measurements under several conditions: e.g., using different sets of items from the same multi-item measurement instrument (internal consistency), over time (*test-retest*), by different persons on the same occasions (*inter-rater*), or by the same person (i.e. raters or responders) on different occasions (*intra-rater*)" [96]. According to the COSMIN consensus the domain reliability, consists of several measurement properties: reliability, internal consistency, and measurement error [96]. *Internal consistency* is defined as the degree of interrelatedness among items in a PRO measure (i.e., the degree to which the items of the questionnaire measure the same variable) [96]. *Measurement error* is defined as the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured [96]. Systematic errors occur constantly and in one direction; for example, the learning effect during muscle testing – this affects the validity more than reliability. However, random errors, that are the result of chance are the biggest issue regarding reliability [92].

Different statistical approaches evaluating aspects of reliability are available and discussed in the literature [92,93,97]. Several approaches are considered relevant, such as the Intraclass Correlation Coefficient [97], Cohen's kappa [98], and the Bland Altman 95% limits of agreement test [93]. But no test alone provide sufficient information, therefore, it is recommended that several complementary tests be used [77,92,99]. There are several guidelines for interpreting reliability statistical results [97,98], but ultimately the results depend on the clinical circumstances, and this needs to be considered when defining acceptable reliability of an outcome measure [77].

1.6.2 Validity

The reliability of an outcome measure is necessary but not sufficient: the validity of the instrument also needs to be evaluated. Validity concerns whether an instrument measures what it intends to measure – in other words, whether one can draw the accurate conclusions from the measurements [77] . Several different terms and 'types' of validity have been described in the literature [77]. According to COSMIN, validity can be divided into: *content, construct* and *criterion validity* [96].

Content validity is defined as "the degree to which the content of a measurement instrument is an adequate reflection of the construct to be measured" [96]. For a PRO measure, content validity is dependent on both the relevance and the comprehensibility of the items in a questionnaire. This is highly group- and setting-specific. The relevance of items could either be judged by an expert panel, and when it concerns a PRO, the patient should be the expert [94,100]. Comprehensibility of a questionnaire can be evaluated in different ways. Cognitive interviewing is one approach that has been applied in the present thesis [101-103]. This technique has been shown to be effective in determining children's understanding of the questions' meaning prior to distribution of a questionnaire [88,102,104,105]. Cognitive interviewing has previously been used to ascertain children's understanding of the items in an adult knee-specific outcome measure when used in children [88].

Construct validity is defined as "the degree to which the scores of a measurement instrument are consistent with hypothesis, e.g., relationships with scores of other instruments or differences between relevant groups based on the assumption that the instrument validly measures the construct to be measured" [96]. When performing hypothesis testing, it is important to include specific hypotheses, which should also include the magnitude and direction of the expected correlations [100].

Criterion validity is defined as the degree to which the scores of a measurement instrument constitute an adequate reflection of a gold standard [96]. Thus, when no gold standard outcome measure exists – e.g., a PRO measure evaluating knee function in children with knee disorders – evaluation of criterion validity is not possible.

1.6.3 Responsiveness

When evaluating the effect of an intervention or treatment in the clinic or for research purposes, it is imperative that the outcome measures used detect changes in the measured construct between repeated assessments. The capacity of an outcome instrument to measure change has been labeled differently in the literature, such as *responsiveness, sensitivity to change* and *longitudinal validity* [96,106]. Despite the differences in terminology, the concept entails comparing a mean score in a group of clinically improved patients and in a group of stable patients [77]. Responsiveness has a close relationship to construct and criterion validity, but the main difference between these concepts is that validity refers to a single score, whereas responsiveness involves the validity of a change score between repeated measurements [96,100]. This thesis applies the COSMIN definition of responsiveness: "the ability of an instrument to detect change over time in the construct to be measured" [96]. Because validity and responsiveness are closely related, the COSMIN panel proposed that similar statistical approaches be used when evaluating responsiveness, i.e. use of hypotheses testing and pre-defined hypotheses including expected directions and the magnitude of correlations [96]. One common approach to evaluate responsiveness is to follow patients during a period of time and administer the outcome measure at the beginning and at the end of the period. An anchor-based method using an external criterion, e.g., asking patients to rate whether they have improved or not, is used and the group that has changed is compared with the one that did not [77]. Patient retrospective ratings of change using a global transition question (as the anchor or external criterion) were first reported in 1989 [107], and this has become the most commonly used method for determining clinical change [108]. However, this approach has been questioned because of problems involving retrospective judgment – namely, the difficulties for patients to remember their initial status (recall bias) [109,110]. Another common approach is to randomly assign patients to a treatment and a control group and to measure health status before and after treatment [77].

Different measures, such as the Cohen's Effect Size (the ratio of the mean change to the Standard Deviation (SD) of the baseline scores) [111] and the Standardized Response Mean (the ratio of the mean change to the SD of change scores) are frequently reported as measures of responsiveness [77]. The appropriateness of these measures, however, has been questioned since they are considered to measure the magnitude of change, rather than a measure of the quality of the measurement instrument [106]. Both Effect Size and Standardized Response Mean are, however, widely used and are commonly reported in order to describe the responsiveness of different PROs for the knee [84,112].

1.6.4 Interpretability

Beyond the psychometric properties' reliability, validity, and responsiveness, it is also important that the score of the outcome measure be interpretable, that the users understand the meaning of the measurement results. The COSMIN definition of interpretability is "the degree to which one can assign qualitative meaning to an instrument's quantitative scores or change in scores"[96]. To evaluate the interpretability of a questionnaire, the distribution of total/change scores in the population and in subgroups, floor and ceiling effects, and estimates of Minimal Important Change (MIC) are often presented [94,96]. A floor effect is present if a questionnaire is unable to demonstrate a worse score in patients who shows deterioration in clinical status, and a ceiling effect is present if the questionnaire is unable to show the improvement of a clinically improved patient [113]. Outcome measures should illustrate the extent to which a change in score reflects the changes in clinical status that a patient perceives as important.

The literature attests several terms and definitions for clinical improvement that are used interchangeably. King et al. [108] have reported that there are, in fact, 14 available terms and definitions of the MIC. The most important difference is the discrepancy in methods concerning whether the patient's perception of change is included in the method or not [108,114,115]. One of the most commonly cited definitions in the literature is by Jaseschke et al. [107]: "the smallest difference in score in the domain of interest, which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management" [107].

In this thesis, the minimal important change is referred to as the MIC. There are two main approaches used to define the MIC: the anchor-based approach (using an external criterion to define an important change; e.g., the patient's retrospective ratings of change or comparing changes in a disease-related outcome), and the distribution-based method (using statistical methods as a value for the MIC).

The most common anchor-based approach for estimating the MIC is the mean change method [107,108,116]. The mean change method defines the MIC as the mean change in scores of patients categorized by the anchor as having experienced minimally important improvement/deterioration [107]. Typically, the mean change in the groups that differ by "a little" is taken as an estimate of the MIC [108]. However, no consensus exists about which method to choose, and large variations in MIC values have been found using different methods within studies and using the same method across studies [115].

Distribution-based approaches focus on the minimal change that falls outside the measurement error and represent the "true score" of a measure. However, the disadvantage is that they do not focus on the importance of the observed change. Several terms have been used for this purpose, including *Minimal Detectable Change (MDC), Smallest Detectable Change (SDC), Smallest Real Change (SRC),* and *repeatability coefficient* [107,108,116-118]. In this thesis both *SDC* and *repeatability coefficient* are used for the same purpose of evaluating changes beyond measurement error [113,119]. De vet el. [116] recommend that anchor-based approaches be used when evaluating the MIC, since they include a definition of what is minimally important. When assessing the interpretability of a score, it is also important to judge whether the SDC is small enough to detect the MIC [116].

1.7 RATIONALE FOR THIS THESIS

There are currently few good predictive indicators or reputable methods for deciding which treatment to choose in children with knee disorders. There is a need to evaluate and develop valid and reliable outcome measures that can define the factors influencing treatment outcomes in order to evaluate existing treatment and plan new treatment strategies to improve patient outcomes.

The studies comprising this thesis extend the knowledge and understanding of outcome measures when treating children with knee disorders in clinic settings or in research. Future research in this area could make an important contribution to improving clinical guidelines and could potentially yield significant health benefits for treating children with knee disorders.

2 AIMS

The overall aim of the thesis was to develop and evaluate outcome measures used in different settings in children with knee disorders.

Specific aims of the studies were:

STUDY I

a) To evaluate the test-retest and intra-subject reliability of isometric knee extensor and knee flexor muscle strength measurements using the Strength Measuring Chair (SMC); b) to evaluate the agreement between the SMC and an Isokinetic Dynamometer (ID); and c) to determine whether there was a difference in measurement results between the instruments in children and adults.

STUDY II

To evaluate two clinical tests easy to use in practice as discriminative tests of medio-lateral knee position. Specific aims were (a) to test the intra- and inter-rater reliability of the Single-limb mini squat test; (b) to evaluate the intra- and inter-rater reliability of the Q-angle; (c) to describe pediatric reference values of the Q-angle; (d) to compare Q-angle measurements across age, sex and limb in healthy children; and (e) to evaluate the association between the Single-limb mini squat test and the Q-angle.

STUDY III

To examine the comprehensibility of the Knee Injury and Osteoarthritis Outcome Score (KOOS) in children 10–16 years of age with knee disorders and document interpretability and response errors in order to recommend modifications for a pediatric version (KOOS-Child).

STUDY IV

To refine the preliminary KOOS-Child through the process of deleting redundant items and to test the final KOOS-Child for construct validity, reliability, responsiveness and interpretability in children with knee disorders.

3 METHODS

Several outcome measures have been used in the four studies that comprise the present thesis (Table I). A brief summary of the study outlines are presented below, followed by a more detailed description in tables and text regarding participants, evaluation methods, testing procedure and data analysis.

3.1 STUDY OUTLINES

STUDY I: A test-retest design was used to evaluate the reliability of the Strength Measuring Chair (SMC) [65] to evaluate isometric knee muscle strength in healthy children and adults as well as the agreement between the SMC and an Isokinetic Dynamometer (ID) (Biodex System 3). Twenty healthy children and 23 healthy adults participated in the study. The subjects were recruited between 2005-2006. The same examiner performed strength tests at three different sessions, two sessions in the SMC and one session in the ID, with a one-week interval in between.

STUDY II: A test-retest design was used to evaluate the reliability of the Single-limb mini squat test (a dynamic clinical measure of medio-lateral knee position) [39,51] and the Quadriceps-angle (Q-angle) (a static clinical measure of medio-lateral knee position) [40]. Additionally, the study was designed to present pediatric reference values of the Q-angle and to evaluate the association between the tests. Two hundred and fifty-three healthy children were recruited from three schools in the Stockholm area between 2007-2008. Three examiners performed evaluation of the Single-limb mini squat and the Q-angle at two different occasions with a one-week interval between assessments.

STUDY III: A qualitative approach was used to evaluate the comprehensibility of the KOOS questionnaire [80,85]. Cognitive interviews were conducted in 34 children with symptomatic knee disorders [103,120,121]. Patients were recruited between 2009-2010 and selected to allow for equal age and sex group representation. All interviews were audiotaped and the data were transcribed and analyzed to determine modifications to the original KOOS, in order to develop a new pediatric version, the KOOS-Child [88,103,122-124].

STUDY IV: A prospective cross-sectional design was used to evaluate the psychometric properties of the KOOS-Child. One hundred and fifteen children with knee disorders were recruited from primary and secondary care clinics in the Stockholm area between 2012-2013. All children completed the KOOS-Child, Child Health Assessment Questionnaire (CHAQ) [125] and EQ-5D-Y [79,126] at baseline to evaluate construct validity. Two additional administrations were performed for analyses of reliability and responsiveness. An anchor-based approach was used to evaluate responsiveness and interpretability.

Table I. Summary of aims, design, study period, participants, outcome measures, and data analysis used in the thesis.

Study	Aim	Design	Study period (data collection)	Participants	Outcome measures	Data analysis
I	Evaluate the reliability of strength measurements when used in healthy children and adults in the SMC, and to assess agreement between the SMC and an ID.	Test-retest design	2005-2006	n = 43 20 healthy children (5-13 years) 23 healthy adults (23-60 years)	Strength Measuring Chair Isokinetic dynamometer (Biodex System 3)	Descriptive statistics $ICC_{1,1}$ and $ICC_{2,1}$ SEM SDC Paired t-test Bland Altman 95% limits of agreement plots
II	Evaluate the reliability of two tests of knee alignment medio-lateral knee position, present pediatric reference values, and evaluate the association between the tests.	Test-retest design	2007-2008	n=253 healthy children (9-16 years)	Single-limb mini squat test Quadriceps angle	Descriptive statistics Three-way ANOVA Limb Symmetry Index Cohen's Kappa Percentage Agreement Sign test $ICC_{2,1}$ SEM Repeatability coefficient Univariate linear regression
III	Examine the comprehensibility of the KOOS in children and to recommend modifications for a pediatric version (KOOS-Child).	Qualitative approach	2009-2010	n=38 children with knee disorders (10-16 years)	Cognitive interviews KOOS	Descriptive statistics Qualitative analysis (problem detection, counts of difficult survey items, sorting of issues into categories suggestions for modifications) Sub-group analysis
IV	Evaluate the psychometric properties of the KOOS-Child when used in children with knee disorders.	Prospective cross-sectional design	2012-2013	n=115 children with knee disorders (7-16 years)	KOOS-Child CHAQ EQ-5D-Y GPE scores VAS (knee function)	Descriptive statistics Floor and ceiling effects $ICC_{2,1}$ SEM SDC Exploratory factor analysis Cronbach's alpha Spearman correlation coefficient Effect Size Standardized Response Mean Mean change method (MIC)

Abbreviations: KOOS - Knee Injury and Osteoarthritis Outcome Score, KOOS-Child - Knee Injury and Osteoarthritis Outcome Score for children, ICC - Intraclass Correlation Coefficient, SEM - Standard Error of Measurement, SDC - Smallest Detectable Change, MIC –Minimal Important Change

3.2 PARTICIPANTS AND DATA COLLECTION

A total sample of 449 subjects participated in the present thesis. Two studies (Study I and Study II) included healthy control subjects (n=296), and the other two studies (Study III and Study IV) included children with various knee disorders (n=153) (Table II).

Table II. Overview of participants in Studies I-IV. Details of patient characteristics can be found in each paper.

Study	Participants	Sex (boys/girls)	Age (min-max)
I	20 healthy children	10/10	5-13
	23 healthy adults	10/13 (men/women)	23-60
II	253 healthy children	126/127	9-16
III	38 children with knee disorders	18/20	10-16
IV	115 children with knee disorders	51/64	7-16

STUDY I

In Study I, a convenience sample of 20 healthy children and 23 healthy adults were included. Criteria for exclusion were current or recent injuries in the lower extremities at time of testing. Four subjects (two adults and two children) were excluded from the analyses due to knee pain experienced during testing in the SMC or ID. Five children could not be tested in the ID due to small body proportions. These data were not included in the analysis. Subjects participating in all three sessions were included in analysis, and the final numbers in the analysis were calculated for: knee flexors in adults, n=21; knee extensors in adults, n=22; knee flexors in children, n=15; knee extensors in children, n=16.

Subjects were evaluated by the same examiner (MÖ) and attended three sessions with muscle strength measurements: two sessions in the SMC (at the Motion Analysis Laboratory, Astrid Lindgren's Hospital, Stockholm, Sweden) and one session in the ID (at the Department of Orthopedics, Karolinska University Hospital, Solna, Sweden), with one-week interval between sessions. The sessions lasted about one hour each. First, the subjects were tested in the SMC, second in the ID to evaluate agreement, followed by the SMC again to evaluate test-retest reliability of the SMC. Isometric torque in the knee extensors and knee flexors were measured in both instruments (three repetitions/muscle group). During testing, subjects were seated with their knee at a 60° angle, hips at 90° angle, and arms across their chests. An adjustable back support was available. Standardized verbal encouragement was used during the tests.

STUDY II

In Study II, a total of 253 healthy children were recruited from three schools in Stockholm, Sweden. Seven children were excluded due to self-reported earlier surgical treatment of the lower extremities, leaving 246 healthy children (9–16 years) included in the final analysis. All children tested twice or by two examiners were included in the reliability analyses (intra-/inter-rater reliability for Q-angle n = 37/85 and for Single-limb mini squat test n = 33/28). An overview of included subjects in the various analyses is presented in Fig. 2. During analysis, subjects were divided into groups depending on age and sex (9–11, 12–14 and 15–16 years). The legs were tested bilaterally for both tests, however, only data collected from the dominant leg were used for reliability analyses.

Data collection took place in conjunction with the children's physical education classes in school. Children were assessed at two different sessions, approximately one week apart. Evaluation of dynamic medio-lateral knee position was performed during the Single-limb mini squat test and the static medio-lateral knee position by the Q-angle. Two physiotherapists assessed the Single-limb mini squat test and one physiotherapist and one orthopedic surgeon measured the Q-angle. Examiners were blinded to each other's measurements during the two tests. An inter-rater training session of measurement was performed for both measurements prior to data collection. The evaluation time lasted approximately 5-10 minutes per session for each subject.

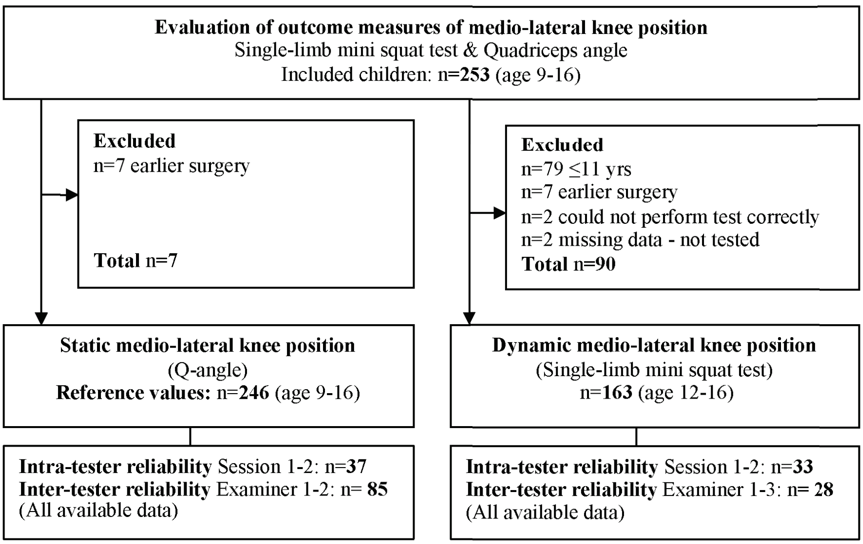


Figure 2. Overview of subjects included in the various analyses in Study II

STUDY III

In Study III, a total of 38 children between 10-16 years with primary or repeated knee disorders were enrolled. Children were recruited from primary and secondary care clinics in Stockholm (eight physiotherapy clinics and the orthopedic clinic at Astrid Lindgren children's hospital). Children were purposefully selected to address subject factors related to injury, sex and age. Inclusion criteria were: unilateral knee injury, Swedish as native language, and first time to be exposed to the KOOS. During the interview session, the researcher noted that two of the children were foreign speaking and two had previously been exposed to KOOS. These children's data were excluded, leaving 34 children for final analysis. The children presented with a mix of knee diagnoses. An initial number of subjects were recruited in each age and sex category as a start, and then the sampling continued until saturation was achieved. Children were also divided into groups based on age and sex (10–12, 13–14, and 15–16) to obtain an approximate equal representation in age/sex groups (Table III).

Table III. Subjects divided in different age groups (n=34)

Group category	**n**	**Age** mean (SD)	**Sex** (boys/girls)	**Diagnoses represented**
1 (10-12 yrs)	14	11.4 (0.8)	7/7	ACL injury, Anterior knee pain, Knee sprain, Meniscal injury, Patellar dislocation, Patellar fracture, PCL injury
2 (13-14 yrs)	9	13.4 (0.5)	4/5	ACL injury, Meniscal injury, Medial patellar plica, Osteochondritis dissecans, PCL injury, Patellar dislocation
3 (15-16 yrs)	11	15.5 (0.5)	6/5	ACL injury, Lateral collateral ligament injury, Meniscal injury, Unspecified knee pain

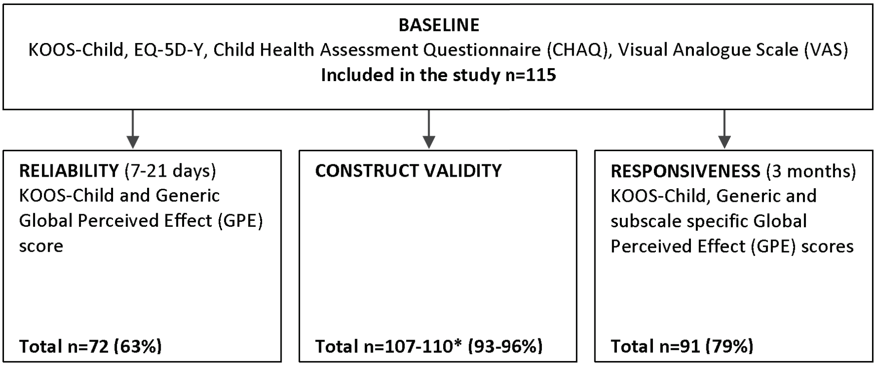
Abbreviations: ACL – Anterior Cruciate Ligament, PCL – Posterior Cruciate Ligament

Cognitive Interviews were used to assess the comprehensibility of the KOOS [103,120,121]. A behavioral scientist (MDI) developed the interview guide and structured training sessions for the interviewer (MÖ). Training sessions were performed in Boston, MA, USA with 6 American-English speaking children age 12-16 prior to the study.

The real interviewing session lasted about 1-1.5 hours. During the interviews, parents were often present, but were advised not to interfere during the interviews. Each child was asked to read the KOOS directions aloud and restate the directions to each section in his/her own words. When answering the survey, the child was instructed to read each question and to circle an item if he or she thought that it was too difficult. During the interview the researcher observed and made notes. The notes and observations made (e.g., hesitations, skipping pages or flipping pages back and forth) were used to implement specific retrospective probes for the child *(supplementary data, Paper III)*. When the child had completed the form, the interviewer discussed the child's interpretation of the form, using general and specific verbal probes. In other words, the interviewer requested the child to rephrase questions, define meanings of words, and explain their responses to identify difficulties in understanding, interpretation or how to fill out the form [120,121,127,128]. All interviews were recorded and transcribed.

STUDY IV

In Study IV, a total of 115 children, 7-16 years of age, with various knee disorders were recruited. As in Study III, patients were recruited from primary and secondary care clinics in Stockholm. Patients were screened for eligibility according to a screening protocol and were included if they fulfilled the following criteria: 1) seeking medical care and/or receiving treatment because of acute or chronic knee-related problems; 2) restricted in physical activities; 3) able to read and complete the forms by themselves; 4) Swedish as native language; and 5) no co-morbidities overriding knee symptoms. A flow chart of the patients included for the different analyses is presented in Fig. 3.



Note. * Number of participants depends on the total number answering both the KOOS-Child and the comparator instruments

Figure 3. Flowchart of subjects participating in Study IV

All children seeking medical care at either their doctor or physiotherapist that fulfilled the inclusion criteria were invited to participate in the study. If they agreed, they were asked to complete the preliminary KOOS-Child developed in Study III, the CHAQ [125], the EQ-5D-Y [79,126], and five subscale-specific VAS formatted questions (specifically designed for this study) to determine construct validity. After 1-3 weeks, and then 3 months after the first session the children were mailed the preliminary KOOS-Child and asked to complete the survey. These repeat administrations were used to evaluate reliability and responsiveness, respectively. At the second and third administrations, children rated their degree of change in their knee condition since last administration using one generic and five subscale-specific Global Perceived Effect (GPE) scores. These GPE scores served as external criteria for clinically relevant change [107]. The seven item GPE-scoring scale ranged from "much worse" to "much better." Children took part in a variety of interventions or in no intervention during the 3-month follow-up period, thus no specific interventions were applied. If needed, the children were reminded once by telephone or e-mail to complete the forms. Children who rated their condition on the GPE score as "somewhat better", "no change" or "somewhat worse" were considered as "stable in knee condition" throughout the analyses.

3.3 EVALUATION METHODS

The methods for data collection used are presented in the following section. They are also summarized in Table IV.

Table IV. Summary of the different outcome measures used in Study I-IV

Outcome measures	Study I	Study II	Study III	Study IV
Strength Measuring Chair (SMC)	X			
Isokinetic Dynamometer (Biodex System 3)	X			
Single-limb mini squat test		X		
Quadriceps-angle		X		
KOOS			X	
KOOS-Child LK1.0 (preliminary version)			X	X
KOOS-Child LK2.0 (final version)				X
CHAQ (shorted version)				X
EQ-5D-Y (shorted version)				X
VAS scores – knee function				X
GPE-scores (generic and specific)				X

Abbreviations: CHAQ – Child Health Assessment Questionnaire, EQ-5D-Y – EQ-5D-Youth version, VAS – Visual Analogue Scale scores, GPE – Global Perceived Effect - scores

3.3.1 Clinical outcome measures

The Strength Measuring Chair
The Strength Measuring Chair (SMC) was designed at the Motion Analysis Laboratory, Astrid Lindgren Children's Hospital, Stockholm, Sweden to measure isometric muscle strength in the lower extremity while the subject was seated [65]. In the SMC, subjects are always tested with hip and ankle joints in 90° positions. The knee joint is adjustable from 90° to 30° positions. The chair is adjustable in size to suit young subjects (approximately 5 years of age) and adults. Force sensors are placed at adjustable distances from joint centers and force generation in tension and compression is measured (Fig. 4). The same sensors can be used to test opposing muscle groups. For example, pushing on a sensor allows for knee flexor strength measurements, while knee extensor strength can be tested by pulling on a strap attached to the same sensor. The instrument is able to measure force generation in 4 different muscle groups: plantarflexors, ankle dorsiflexors, knee flexors, and knee extensors on both sides simultaneously.

In Study I, the focus was solely on knee flexors and extensors. To restrict motion during muscle strength testing, three straps were attached to the subjects: one around the waist, one over the thigh, and one around the lower shank. The moment arm was measured from the lateral side of the knee joint center and to the sensor's middle along the sensor plate axis. The voltage signal from the force sensors was converted to a digital signal and sent to a personal computer (PC). The instrument measured knee muscle strength as the product of the compression force (in Newtons) of the sensor and the moment arm (in meters). A PC user interface implemented moment arms and displayed the generated torque in newton-meters (Nm). The data was collected and exported to Microsoft Excel. Taring of the sensors was implemented to offset the limb weight on the sensor; in other words, the subject's limb was weighed prior to testing to correct for the influence of gravity on the data. Data were sampled at 1000Hz and analyzed as a mean over a 10-second interval. The sensors of the SMC were calibrated shortly before Study I began [65].

Figure 4. The Strength Measuring Chair (SMC)

Isokinetic Dynamometer
The Isokinetic Dynamometer (ID), the Biodex System 3 dynamometer® (Biodex Medical Systems, NY, www.biodex.com) was used in Study I to evaluate the agreement between the ID and the SMC. The ID has a central torque-measuring device to measure muscle strength and can measure both isokinetic and isometric muscle torque (Nm). To compare the instruments the isometric mode of the ID was applied. In the ID, subjects were tested in the same measuring positions as in the SMC. The instrument was originally designed for adults and in cases where the seat was too large for the smallest subjects a pillow was used to compensate for the large seat depth. As in the SMC, the subject's limb was weighed prior to testing to correct for the influence of gravity on the data.

The Quadriceps-angle
The Quadriceps (Q-angle) was used to determine the static medio-lateral knee position in Study II [40]. The Q-angle (in degrees) was measured with one arm from anterior superior iliac spine of the pelvic bone to the center of the patella, and the other arm from the center of the patella to the tibial tubercle on the tibia. Subjects were in supine position with the quadriceps relaxed, the hip and knee joint extended, and with leg and foot in a neutral position during measurements (Fig. 5) [129]. The angle was measured with a standardized goniometer with extended arms, modified for the present study (Olmed Ltd) with the purpose of achieving the most possible accurate alignment from the measuring points. A disagreement regarding the validity and reliability of the Q-angle can be found in the literature [68,72,75].

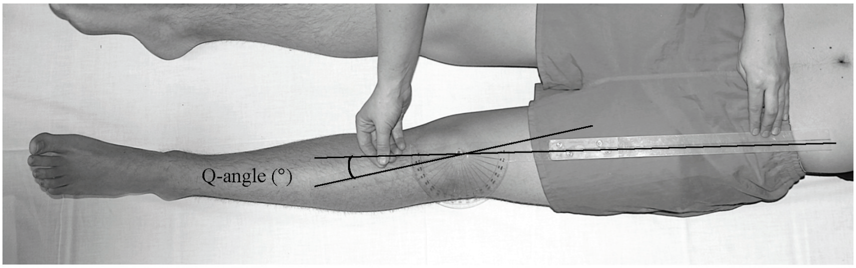


Figure 5. Quadriceps-angle (Q-angle)

Single-limb mini squat test

In Study II, the Single-limb mini squat test was used as a discriminative test of medio-lateral knee position (Fig. 6). The test is a functional dynamic test that evaluates two aspects of knee function, that are measured simultaneously. The qualitative aspect of knee motion entails observation of movements during test performance (assessment of "knee medial to foot" position). The quantitative aspect involves measuring the maximum number of knee bendings on one leg in 30 seconds (as a measure of knee function) [39,51]. In Study II, the focus was on the qualitative aspect and assessment of the knee position in relation to the ankle joint, even though the number of knee bendings was registered simultaneously. A positive outcome of the test was defined as the knee observed to be medial to the foot (second toe) any time during the 30-second performance of the test.

The testing procedure used was as follows: the subject stood on one leg with fingertip support for balance and to direct the hip to decrease the risk of rotation from hip and trunk. The foot was placed on a T-shaped mark taped to the floor. The foot was aligned to the stem of the "T" by placing the second toe on the stem. The subject was asked to bend his/her knee without bending the trunk forward from the hip, until he/she could not see the foot and then return to extended position. This position corresponds to approximately 50° knee flexion. The instructions were to perform as many knee flexions as possible during 30 seconds. The subjects were not aware of the observation of knee position during the test.

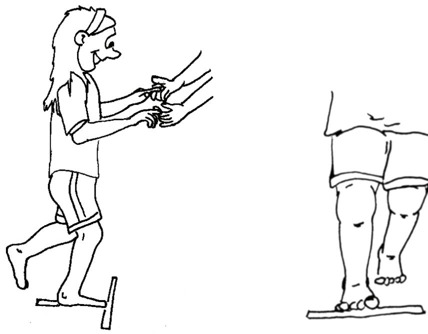


Figure 6. Single-limb mini squat test (sagittal and frontal plane)

Patient-Reported Outcome Measures
Several PRO measures were used in Study III and Study IV (Table IV). Instrument characteristics and psychometric properties are described below.

Knee Injury and Osteoarthritis Outcome Score (KOOS)
The KOOS (Study III) is a self-administered disease-specific questionnaire designed to assess patient-reported knee outcomes among adults with joint injury or degenerative disease [80,84,85]. The KOOS contains 42 items divided into five domains; pain, symptoms, activities of daily living (ADL), sports/recreation (Sport/Rec), as well as knee-related quality of life (QoL), and presents separate subscale scores as a profile (*original KOOS items are presented in Paper III*). KOOS is widely used, and has shown to be highly reliable, valid, and responsive to change when used in adult subjects with various knee disorders in several studies [80,84,85] (www.koos.nu).

The time period considered when answering the KOOS is the previous week. Standardized answer options are given on a Likert scale (5 boxes) [77] and each question is assigned a score from 0 to 4. The score is then normalized to a 0-100 scale for each subscale (100 indicating no symptoms and 0 indicating extreme symptoms). Scores between 0 and 100 represent the percentage of total possible score achieved. The KOOS takes about 10 minutes to complete.

KOOS-Child
KOOS-Child (Studies III and IV) is a modified version of the KOOS, developed as an instrument to assess the pediatric patient's self-reported knee problems (Study III). The development process consisted of two steps. In Study III the comprehensibility of the KOOS was evaluated and the preliminary version of the KOOS-Child (KOOS-Child LK1.0) was developed *(supplementary data, Paper III)*. This version was tested and further modified in Study IV, where the KOOS-Child showed to be reliable, valid and responsive to change when used in a pediatric population aged 7-16 years with knee disorders. The structure of the KOOS-Child subscales are the same as in the original KOOS, however 4 new items have been added and 9 items have been deleted, leaving 39 items in the final version of the KOOS-Child (KOOS-Child LK2.0) *(appendix).* The KOOS-Child takes about 10 - 20 minutes to complete depending on age and reading skills. Scoring of the KOOS-Child is similar to the original KOOS and detailed scoring information is available in the user's guide *(appendix).*

EQ-5D-Youth version (EQ-5D-Y)
The EQ-5D is a broadly used generic questionnaire originally designed to measure health-related quality of life in adults [78]. A lack of comprehension of the EQ-5D when used in children has been shown. Therefore the adult version has been modified to suit a pediatric population (EQ-5D-Y) from the age of 8 years [79,126]. The EQ-5D-Y consists of five domains: mobility, self-care, usual activities, pain and discomfort, and anxiety and depression. Each item has three response categories, "no problems", "some problems" and "a lot of problems". No overall score can be calculated, which means that data are presented as the percentage of individuals reporting each level of problem for each item [79,126]. The EQ-5D-Y also includes a VAS score for rating overall health

status; however, this was not included in Study IV *(appendix)*. The EQ-5D-Y has shown good psychometric properties when used in healthy children and children with various types of chronic health conditions [126].

Child-Health Assessment Questionnaire (CHAQ)

The CHAQ is an instrument measuring self-assessed disabilities in children with inflammatory joint diseases. It has shown to be a reliable and valid tool for functional, physical and psychosocial assessment in children with Juvenile idiopathic arthritis [125]. It contains 8 domains that evaluate the degree of difficulty in performing activities of daily living, each of them scored on a four-point ordinal scale; 0 (without any difficulty), 1 (with some difficulty), 2 (with much difficulty), and 3 (unable to do). The eight domains together gives a disability index that is also complemented by two VAS scores, one for pain and one for rating of overall well-being. In Study IV, a shortened version of the CHAQ was used, including the domains "arising," "walking," "grip," and one item from the "activity" domain as well as one VAS score, "pain" *(appendix).* The domains used in Study IV were scored separately.

Visual Analogue Scale scores - of knee function

Visual analogue scales (VAS) are used to measure a variety of subjective phenomena. With lack of existing pediatric PROs measuring knee function, five subscale-specific VAS scores were specifically designed in Study IV to serve as a comparable instrument to the different subscales of the KOOS-Child *(appendix)*. All children were asked to make a mark on a 0-100 mm horizontal line (no problems to worst possible problems). The use of VAS scores has been shown to be valid when measuring knee function in adults [84], however its psychometric properties when used in a pediatric population with knee disorders has not been evaluated.

Global Perceived Effect (GPE) scores – an external criteria of health status

In Study IV all children were asked to rate their change in knee condition between the different administrations (between baseline and first follow-up at 1-3 weeks, and between baseline and 3 months follow-up), using one generic and five subscale-specific Global Perceived Effect (GPE) scores designed for this study *(appendix).* They were used as anchor/external criteria for a clinically relevant change [107]. The 7-item GPE scoring scale ranged from "much worse" to "much better". GPE scores are widely used, and have been shown to have excellent test-retest reliability. However, the use of GPE scores has been questioned as they tend to be influenced by current status, with the effect more obvious with longer periods of recall [107-110,130].

3.4 STATISTICAL METHODS AND DATA ANALYSIS

Statistical methods and data analyses used are presented in detail below. A summary of the statistical approaches can be found in Table V. All statistical analyses were performed using Statistica Software (Studies I, II, IV), and SPSS ver.19 (Study IV), and SPSS ver.21 (Study I).

Table V. Overview of statistical methods used in Studies I-IV.

Statistical method	Study I	Study II	Study III	Study IV
Descriptive statistics	X	X	X	X
Intraclass Correlation Coefficient, (ICC)	X	X		X
Standard Error of Measurement (SEM)	X	X		X
Smallest Detectable Change (SDC)/ Repeatability coefficient	X	X		X
Bland Altman 95% limits of agreement plots	X			
Paired t-test	X			
Three-way ANOVA		X		
Limb Symmetry Index		X		
Cohen's Kappa		X		
Percentage Agreement		X		
Sign test		X		
Univariate linear regression analysis		X		
Qualitative analysis			X	
Exploratory Factor Analysis				X
Cronbach's alpha				X
Spearman correlation coefficient				X
Effect Size				X
Standardized Response Mean				X
Mean change method				X

3.4.1 Descriptive statistics (Study I-IV)

Data types obtained in the different studies are ratio/interval, ordinal, and categorical data. Torque (Nm) and Q-angle (degree) values were ratio data, whereas data from the Single-limb mini squat test (medial or neutral/lateral knee position) were categorical, and the data from questionnaires were ordinal. Thus both non-parametric and parametric statistics were used.

In Study I, muscle strength was presented as the average values of 3 peak torque measurements in Nm complemented by standard deviations at 3 different sessions in the SMC, and the ID for subjects attending all 3 sessions.

In Study II, reference values for the Q-angle (from Session 1, examiner MÖ) were presented as degrees, and for the Single-limb mini squat test the amount of subjects classified as "knee medial to foot" were presented. During the Single-limb mini squat test, two subjects were classified as "knee lateral to foot". They were included in the "knee neutral to foot" group during the analysis. In the present thesis, the number of knee bendings was also presented as number of knee bendings/30 seconds.

In Study III, interview results were presented as general observations. Item-specific comments were presented as key examples, and as the numbers of reported issues.

In Study IV, the raw data from the KOOS-Child are ordinal, which implies the use of non-parametric statistics; however, it is common that data from questionnaires are presented as mean and SD, such as in the original version of the KOOS [80,85]. Thus to be able to compare our data and the distribution of KOOS-Child scores, the data were presented as both median (min-max) and mean (SD) scores.

3.4.2 Study I and II

Reliability

The Intraclass Correlation Coefficient (ICC) is an index designed to evaluate reliability [97]. The ICC is calculated from the mean square values derived from a repeated-measures ANOVA. The ICC is dependent upon the heterogeneity of the group, and will be high if the variance of measurements between subjects is higher than that within subjects. There is no standard acceptable level of reliability. ICC ranges from 0 (low) – 1 (high reliability) [92]. In Study I, the ICC values were interpreted according the recommendations by Fleiss et al. [131], wherein an ICC > 0.75 represents excellent reliability, and 0.4-0.75 as fair to good reliability [131] whereas in Study II by Landis and Koch[98] (0.80-1.00 almost perfect agreement, 0.61-0.80 substantial agreement, 0.41-0.60 moderate agreement, 0.21-0.40 fair agreement and 0-0.20 poor agreement)[98]. In Study I, the ICC values were complemented with a 95% Confidence Interval (CI). In Study II, (Single-limb mini squat test data) the use of Cohen's Kappa and percentage of agreement were used to evaluate the reliability. The results were interpreted according to Landis and Koch [98].

In study I, a Bland Altman 95% limit of agreement plot was used to detect systematic variations and trends [93]. The plots were used to complement the ICC and to check for systematic differences. In the plot, the difference in muscle strength results (mean peak torque values) between Session 2 (ID) and Session 3 (SMC) were plotted against the average of the two means from sessions 2 and 3 for each subject.

In Study I, a paired t-test was also used to check for significant differences between the sessions (peak torque values). In Study II, an ANOVA was used to test whether there was a systematic difference between examiners and test occasions. In both studies the significant level was set to $p<0.05$.

ICC gives no information about size of disagreement between measurements and should be complemented by the Standard Error of Measurement (SEM). The SEM is a measure of the precision of a test instrument. The SEM was calculated as √mean square – within targets) [92,99]. To determine the clinical change apart from the measurement error the Smallest Detectable Change (SDC)/repeatability coefficient was calculated (SEM*1.96*√2) [117]. The SDC and the repeatability coefficient is the same; however, different terminology is used in the literature [108]. The SDC can be interpreted as, "the magnitude of change below which there is more than a 95% chance that no real change occurred" [118].

Differences between legs, sex and age groups
In Study II, a three-way ANOVA with one within-groups factor side (dominant/non-dominant leg) and two between-groups factors for sex and age groups to evaluate any possible differences in Q-angle between dominant/non-dominant leg, sex, and age groups was calculated. Possible leg asymmetry for the Q-angle was evaluated using a Limb Symmetry Index: highest Q-angle value / lowest Q-angle value * 100. Since the Single-limb mini squat test outcome data were categorical a Sign test was used to analyze possible leg asymmetry.

Association between tests
In Study II, the evaluation of possible association between the Single-limb mini squat test and the Q-angle were evaluated by a univariate linear regression analysis.

3.4.3 Study III and IV

Comprehensibility and content validity
In Study III, cognitive interview transcript data were analyzed by the research team to evaluate the comprehensibility of the KOOS to ensure good content validity of the KOOS-Child. The 'main researcher' (MÖ) extracted all comments from the transcripts. Themes were created separately by the team. Thereafter, themes and issues that arose from the interviews were discussed by the team until an agreement was reached. Triangulation was also made; in other words, the other researchers of the group made the same extracting process of four random transcripts to ensure validity. Analysis focused on problem detection such as counts of difficult items, identifying reactions and similarities in paraphrasing of questions as well as suggestions for modifications [103,122-124]. Problematic issues were also sorted into the following categories: comprehension, item format, response format, or mapping issues. To illustrate the areas of difficulties and lack of comprehension, as well as for revising items, quotes were extracted from the interviews. Modifications to the questionnaire were then made based on the results from the analysis [88,103,123]. Subgroup analysis between age, sex, and treatment were made to evaluate if certain issues were more common among specific groups. The KOOS was modified simultaneously in Swedish and English using the same process of analyses.

Confirmation of subscales and item reduction
In Study IV, predefined criteria for possible deletion of an item from the preliminary version of the KOOS-Child were: A ceiling effect of 70% or above and/or an ICC of less than 0.41, and/or >5% missing responses [98]. First and second administrations of the preliminary KOOS-Child were used for this purpose.

Factor analysis is used to evaluate whether the items of a questionnaire representing different dimensions of the construct measured could be sorted into different subscales. Failure to load on a single major factor suggests that the items do not measure the same construct [77,94,133,134]. An Exploratory Factor Analysis (EFA) using principal component analysis with varimax rotation was performed for each KOOS-Child subscale separately in order to evaluate subscale unidimensionality [134]. An index of the amount of variance accounted for by each factor, Eigenvalue above 1, was set as the limit [77].

Reliability
An ICC was calculated to evaluate the test-retest reliability for all final KOOS-Child subscales. ICC was interpreted according to recommendations by Landis and Koch [98]. ICC values were complemented with a 95% CI. As in Study I and II, the ICCs were also complemented with both SEM and SDC to evaluate the measurement error. SDC was calculated at both group (SEM*1.96*√2*√n) and individual level (SEM*1.96*√2) [117,132]. Cronbach's alpha was calculated for each subscale of the KOOS-Child to evaluate internal consistency (inter-item correlation) for all KOOS-Child subscales [96]. A score >0.70 was set as sufficient item homogeneity [113].

Construct validity
Correlations between the final KOOS-Child subscales and the other questionnaires (CHAQ, EQ-5D-Y and VAS scores of knee function) were evaluated using a Spearman Correlation coefficient. Correlation coefficients > 0.5 were considered strong, 0.35-0.5 moderate and < 0.35 as weak [135]. Predefined hypotheses were established regarding expected correlations between similar and dissimilar constructs. Construct validity were defined to be good if 75% of the hypothesis were confirmed [113].

Responsiveness
Spearman correlation coefficients were used to evaluate the correlation of change scores between the KOOS-Child and the specific GPE scores. A predefined hypothesis was set for the correlations at a threshold of at least 0.3 [136]. Effect Size (baseline score – follow-up score / SD of baseline) and the Standardized Response Mean (baseline score – follow-up score / SD change in score) were also calculated. Effect sizes of 0.2, 0.5, and 0.8 were interpreted as small, medium or large respectively [111].

Interpretability
Interpretability of the KOOS-Child was evaluated by presenting the distribution of total/change scores in the study sample, floor/ceiling effects and evaluation of the Minimal Important Change (MIC). A floor or ceiling effect was considered present if >15% of the respondents reported the lowest/worst (0) or the highest/best score (100) for a subscale[113]. The MIC was determined with an anchor-based method, employing the subscale specific GPE scores as the global index of change [113,116]. The mean change method was used [107]. A clinically meaningful change was defined as a mean change score between baseline and 3-months follow-up for those patients reporting to be "better" or "much better".

3.5 ETHICAL CONSIDERATIONS

All studies were approved by the regional ethical board of Stockholm and conducted according to the Declaration of Helsinki. Participation was voluntary and all subjects were given both oral and written information about the studies. Written consent to participate in the study was obtained from all adults, and by parent or legal guardian where applicable in the pediatric population.

4 SUMMARY OF RESULTS

4.1 EVALUATION OF AN INSTRUMENT FOR MEASURING KNEE MUSCLE STRENGTH

The overall aims of Study I were to evaluate the reliability of isometric knee extensor and knee flexor muscle strength measurements using the Strength Measuring Chair (SMC) and to evaluate the agreement between the SMC and a commonly used Isokinetic Dynamometer (ID) in healthy children and adults.

The results showed a significant difference in muscle strength measurements between the SMC and the ID. The mean torque values were consistently smaller in the SMC than in the ID; in adults 19% lower knee flexor, and 16% lower knee extensor muscle strength was found, and in children, 65% lower knee flexor, and 63% lower knee extensor muscle strength was found (Table VI).

Table VI. Muscle strength, computed as the average of 3 peak torque measurements (Nm) from sessions in the Strength Measuring Chair (SMC) and in the Isokinetic Dynamometer (ID).

	n	Session 1 (SMC)	Session 2 (ID)	Session 3 (SMC)	P value*
Adults/ Knee flexors	21	75.3 ± 26.5	92.8 ± 35.1	75.2 ± 32.3	<0.05
Adults/ Knee extensors	22	147.0 ± 56.9	174.7 ± 63.8	143.2 ± 56.1	<0.05
Children/ Knee flexors	15	12.4 ± 4.5	34.4 ± 17.7	12.3 ± 4.9	<0.05
Children/ Knee extensors	16	20.9 ± 9.1	56.8 ± 32.5	18.6 ± 8.5	<0.05

Note: * Significant p-values represent the difference between Session 2 (ID) and Session 3 (SMC). Non-significant differences were found between Sessions 1 and 3 in the SMC. N denotes the number of subjects measured at all of Sessions 1 - 3. Data analyzed from trials with a 60º knee angle (mean ± SD).

4.1.1 Reliability

An excellent intra-subject reliability (ICC 0.93-0.99) was observed in both groups for knee extensor and knee flexor muscle strength measurements from both instruments. Test-retest reliability was evaluated for the SMC, and was found to be excellent according to the ICC values in both the adult and pediatric groups (Adults: ICC 0.87-0.92, SEM 10.83-17.33, SDC 30.02-48.05 Nm; Children: ICC 0.90-0.93, SEM 1.63-2.89, SDC 4.51-8.01). However, a significant difference was observed in knee extensor strength measurements in the pediatric group (p=0.03), wherein the SMC measured slightly greater muscle strength in Session 1.

4.1.2 Instrument agreement

A systematic difference was found in both groups between the two instruments, with more differences noted in the pediatric group (Adults: ICC 0.72-0.85, SEM 14.49-28.45, $p<0.05$; Children: ICC 0.17-0.23, SEM 19.78-33.25, $p<0.05$). Bland Altman 95% Limits of Agreement plots were used to show possible systematic variation and trends in the groups and between instruments. The graphs illustrated that the ID systematically recorded greater muscle strength values than the SMC. The systematic difference was nearly consistent in adults, but increased proportionally with muscle strength in children (Fig. 7 A-D). Two groups can be distinguished in each graph: in Fig. 7 A, women and men, and in Fig. 7 C, younger and older children.

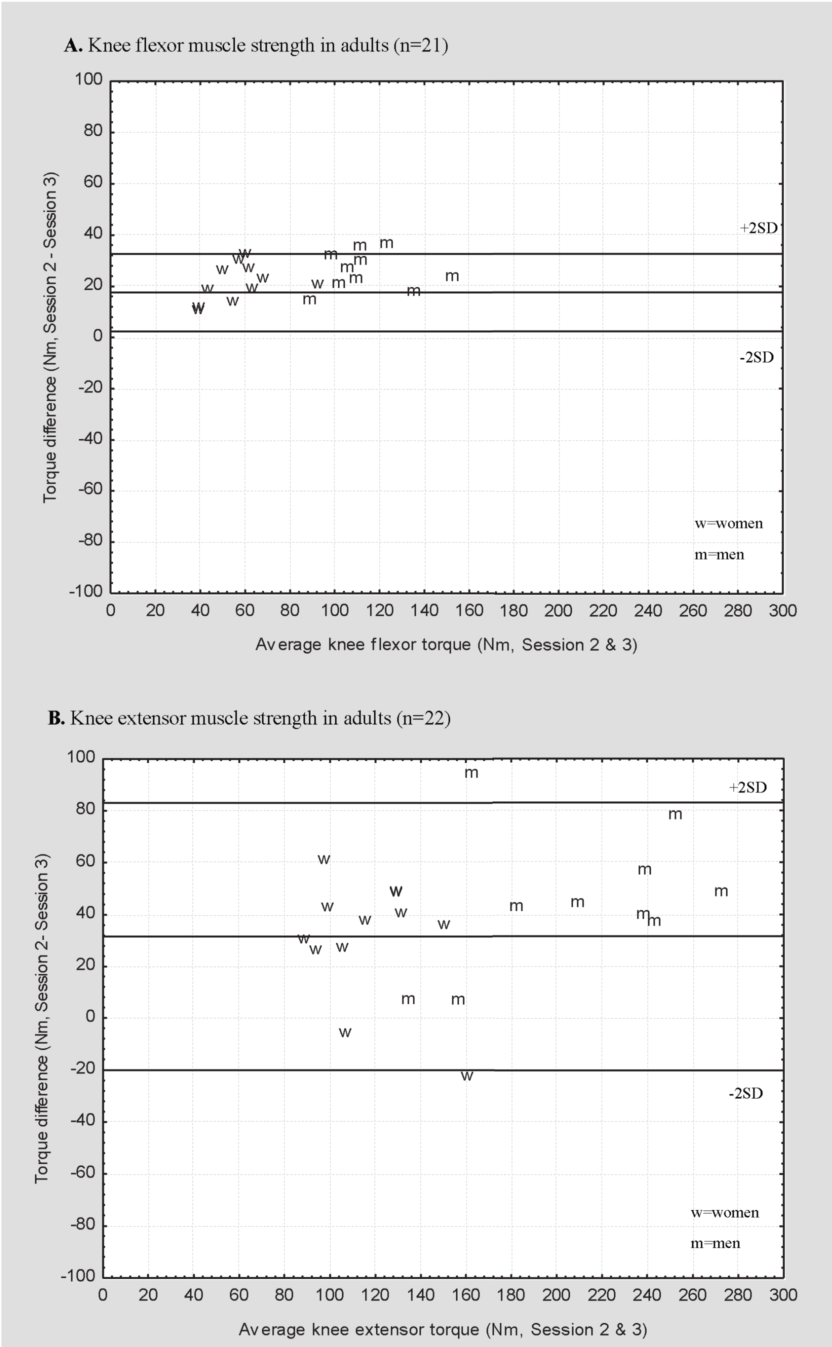


Figure 7 A-B. Bland Altman plots[93] showing the differences in mean peak torque plotted against the average of the two mean torques from Session 2 (Isokinetic Dynamometer) and Session 3 (Strength Measuring Chair) for each subject.

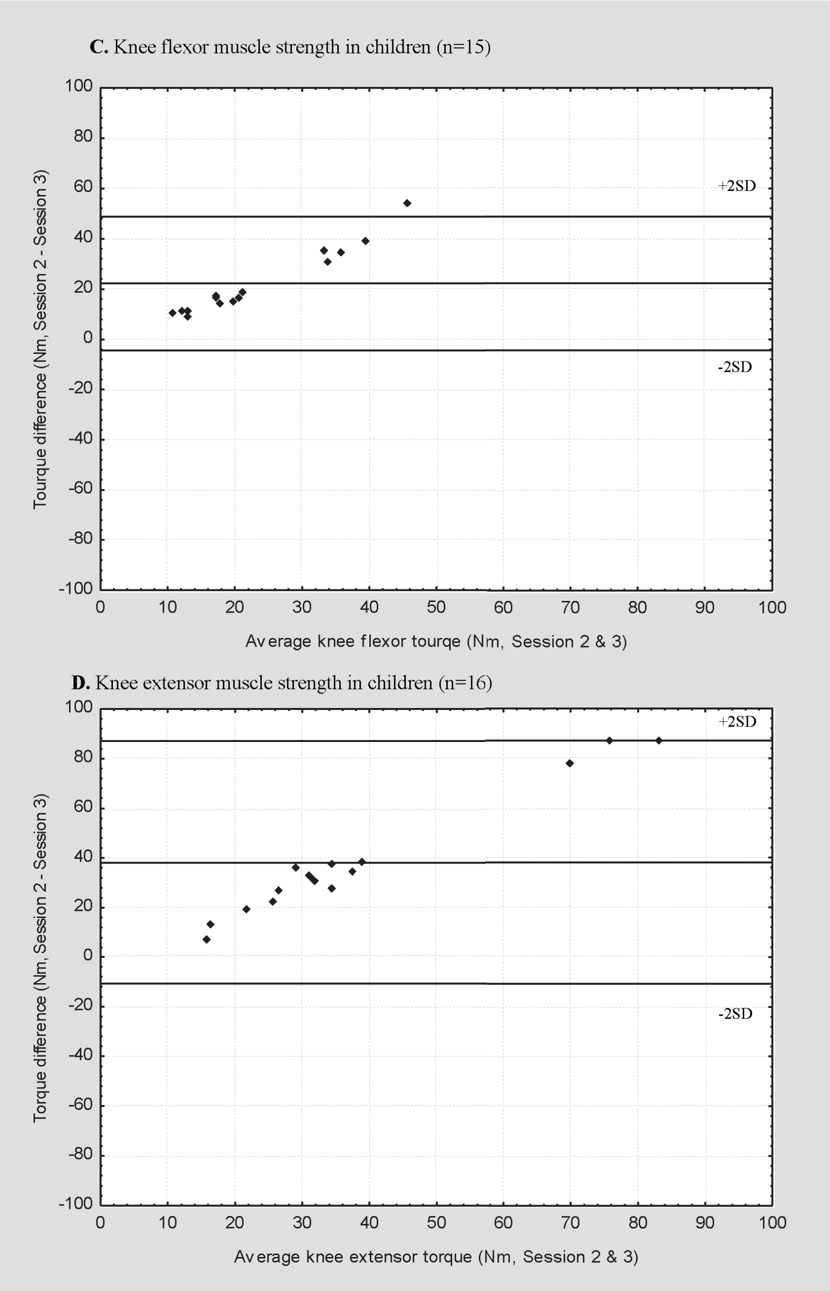


Figure 7 C-D. Bland Altman plots[93] showing the differences in mean peak torque plotted against the average of the two mean torques from Session 2 (Isokinetic Dynamometer) and Session 3 (Strength Measuring Chair) for each subject.

4.2 EVALUATION OF TWO CLINICAL MEASUREMENTS OF KNEE POSITION

The overall aim of Study II was to evaluate two clinical tests, the Single-limb mini squat test and the Q-angle, as discriminative tests of medio-lateral knee position, with respect to reliability, reference values and the association between the two tests.

4.2.1 Single-limb mini squat test

The outcome of the Single-limb mini squat test is twofold: the qualitative observation of knee position as well as the quantitative aspect of number of knee bendings/30s. In Study II, both quantitative and qualitative data were collected. However, since the aim was to evaluate reliability when evaluating the medio-lateral knee position during performance, the focus of paper II was solely on the qualitative aspect. In Table VII, the mean values of quantitative data are presented as a complement. 36% of the subjects were classified as “knee medial to foot” (examiner MÖ, session 1). No side-to-side differences (dominant/non-dominant leg) could be seen (p=0.27).

During assessment of dynamic medio-lateral knee position in healthy children a moderate reliability was noted with the Single-limb mini squat test (Intra-rater: kappa 0.48, 95% CI 0.16–0.79, 76% perfect agreement; Inter-rater: kappa 0.57, 95% CI 0.30-0.85, 79% perfect agreement).

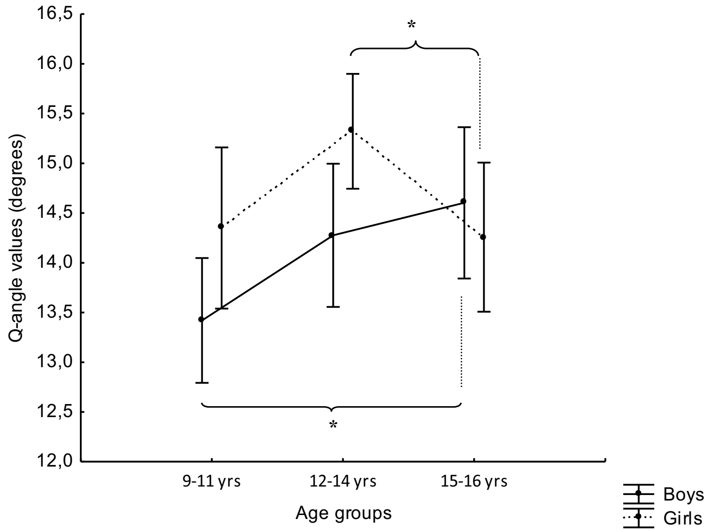
Table VII. Single-limb mini-squat test (maximum numbers of knee bendings/30s)

Children	**n**	**Age** (yrs) Median [min-max]	**Knee bending/30s** Mean (SD)
Boys 12-14	37	12.4 [11.5-14.3]	20.8 (4.6)
Boys 15-16	34	15.4 [14.5-16.4]	18.8 (5.7)
Girls 12-14	57	12.8 [11.6-14.4]	20.4 (5.3)
Girls 15-16	35	15.2 [14.5-16.3]	19.1 (4.3)
All children	163	14.1 [11.5-16.4]	19.9 (5.0)

Note: Data presented as mean (SD) for dominant leg collected during Session 1.

4.2.2 Quadriceps-angle

Reference values (mean ± SD) of the static medio-lateral knee position measured by the Q-angle ranged between 13.5 (1.9) – 15.3 (2.8). No differences between dominant and non-dominant leg (Limb Symmetry Index = 97.13%) or sex differences in Q-angle values were found. A difference was found between age groups (p = 0.034) (Fig. 8). This difference was not statistically dependent on dominant/non-dominant side or sex; however, the interaction, age-group*sex, indicated different effect of age for boys and girls. Further analysis showed a difference between the youngest and the oldest boys as well as when comparing girls in the middle age group and the oldest girls (Fig. 8).



Note: N denotes the number of boys/girls in each age group (9–11) n=30/50, (12–14) n= 59/38, (15–16) n=35/34. *P <0.05

Figure 8. Age differences in Quadriceps-angle values represented by sex (Mean, 95% CI).

A fair to moderate reliability was found for the Q-angle measurements when measuring healthy children (Intra-rater: ICC 0.42, 95% CI 0.11-0.66, SEM 1.4º, repeatability coefficient 3.9º, n.s.; Inter-rater: ICC 0.35, 95% CI 0.14-0.52, SEM 1.9º, repeatability coefficient 5.3º, n.s.). This means that the ICC values were below the suggested criterion values for good reliability in terms of both intra- and inter-rater reliability [98].

4.2.3 Association between the two tests

No associations could be found between dynamic and static medio-lateral knee position assessed by the Single-limb mini squat test and the Q-angle (R^2 = 0.005, n.s.).

4.3 DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF THE KOOS-CHILD

The aims of studies III and IV were to evaluate the comprehensibility of the KOOS when used in children, to modify the KOOS to create a pediatric version (KOOS-Child) to ensure good content validity when used in children, and to evaluate the psychometric properties of the KOOS-Child when used in children with knee disorders. Detailed information on psychometric properties can be found in Paper IV.

4.3.1 Comprehensibility and content validity

The interview results are presented as general observations about the questionnaire. Overall, 2,128 comments about the 42 items were generated in the interviews. Comments were sorted into 4 major categories related to difficulty in understanding and in responding to items, language comprehension/jargon, item format, response-set format and response mapping. Key-examples of problematic issues are presented in Table VIII.

Many children had difficulty understanding medical terminology and complex wording. They found the instructions confusing and stated that it was easy to forget the question stem. Time frame issues were common, and many children had difficulty understanding the distinction between phrases such as “degree of pain” or “degree of difficulty” as well as “how often” and “amount of,” and considered these phrases equivalent constructs. Children also thought that several items were irrelevant because they did not see how the activities could affect an injured knee (e.g., A7, A8, A11, A13, A15 and A17). Some children had difficulty in selecting a response (mapping) if they had not performed the activity or were not allowed to perform it due to the risk of further injury. Double-barreled items (S2, S3, S7, P6, P8, and A13), defined as questions containing 2 or more constructs such as, for example, *“What amount of knee pain have you experienced the last week when going* ***up or down*** *stairs?”* also presented problems. The response option “moderate” was constantly perceived as confusing. Overall, mapping issues (when responses available are not considered suitable) resulted from misinterpretation of items and from design issues related to the item such as double-barreled format. The children also showed a lack of comprehension for the actions in the Sports/Rec subscale and presented many different interpretations of these items. Overall, many children commented that the activities were not representative of children’s daily activities.

4.3.2 Differences in comprehension between groups

Children of different ages raised different concerns with the questionnaire. Overall, the youngest age group raised more issues than the older groups. There were no sex differences regarding comprehension of KOOS-items, except for one item (P1) where boys had more comments than girls (71% vs. 47%). The children who had been treated by either a physiotherapist or had had knee surgery had fewer comments than the untreated groups.

Table VIII. Quotes illustrating different problematic areas

Original item	Comments when probed during interview
Comprehension/jargon	
S2. Do you feel **grinding,** hear clicking or any other type of noise when your knee moves?	"I don't know, 'grinding' makes me think of about when you grind something, like a vegetable
SP5. What degree of difficulty have you experienced during the last week due to your knee **when kneeling** (in Swedish "ligga på knä")?	"Standing on all fours or something…" "I think you bend your knees and lay on your back. That´s how I would do it" "Is it when you lie on your side?" "Like, lying down on your belly and then you have your knees to the floor"
SP1. What degree of difficulty have you experienced during the last week due to your knee **when squatting?**	"I don't really know how to do that. I think you´re sitting on your feet but I don't know..."
Q3. How much are you troubled with lack of confidence in your knee? (In Swedish: I hur stor **utsträckning** kan du lita på ditt knä?)	"If I could trust that I have been able to stretch it?"
Item format *(time-frame, double-barreled)*	
S1. These questions should be answered thinking of your knee symptoms during the **last week.** Do you have swelling in your knee?	"I answered regarding as far back as I can remember."
P6. What amount of knee pain have you experienced the last week when going **up or down** stairs? (none/mild/moderate/severe/extreme)	"I chose "mild" because I have more pain walking down stairs; walking up stairs is no problem"
Response-set format *(terminology)*	
P3. What amount of knee pain have you experienced the last week when straightening your knee fully? **(none/mild/moderate/severe/extreme)**	"Moderate is quite difficult to understand I think…"
Q1. How often are you aware of your knee problem? **(never/monthly/weekly/daily/constantly)**	"I don't like every month and every week. If it is every week, like, it will eventually be every month as well…"
Mapping *(responses available - not considered suitable)*	
SP1. What degree of difficulty have you experienced during squatting the last week due to your knee? **(none/mild/moderate/severe/extreme)**	"I haven´t squatted, but is this if I would do it?" "I don't squat because of my surgery. It's a long time since I squatted." "If I haven´t done this, can I skip it?" "I have not tried that either, I don't think it would hurt so much… I guess I´ll take 'mild'."
A16. For the following activity, please indicate the degree of difficulty you have experienced in the last week due to your knee: Heavy domestic duties (scrubbing floors, etc.). **(none/mild/moderate/severe/extreme)**	"I have never tried this, doing heavy housework. I guess it's not very difficult; it's mostly arms and stuff so I go with 'none'"
A8. Please indicate the degree of difficulty you have experienced in the last week due to your knee: Going shopping **(none/mild/moderate/severe/extreme)**	"I have never gone shopping... what should I fill in then?"

4.3.3 Modifications of the KOOS to the KOOS-Child

To develop the KOOS-Child, modifications of the KOOS were made based on qualitative feedback from the children. Terminology was changed throughout the questionnaire based on the children's suggestions. Since many children had difficulties understanding the instructions and sub-headings they were deleted or modified. Changes in questionnaire layout were also made to make the transition between constructs such as "degree of pain" and "degree of difficulty" more clear, for example.

Most children understood the 5-point Likert response format; however, while most children could interpret the meanings of the words in the response set due to their location, they could not define all the words, and suggested replacing, for example, "moderate" with "some," and "mild" with "a little". This suggestion was implemented in the new questionnaire.

Some children reported that the KOOS did not reflect problems they experienced due to the limited number of items dedicated to strenuous and children's daily activities. Four new items were therefore added: two in the Sport/Rec subscale, and two to the QoL subscale.

Double-barreled items caused problems and were either modified or divided into separate questions. Most children did not carry over the concept of answering the items using the previous week as a time frame. Thus, the time frame was added to each item in the revision. A few items (P1 and the QoL subscale) did not consider "last week" as time frame and when the time frame shifted to a longer duration, many children were confused. Thus, in the revised KOOS, beyond that the time frame was added to each item, the response set of P1 was changed to make the item more easily understood.

When children had been told not to perform an activity, they were unsure of how to answer the questions and suggested adding a "not applicable" box. This was, however, not applied for a couple of reasons: 1) the items are supposed to be answered with regard to perceived difficulty and not if performed or not and 2) such an alternative would make it impossible to determine a change over time. Instead, information was added in the user's guide on how to instruct the patients regarding this matter.

Since the children had many different ideas about how to interpret especially the actions in the Subscale Sports/Rec, illustrations were also added to increase comprehension (Fig. 9).

"Squatting" "Twisting/pivoting" "Kneeling"

Figure 9. Examples of illustrations added to the KOOS-Child questionnaire

As a summary the preliminary version of the KOOS-Child consists of 48 items divided into the same 5 subscales as the original KOOS: Pain (11 items), other Symptoms (7 items), ADL (17 items), Function in sports and play (Sports/Play) (7 items) and Knee-related Quality of Life QOL (6 items). The scoring system from the original KOOS has been kept. Each item receives a score from 0 (no symptoms) to 4 (extreme symptoms). The scores are then normalized to a score ranging from 0 (extreme symptoms) to 100 (no symptoms). In the first modification phase, no items were deleted, but instead considered for possible item deletion after evaluation of the psychometric properties in the next phase (Study IV).

4.3.4 Confirmation of subscales

According to the predefined criteria for possible item reduction, 9 items (P5, P7, P8b, A4, A6, A8, A9, A11, A15) were deleted from the preliminary KOOS-Child. These 9 items showed a ceiling effect >70%. This was in accordance with the results from the qualitative analyses in Study III, where these items were considered irrelevant or redundant. Thus 39 items were kept in the final KOOS-Child. An Exploratory Factor Analysis (EFA) performed on subscale level of the final version demonstrated that items in all subscales except for Symptoms loaded on one factor (Eigenvalues 3.1-5.5, Symptom: two factors, Eigenvalue >1), indicating unidimensionality of all the subscales except for the Symptom subscale being somewhat less homogenous.

4.3.5 KOOS-Child profile

The distribution of KOOS-Child subscale scores from the final version are presented as mean values ± 95 % CI at subscale level in Fig. 10. The percentage of missing responses at item level were ≤5%, and at subscale level ≤2 % for baseline and follow-up sessions.

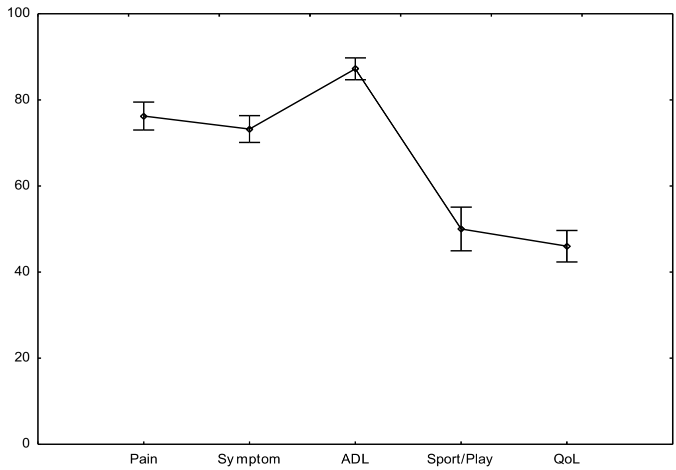


Figure 10. Mean subscale scores (±95 % CI) of the final KOOS-Child at baseline for the included 115 children (56% girls, aged 7-16) suffering from various knee disorders and treated in primary and secondary care at different time points from injury/onset.

4.3.6 Reliability

Test-retest reliability was excellent for all subscales, except for the symptoms subscale showing a slightly lower reliability (ICC 0.78-0.91, 95% CI 0.67-0.94, SDC_{ind} 14.6-22.6, SDC_{group} 1.7-2.7). The inter-item correlation in the subscales, in other words, the internal consistency, was evaluated with Cronbach's α, and sufficient homogeneity of the final version was found for all subscales: Pain α=0.85, ADL α=0.90, Sports/Play α=0.89 and QoL α=0.80, except for the Symptom subscale α=0.59.

4.3.7 Construct validity

Construct validity was evaluated using hypothesis testing. All 7 predefined hypotheses were confirmed, indicating an excellent construct validity of the KOOS-Child (Table IX)

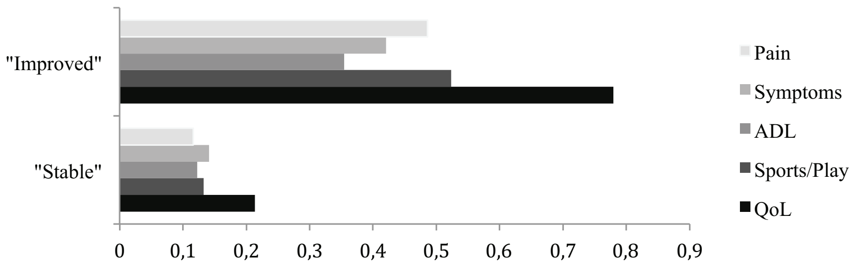
Table IX. Construct validity

Predefined hypotheses	Observed correlation	Hypothesis confirmed (Yes/No)
1. All the KOOS-Child subscales and the **CHAQ "grip"** subscale are hypothesized to have a weak correlation (<0.35). (n=107)	pain: -0.12 symptom: -0.27* adl: -0.23* sports/play: -0.25* qol: -0.19	Yes
2. The KOOS-Child **subscale pain** and the **CHAQ item "pain"** are hypothesized to have a strong correlation (>0.5) (n=110).	-0.73*	Yes
3. The KOOS-Child **subscale sports and play** and the **CHAQ item "run and play** are hypothesized to have a strong correlation (>0.5) (n=107).	-0.72*	Yes
4. The KOOS-Child subscale **activity of daily living** and the **CHAQ "walking"** are hypothesized to be at least moderate (>0.35) (n=109).	-0.63*	Yes
5. The KOOS-Child subscale **"symptom"** should correlate moderately (0.35-0.49) to the respective subscale specific **VAS question** (n=110).	-0.44*	Yes
6. The KOOS-Child subscales **"pain, ADL, sports and play** and **quality of life"** should correlate strongly (>0.5) to the respective subscale specific **VAS question** (n=110).	pain: -0.71* adl: -0.65* sports/play: -0.75* qol: -0.65*	Yes
7. The correlation between the KOOS-Child subscale **quality of life** and the **EQ-5D-Y item "usual activities"** is hypothesized to be at least moderate (>0.35) (n=108).	-0.57*	Yes

Note: Spearman correlation coefficients: > 0.5 strong, 0.35-0.49 moderate and < 0.35 weak. *Correlations significant at <0.05. N denotes the number of subjects responding to both items/subscales used in the respective analyses. Abbreviations: CHAQ - Child Health Assessment Questionnaire, VAS - Visual Analogue Scale

4.3.8 Responsiveness

The predefined hypothesis (a correlation ≥ 0.3 between KOOS-Child subscale change scores and specific GPE scores) was confirmed (Pain 0.43, Symptoms 0.38, ADL 0.48, Sports/Play 0.44 and QoL 0.57, $p<0.05$). As hypothesized, Effect Sizes were higher for patients reporting an improvement in their knee condition compared to those who reported themselves as stable (Fig. 11). These results indicate that the KOOS-Child is responsive to change in children with knee disorders.

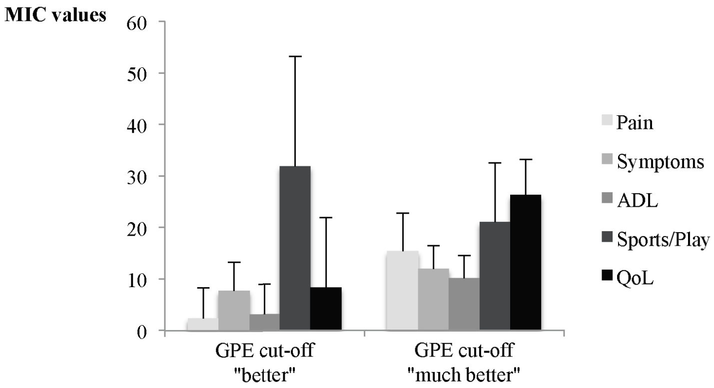


Note: Effect Sizes for patients collapsed into cut-off group “Stable” (Pain n=35, Symptoms n=34, ADL n=35, Sports/Play n=34, QoL n=34) and cut-off group “Improved” (Pain n=48, Symptoms n=49, ADL n=48, Sports/Play n=49 QoL n=48). N denotes the number of patients in each group responding to each subscale at both sessions and to the Global Perceived Effect score.

Figure 11. Effect Sizes for KOOS-Child in children with various types of knee disorders

4.3.9 Interpretability

The KOOS-Child showed no floor or ceiling effects for the different subscales (≤15%). The clinical important change perceived by the subjects was evaluated by the Minimal Important Change (MIC). MIC values for the KOOS-Child subscales are presented in Fig. 12. MIC scores were greater than the SDC both for patients reporting to be "better" and "much better."



MIC values	n	Pain	n	Symptoms	n	ADL	n	Sports/Play	n	QoL
GPE cut-off "better"	18	2.3	19	7.7	13	3.1	7	31.9	11	8.3
GPE cut-off "much better"	28	15.4	25	11.9	26	10.1	19	21.1	19	26.3

Note: Mean change KOOS-Child MIC values with a 95% CI for patients divided into two groups depending on subscale-specific GPE-score cut-off: "better" and "much better." N denotes the number of patients in each group. Abbreviations: GPE score – Global Perceived Effect score

Figure 12. Minimal Important Change (MIC) values of the KOOS-Child

5 DISCUSSION

5.1 GENERAL

Valid and sensitive instruments for measuring knee function are of paramount importance to develop and apply optimum treatments and to detect early changes in knee function in children with knee disorders. Because of a lack of appropriate reliable and valid measures of knee function for children, instruments that have been developed for adults are often used. However, differences in body size, anatomy, cognitive maturity, and lexical ability, among other things, may lead to an inaccurate evaluation and selection of treatments for children [7]. An outcome measure is not valid per se: therefore, it is crucial that any new or modified instruments be evaluated for psychometric properties in the context in which it is meant to be used [77].

The overall aim of this thesis was to evaluate and develop outcome measures aimed to assess functional outcome of the knee in children with knee disorders. The main findings of the studies highlight the importance of using instruments that have been specifically standardized for the study population.

Skeletally immature individuals suffering from a knee injury differ from adults in many ways, and this means that outcome measures need to be adapted and adjusted before they are implemented, so as to decrease the risk of measurement errors. The results from Study I demonstrate difficulties using the Biodex – a gold standard instrument when measuring muscle strength in adults – specifically, problems with stabilizing and standardizing the testing position in the youngest children. The results of Study III shows a lack of comprehensibility of the KOOS, originally developed for adult subjects, when used in children. These data illustrate that using adult outcome measures in a pediatric population may lead to errors in assessment, and eventually this may also result in errors in intervention, diagnosis and prognosis.

The results from Studies III and IV illustrate the importance of a population- and context-specific PRO measure for children with knee disorders. Several PRO measures developed and validated for adults – such as the IKDC, the KOOS, the Tegner activity scale, and the Lysholm knee-scoring scale – have frequently been used in studies evaluating knee function in children with knee disorders [22,137]. Even though some studies report good psychometric properties when the IKDC is used in children, the results are conflicting [88,137]. In a recent report the comprehensibility problems using the IKDC has been clearly established[88]. These results, and the lack of comprehensibility of the KOOS highlighted in our studies, illustrate the difficulty in drawing conclusions from studies using outcome measures not tailored for the target population.

A review by Moksnes et al. [22] aimed to evaluate the methodological quality of 31 studies assessing the outcome after operative and non-operative treatment in skeletally immature children with an ACL injury. The results indicate that the interpretation of results should be performed with caution owing to widespread methodological deficiencies. A major deficiency was the lack of validated outcome measures with adequate sensitivity. Most of the studies included a PRO measure for evaluating knee function. However, none of the PRO measures used in these studies had been validated for the specific population [22].

5.2 RESULTS

5.2.1 The challenge of measuring knee muscle strength in children

Measuring muscle strength in an easy and standardized way in children is a challenge[41]. In study I, we aimed to assess the reliability of the Strength Measuring Chair (SMC) in measuring knee muscle strength in healthy children and adults, as well as its agreement with a widely used ID. While we could observe high test-retest and intra-subject reliability of the SMC, we found a large disagreement between the instruments, particularly in the children's group.

The SMC was developed as a single instrument to be used mainly in clinical settings and in research at the Motion Analysis Laboratory. The goal of developing the SMC was to produce a feasible tool, easy to use in the clinic with standardized measuring positions for both small children and adults [65]. We believe that the large disagreement observed between the two instruments, especially in children, is interesting and highlights the challenges involved in measuring knee muscle strength in children and the importance of using an instrument developed or adapted for the specific population [41,61]. The impression gained from the testing sessions was that the ID's testing equipment was too large for the children and that this influenced the measurements' precision. Since the SMC was developed from the beginning to be adjustable in size to suit both small children and adults, the same problem was not perceived in use of the SMC. The experienced difficulty in standardizing the measuring position in children in the ID is in accordance with previous findings by Tsiros et al. [61]. As a clinical recommendation with respect to the disagreement observed, we believe the SMC and ID should not be used interchangeably.

Previous investigations of knee extensor and knee flexor muscle strength have reported a wide range of peak isometric muscle strength values in both children and adults. There are differences in measuring devices, testing positions and units of measurements in all these studies, which must be taken into consideration before comparing reported results [56,57,63]. The use of torque as the measured unit and standardized position are of great importance to enable comparison of results.

5.2.2 Evaluation of static medio-lateral knee position

The clinical relevance of the Q-angle as an outcome measure of static medio-lateral knee position has been widely debated in the literature [72]. Despite this lack of agreement, the Q-angle is often used in the clinic and discussed as one of many risk factors for future knee injuries [68,138,139]. In Study II, an important finding is that the Q-angle showed a low reliability according to the ICC values found (ICC 0.35-0.42). No sex or side-to-side difference was found, but a significant difference in Q-angle values across age groups was observed.

Reliability and the measurement error of an outcome measure influence the generalizability and clinical usefulness of a tool. One important finding in our study is the high measurement error, illustrated by the Smallest Detectable Change (SDC) for the Q-angle. The SDC ranged between 4° and 5°, which means that a difference of greater than 5° is needed to detect a "true" difference beyond the measurement error. We believe this is a rather high value, considering that the mean values in this population were found to range between 13° and 23°.

The significant difference in Q-angle values between different age groups and the lack of a difference between boys and girls is inconsistent with data from earlier publications [68,74,75]. Bayraktar et al. [74] evaluated the Q-angle in supine position with extended legs in boys 9-19 years of age: their results demonstrate a negative association in Q-angle values with age wherein the youngest boys showed higher Q-angle values than the older boys did [74]. In another study, girls (ages 13-19) were shown to have significantly larger Q-angles (measured in standing) than boys of the same age [68]. Similar results have been reported for adults, women showing larger Q-angles than men [73]. The inconsistency between our results and those previously reported may be explained by the differences in measuring positions applied and by differences in how subjects were divided into age or maturation groups. However, the statistically significant change in age groups in our study has questionable clinical relevance for several reasons. First, the mean difference found was only 1.2° between the groups, with correspondingly large measurement error. Second, there was a low reliability of the Q-angle. It is important to note that ICC values depend on the homogeneity of the sample and that a more heterogeneous group (e.g., children with different types of knee disorders) might have produced higher ICC values.

Measuring positions and the standardization of Q-angle measurements vary in the literature [68,72-75]. In Study II, we chose to measure the Q-angle as it is traditionally described – the subject lying supine with the hip and knee extended, quadriceps relaxed, and foot in a neutral position [129] – because we believed this position would be the easiest to standardize when measuring children. Another procedure is to measure the Q-angle in a standing position or a supine position with the knee slightly bent [72]. Measuring the knee in slight flexion keeps the patella more centralized and may reflect the gait stance better. Measurements with a slightly bent knee have also been suggested as more clinically meaningful than measurements in which the knee is fully extended. There is today no consensus regarding whether the reliability and validity of the Q-angle measurements vary when measured in different knee flexion ranges [72].

Because of the low reliability found in Q-angle measurements, we do not recommend the Q-angle as a measure of static medio-lateral knee position before more research has been conducted. Given that the Q-angle is seldom used as the sole measure today, it is important to consider alternative reliable methods when evaluating medio-lateral knee position in children.

5.2.3 Evaluation of dynamic medio-lateral knee position

The Single-limb mini squat test is a reliable outcome measure for evaluating dynamic medio-lateral knee position in adults with and without osteoarthritis [39,51]. It is a feasible test, easy to use in clinical and research settings and it is easy to instruct the subjects and perform in a standardized way. As mentioned before, the test has two possible outcomes: (1) knee function expressed as the number of knee bendings/30 seconds and (2) the quality of movement as the observed medio-lateral knee position during performance of the test. In Study II, the Single-limb mini squat test was evaluated for reliability when used in children using visual analysis of the medio-lateral knee position, scored as “knee medial to foot” or “knee neutral to foot”. It can always be argued that it is difficult to capture the medio-lateral knee position during the test by using visual observations alone; however, we believe the findings in this study support

the use of the Single-limb mini squat test in a pediatric population because the intra- and the inter-rater reliability results agreed in more than 75% of the cases.
One aspect that should be mentioned is the effect of possible fatigue during the test. We believe that fatigue during this kind of test can be considered a positive property, since this reflects a realistic situation during physical activity. Previous studies have recommended that testing dynamic function in adults should be performed under fatigue conditions [140,141].

The Single-limb mini squat test has shown good validity when assessing medio-lateral knee position in adults. Validity was evaluated using movement analysis, and in the 2-D results, the peak thigh angle and peak tibial angle with respect to the horizontal indicated that the knee was more in valgus in subjects with the knee medial to foot than in subjects with the knee neutral to foot. In 3-D, however, the actual movement was mainly exhibited as an increased internal hip rotation; no difference could be seen in knee valgus between the groups [39]. Until now, the validity of the Single-limb mini squat test has not been evaluated when used in children, and this needs to be considered in future studies.

5.2.4 Development and evaluation of the KOOS-Child

Today, the number of PRO measures is increasing, and there are numerous self-administered instruments for measuring knee outcomes in adults. The target population must however be considered, since there is no standardized questionnaire for all groups of patients or knee disorders [84,112]. Lexical comprehension is an especially important aspect to consider when questionnaires are used in a pediatric population [88]. In the clinic, questionnaires designed for adults are often used for children due to a lack of suitable, validated instruments for this population. Thus, the potential exists for inaccurate evaluation and inaccurate selection of treatments for children. A PRO measure will never be perfect, but the goal is to use or develop one that is as good as possible.

Development process
In the process of developing a pediatric questionnaire for children with knee disorders, the KOOS-Child, we found that the original KOOS was not appropriate in its original form, especially for younger children (ages 10–12). Thus, the KOOS required modifications for this purpose. In general, our results show that children raised many concerns regarding comprehension of items, directions, time intervals, and medical terms. Comprehension problems associated with double-barreled questions, medical jargon and difficult terms are well known to lead to misunderstanding and to default or missed responses [120,121,123,142].

Based on our results, we modified the KOOS in several areas: general instructions, lexical/language, item format, response format, and mapping. To our knowledge, there is no knee-specific PRO measure valid for use in children currently available in Swedish. In a recent study by Iversen et al. [88], the IKDC was modified for children (producing Pedi-IKDC), using the same methodology as in our study. In accordance with the earlier study, our results illustrate that modifying questionnaires for use with children is indicated and that cognitive interviewing is an appropriate methodology for identifying areas of concern and making recommendations for improvement [88].

A PRO-measure developed for evaluating knee function needs to identify aspects important to the targeted group, be stable over time, and be sensitive to change in knee function [77,84]. To ensure good content validity, the instrument's content needs to adequately reflect the construct to be measured in the specific target population. During the interviews, many children voiced concerns that their chief complaint items were not fully captured by the KOOS. When further probed, the children expressed that activities included in the KOOS did not correspond to what they did during their daily lives. To ensure a good content validity of the KOOS-Child, questions were therefore added based on their suggestions – questions, for instance, about participating in sports, school activities, and social participation. In the second phase, in which the psychometric properties were evaluated, the high ceiling effects of nine items also reflected this: the qualitative results accorded with the quantitative analyses, supporting the deletion of these items from the final version.

Evaluation of psychometric properties

In Study IV, the evaluation of the KOOS-Child showed good psychometric properties when used in children with knee disorders. According to the ICC values, our results showed excellent test-retest reliability (ICC 0.85-0.91) for all the subscales except the symptom subscale, which demonstrated slightly lower reliability (ICC 0.78). This is in accordance with the reliability (ICC 0.61-0.95) shown of the original version of the KOOS when used in adults with knee injuries [84]. As recommended in the literature, the ICCs were complemented by the standard error of measurement (SEM) and the SDC [96,99]. Our results demonstrate that small changes (2-3 points) can be measured over time at group-level but a considerably larger change (15-23 points) will be needed to detect a true change in an individual over time. This is a previously known phenomenon in PRO measures [132].

The KOOS-Child consists of five different subscales. The unidimensionality and homogeny of the subscales were evaluated using Exploratory Factor Analysis (EFA) and Chronbach's alpha, a coefficient of internal consistency (application and possible shortcomings of EFA are discussed separately in the section on methodological considerations). The symptom subscale was found to be somewhat less homogenous than the other subscales. This was an expected result, since this has previously been shown for the original KOOS [84,85]. The lower homogeneity can be explained by the variation found in symptoms experienced by patients with different knee disorders.

The validity of a PRO measure is a complex property to evaluate especially when there is no gold standard instrument to compare the instrument with. In Study IV, several instruments were used to cover the different components of the KOOS-Child: the EQ-5D-Y [79,126], the CHAQ [125] and the subscale-specific VAS-scales. These comparator instruments have not been evaluated for use in children with knee disorders but were chosen as the best available options. Predefined hypotheses were set up and confirmed statistically, indicating excellent construct validity of the KOOS-Child when used in this specific setting and population. To our knowledge, there are no recommendations or standards for constructing hypotheses or regarding the number of hypotheses. This is a limitation, given that the number influences the percentage of confirmed hypotheses.

One of the most important properties of a PRO measure is its ability to detect change over time. Analysis of responsiveness and the Minimal Important Change

(MIC) of an outcome measure is a continuous process that is strongly recommended as strengthening the applicability of the outcome measure [94,116]. There is an ongoing debate about the most appropriate statistical approach when evaluating responsiveness [143]. Traditionally, responsiveness parameters, such as the Effect Size and Standardized Response Means, have been used and can be found in a large number of studies [84,112,143]. When the Delphi study reached consensus for the COSMIN checklist, the Effect Size and Standardized Response Mean were considered measures of magnitude of the intervention rather than measurements of the quality of the measurement property [96]. The COSMIN panel proposed that responsiveness should be evaluated similar to construct validity using hypotheses testing, i.e. by comparing changes in scores of the actual instrument to expected correlations with changes in other measures. The aim of the hypothesis testing, according to COSMIN, is to show that the instrument measures what it is supposed to measure, and also that it measures the right amount of change and does not over- or underestimate it [96].

In Study IV, responsiveness of the KOOS-Child was evaluated using a global index measure of health status, the Global Perceived Effect (GPE) scores, compared to the change in KOOS-Child scores between baseline and three-month follow up. Correlations between the GPE scores and KOOS-Child subscale change scores were found to be 0.38-0.57. As a complement, Effect Size and Standardized Response Mean were evaluated since these measures are still the most widely reported in the literature for knee-specific PROs [84,112]. As hypothesized, the Effect Size and the Standardized Response Mean were higher for the group that had clinically improved than for those who had rated themselves as stable. The children took part in a variety of interventions, or in no intervention at all during the three-month follow-up period, which means that the effect sizes determined in this study cannot be related to a specific clinical intervention effect nor can it be used to determine future sample sizes needed in clinical studies. Moreover, there is a possibility that a longer period would have been needed for some children to experience a clinical change depending on diagnosis.

Interpretability, the degree to which qualitative meaning can be ascribed to quantitative scores is crucial for score interpretation and application in both clinical and research settings. During the evaluation of the KOOS-Child, we employed several methods to illustrate the interpretability of the questionnaire when used in children with knee disorders. The presence of floor and ceiling effects, were first evaluated, and we found these to be below the predefined criteria (≤15%) for all KOOS-Child subscales. These data indicate that it is possible to measure improvement and deterioration over time in groups of patients. The subscales Sports/Play and QoL had the lowest floor and ceiling effects, a result that accords with the children's impression of their conditions expressed during the cognitive interviews.

The next step was to evaluate the MIC of the KOOS-Child. In Study IV, we chose to use the MIC definition by COSMIN [96,119], selecting the most common anchor-based approach, which includes the patient's perception of change represented by a GPE score [107]. The reliability of the GPE-score is, however, discussed in the literature because it tends to be influenced by present health status [107-110,130]. The MIC for the KOOS-Child in our study was defined as children reporting to be "better" and "much better" on a subscale-specific GPE score during the three-month follow-up period. There is today no consensus regarding the best cutoff for the MIC [108]. In many studies the alternative "somewhat better" or similar is chosen. We chose to collapse the groups "somewhat worse", "no change," and "somewhat better" to represent the stable group.

Our results show incongruence between the groups in the degree of improvement needed in the KOOS-Child subscale scores to detect a clinically important difference. One reason for this incongruence might be the dependence on the GPE scores [115]. The other reason may be the small sample size in each group (“better” n = 7-19; “much better” n = 19-28). Larger sample sizes are needed to more accurately evaluate the MIC for the KOOS-Child. The MIC values found are not surprising though, e.g., only seven individuals were included in the “better” group and the small sample is represented by the high values and large confidence intervals found for the subscale Sports/Play. Our MIC values indicate that a larger change in KOOS-Child scores are needed in order to detect an important change in the group reporting to be “much better” than is needed in the group “better”, a finding that seems reasonable.

Comparison of KOOS and KOOS-Child profiles in different populations

In a recent prospective study by Moksnes et al. [144], the KOOS was used as a measure, in order to evaluate functional outcomes in 46 skeletally immature Norwegian children (< 12 years) with an ACL-injury two years after receiving non-operative treatment. This was the first to include both physical function and PRO measures in the evaluation of knee function in children with ACL-injury. The main results indicated that a non-operative treatment algorithm may be appropriate for skeletally immature children with ACL-injuries, although a reduced participation on the highest activity level may be necessary [144]. However, a shortcoming of this study was the use of a PRO measure developed specifically for adults. The authors discuss this issue, as a limitation and recommend interpreting the KOOS results with caution and note that comparisons of the KOOS scores from studies on adult patients should be avoided [145]. Figure. 13, plots the KOOS subscale scores reported from the Norwegian cohort and KOOS-Child subscale scores from Study IV to illustrate the dissimilarities in profiles found.

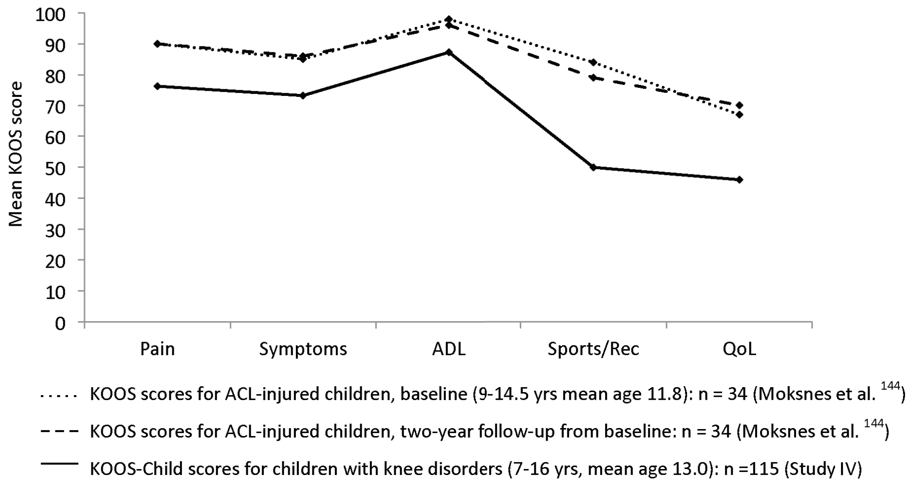


Figure 13. KOOS and KOOS-Child profiles presented for the Norwegian cohort (skeletally immature individuals with ACL injury) and from Study IV (children with various knee disorders).

There may be several potential explanations for the differences in the KOOS/KOOS-Child profiles seen in children (Fig. 13). First, it is important to recognize the differences in samples. Our sample comprises a group of children with various knee disorders who were seeking medical care owing to symptomatic knee problems and restrictions in physical activity. In the Norwegian cohort, the ACL-injured children were evaluated for the first time (baseline), in mean 11.7 (SD 11.5) months after injury [144].

Additionally, there is a potential of a ceiling effect among children with ACL-injuries when evaluated using the adult KOOS. Evidence to support this assumption is the fact that nine items have been deleted and four new items have been added to the KOOS Sports/Play and QoL subscales in the creation of the KOOS-Child in order to better reflect the problems reported by children with knee disorders. In summary, the difference seen in the plot highlights the challenges of interpreting data collected with an outcome measure not specifically designed for the studied population.

In Fig. 14 the KOOS profiles from a cohort of 7 331 young adults with ACL-injuries from the Swedish ACL registry [146] and normative KOOS values for 134 healthy young adults [147], as well as the KOOS-Child scores from Study IV are presented. These data illustrate the similarities in pattern of the KOOS-Child profile between children with various knee disorders and young adults with ACL-injuries. We believe these data strengthen the results of the KOOS-Child scores found in Study IV.

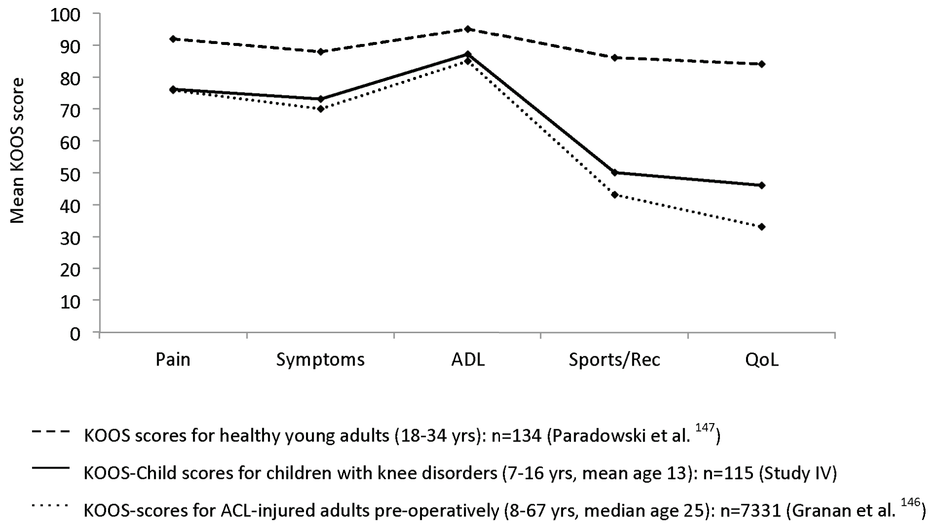


Figure 14. KOOS and KOOS-Child profiles presented for healthy young adults, adults with ACL injury (pre-operatively) and from children with various knee disorders (Study IV) respectively.

The KOOS-Child was developed in Swedish and English simultaneously in Study IV; however, further psychometric validation may be needed in populations of English-speaking children. Further psychometric testing in English-speaking children may identify cultural differences, such as dissimilarities in activity pattern and activity level.

Since the KOOS-Child is a modified version of the KOOS, the same administration of scoring and handling of missing items/ambiguous answers was followed (www.koos.nu). In many adult orthopedic knee scales, an aggregated score is calculated for all items [84]. A single score is often preferred for simplicity, whereas subscale scores allow evaluation of separate constructs and enhance the clinical interpretation of results. The use of separate subscale scores also ensures content validity (e.g., in groups with different physical activity levels). In the KOOS-Child, we have followed the recommendations of the KOOS to keep the subscale scores separate.

The question about transition between the pediatric version (KOOS-Child) and the adult version (KOOS) remains unanswered today. Study III and IV were not designed to determine a definite age cutoff for the use of the KOOS; rather the studies were designed to develop a version appropriate for use with children. A recommendation is that PRO measures should not require reading skills beyond the 6th grade-level (12-year-old) to avoid missing values and unreliable answers [77]. This recommendation supports the use of the KOOS-Child even in older subjects. However, researchers should also consider the content and the relevance of the items to the specific target group. Future studies will be needed to resolve possible issues with longer-term follow up of children with knee disorders into adulthood.

5.4 METHODOLOGICAL CONSIDERATIONS

5.4.1 Subjects

A total of 449 subjects participated in the studies constituting this thesis. Two studies (I and II) considered healthy control subjects (n = 296), and the other two studies (III and IV) involved children with various knee disorders (n = 153).

Before using an outcome measure in patients with disorders, it is important to have reliable reference data. Therefore, the first step in Studies I and II was to evaluate the reliability of the outcome measures in healthy subjects. In Study I adults were recruited as a reference for the pediatric group. We believe that it was important to evaluate both children and adults since the instrument was supposed to be applied to both groups in clinical and research settings.

In the studies, both children and adolescents were included. The definitions of child and adolescent are arbitrary, and adolescence is a difficult period to define because of the wide variations in its onset and termination. According to MeSH definitions [148], children are those aged 6-12 and adolescents are aged 13-18 [148]. However, the definition of being a child or adolescent depends on diagnosis, maturity, type of clinic etc. In the studies, we therefore chose to use the term children for all subjects under the age of 18.

The sample size in a reliability study does not need to be large. According to Fleiss [131], 15-20 will usually be enough for a quantitative variable, but more will be required for estimating the reliability of a categorical variable [131]. In Study I, we included 20 children aged 5-13, however, the age distribution in the children's group was not equally distributed (with more children aged 6-8 years). This imbalance may have affected the disagreement found between the instruments; a larger group of children with a broader age range would be required to confirm and further evaluate the systematic disagreement found.

In Study II the required number of subjects for reference values was estimated to be 250 and for the reliability calculations to be at least 32 [149]. During data collection, children between the ages of 9 and 11 were observed discussing their performance scores of the Single-limb mini squat test with their classmates. Subsequently, these children began performing the test with less accuracy owing to a growing competition – even though being tested individually in a private area. Therefore, we did not include the data collected from these youngest children for the Single-limb mini squat test. Nevertheless, the final number of subjects in the analyses was regarded as sufficient. All the subjects in Study II were divided into the age groups 9-11, 12-14 and 15-16. This division might have influenced the interpretation of the results and needs to be considered when comparing the results in other studies that use different age or maturation groups [68,74,75].

There is no gold standard for calculating sample sizes for qualitative studies. Thus, subject factors, the number of items, and the application of the questionnaire are important to consider [103,150,151]. Although no techniques are available to calculate sample sizes for qualitative study designs, sample sizes of 30 or more are common [152,153]. To obtain a broad representation of children with different knee disorders in Study III, children were selected from both primary- and secondary-care clinics. The children were purposefully selected to allow for an approximately equal group representation of age and sex. We selected an initial number of subjects and then continued until saturation was reached. In total, 34 subjects were included, a relatively large sample size for this type of study. Since self-reported questionnaires rely on lexical comprehension, and this is especially important to consider when such instruments are used in a young population [88], we recruited children from the age of 10 years. Additionally, children under the age of 10 suffer knee disorders less often. In Study IV, one 7-year-old child and five 9-year-old children were included below the age of 10. The 7-year-old had missing data for both follow-up sessions and one of the 9-year-olds had missing data at the test-retest follow-up. Even though these children were below age 10, we chose to include their data in the analyses so as to achieve the highest possible external validity. It is important, though, to instruct parents to help the children read the questionnaire when needed.

5.4.2 Psychometric evaluation

There are several available guidelines for evaluating the methodological quality of a study on measurement properties [77,96,154-156]. In order to achieve consistency in terminology regarding psychometrics, the definitions by COSMIN were used in the present thesis [96]. The COSMIN checklist was also used as a guide when designing Study IV. Compared to other guidelines, the advantage of the COSMIN is that it is consensus based and focus on health status measurement. As already mentioned, the COSMIN checklist was developed based on the results of a Delphi process that included 57 international experts in the field. The COSMIN was developed in order to evaluate the methodological quality of a study on the measurement properties of a PRO measure, not the quality of the PRO measure itself [95,96,100].

5.4.3 Measurement theories

A measurement theory is a theory addressing the ways the scores generated by the items represent the construct to be measured [94]. The most commonly used measurement theory when evaluating PRO measures for psychometric properties – also used in Study IV – is the Classical Test-Theory (CTT). Another relatively new approach to evaluate questionnaires must be mentioned though – namely, the Item Response Theory (IRT) [77]. The IRT has some advantages over the CTT that need to be addressed. The Rasch model is a mathematical model applied in IRT that has been used to develop and internally validate questionnaires [157]. One advantage of the IRT is that the Rasch model uses a logistic function that creates a scale with interval properties, in contrast to the CTT, in which ordinal scales are treated as interval scales [77,157]. In general, CTT focuses primarily on test level, and there is an assumption that each item in a scale contributes equally to the final score, whereas the IRT focuses on item-level information and can, for instance, provide detailed information about the difficulty of an item or detect items that do not fit into the construct to be measured. One of the major assumptions in IRT is that the scale needs to be unidimensional – the scale should measure only one attribute – however, this is also believed to be one reason why the IRT is not as widely used as the CTT is [77]. In the future, Rasch analysis could be considered in order to validate the KOOS-Child further.

5.5 STATISTICAL CONSIDERATIONS

5.5.1 Factor analysis

Factor analysis is a widely used method for evaluating whether items of a questionnaire can be grouped into clusters representing different dimensions of the construct to be measured [77,94,133,134]. There are two main types of factor analysis, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). When performing an EFA, no expectations are made in advance, in contrast to CFA, in which the model should be based on previous expectations and prior hypothesis [77,94,133].

In Study IV, we performed an EFA using principal component analysis with varimax rotation for each KOOS-Child subscale separately in order to evaluate subscale unidimensionality [134]. The main reason for this was the addition of four new items to the questionnaire. EFA is the most commonly used method, but CFA has received more attention in the literature in recent years. There is an ongoing debate regarding the best application of these methods [77,94,133].

Since we modified the KOOS-Child from an existing version (in Study IV), one could argue that a CFA would have been the most appropriate model to use. However, the reasons for choosing EFA were multiple. First, both the EFA and CFA require a reasonable amount of data in order to produce reliable results. A sample size of ≥ 100 subjects and seven times the number of items is recommended [113]. The KOOS-Child consists of 39 items, which means we would have needed a much greater sample size of 273 subjects, in order to perform a CFA or an EFA on the complete scale at once. Even though knee injuries in children constitute one of the most common sports injuries, the population is rare, and the number of items in the KOOS-Child is also high. We believe that collecting further data to obtain the required sample size would have prolonged the study period markedly. Another reason we performed an EFA on the subscale level was to achieve a pediatric version of the KOOS that was similar in structure to the adult version in order to facilitate comparison of the two and to make transition between the scores possible in the future.

Factor analysis is complex, but it is a flexible analytical process. The clinical reality versus the statistical result must be considered when interpreting the results [94]. Our results show that the items in the subscales Pain, ADL, Sports and QoL loaded on one factor with an Eigenvalue of 3.1-5.5. However, the Symptom subscale loaded on two factors with Eigenvalues > 1, indicating unidimensionality of all the subscales except the Symptom subscale, which was somewhat less homogenous. From a clinical point of view, we decided that the symptoms subscale should be kept even if the model showed two possible factors. The first reason is that knee injuries are often concomitant; therefore different symptoms will be perceived. Since pathology often is unknown at time of injury, a self-reported instrument needs to assess symptoms that are not necessarily correlated. Second, it is previously known that the KOOS symptom subscale is not unidimensional [85].

In the final version of the KOOS-Child, nine items were deleted owing to a high ceiling effect and in response to the results of the cognitive interviews in Study III (three from the Pain subscale and six from the ADL subscale). Before we deleted these items, these two subscales loaded on three factors each with an Eigenvalue > 1. That the deleted items resulted in the same subscales loading on one factor each, we believe, supports our choice to perform an EFA at subscale level.

5.5.2 Interpretability - Minimal Important Change

In study IV we used the most common anchor-based approach, the mean change method, to calculate the MIC [107]. The disadvantage of this method is that it does not consider the misclassifications or overlap in scores between the groups, e.g., if subjects in the "unchanged group" have the same KOOS-Child scores as those in the group "better" or "much better". The Receiver Operating Curve (ROC) method is another anchor-based approach that would have been the preferred approach, if we had had a larger sample size and a more even distribution between the subgroups according to GPE score. The advantage with the ROC method is that the degree of misclassification is taken into consideration [113]. In our study, however, the ROC analysis was considered unreliable because the number of patients in each subgroup (according to the GPE scores) was low, especially in the unchanged group. Sample-size requirements for determining ROC MIC values are considered to be at minimum 50-100 patients in each sub-group [113]. This method should be considered in future studies with larger sample sizes.

5.7 LIMITATIONS

All studies have limitations and before the conclusions were drawn from the studies in this present thesis, the limitations have been thoroughly discussed.

First, a general comment about terminology that could possibly be seen as a limitation: in orthopedic literature, there is a lack of consensus on the terminology regarding different outcome measures evaluating knee function in subjects with knee disorders. An example is the different terminology found in the literature for measures of physical function such as, performance-based measures, performance tests, dynamic performance measures, physical performance measures, and functional performance tests. In this thesis, all outcome measures used in the separate studies were sorted under clinical outcome measures and titled by their properties, such as measures of muscle strength, measures of static and dynamic knee position, and PRO measures.

Another approach would have been to sort all outcome measures according to the International Classification of Functioning, Disability and Health (ICF) [158]. The ICF is a framework that is used as a classification system of health and health-related domains to achieve an overall picture of the functional status assessment, goal setting, and treatment planning, as well as outcome measurement [159]. However, since the ICF was not used as a framework from the beginning, and since the terms performance and function creates confusion as they have different meanings in ICF [160] and in orthopedic literature, the ICF was not incorporated as a framework at this point.

It would be of importance to reach a consensus in the future concerning how to categorize and define different types of outcome measures in the orthopedic literature, preferably in a way that works with the ICF.

The sample in Study I were not equally distributed by age, which is a limitation. A larger group of children with a broader age range would be required to confirm and further evaluate the systematic disagreement found between the SMC and the ID. Some of the participants also experienced pain during testing using both instruments. Since we did not register the details about the pain sensations, we can only speculate about the cause. This highlights the importance of including standardized measures of pain in any study that includes testing of muscle strength. Another limitation of Study I is the lack of evaluation of the ID's test-retest reliability using the special adaptations in the pediatric group. This needs to be considered in future studies.

In Study II, all the included children were stratified for age and sex in the statistical analyses. The differences found between age groups in our study differ from results presented earlier in the literature. This may be explained by the differences in how children were divided and one limitation is that no registrations about maturation status were made. Another consideration is the lack of information on the height and weight of the children, which would have facilitated the interpretation of the reference values of the Q-angle.

For assessment of test-retest reliability in Study IV, the children answered the KOOS-Child during their visit at the clinic the first time and then at home during the follow-up session. These differences in administrative setup could be regarded as a limitation. However, despite these differences, a high reliability was achieved, which we believe indicates that this common way of using questionnaires in clinical practice is of little importance for the administration of the KOOS-Child.

The reliability analyses of the KOOS-Child were made based on the preliminary version, which for practical reasons is often the case in such studies. This could, however, be seen as a limitation. When planning Study IV, this issue was discussed. The other approach would have been to perform a pilot test of the preliminary version and then do the psychometric testing on the final version. This would, however, have prolonged the study time markedly because we wanted to cover a broad spectrum of children of different sexes, ages, and knee diagnoses. Given these circumstances, this approach was not considered realistic. We believe, however, that since the results from the quantitative analysis (Study IV) were confirmed by the results obtained from the qualitative study (Study III), our decision was supported.

Additional limitations in the psychometric testing of the KOOS-Child in Study IV include the following: difficulties finding appropriate instruments to compare with the outcome measure being tested, the use of shortened versions of the comparator instruments, and the fact that the number of hypotheses influences the percentage of confirmed hypotheses. To our knowledge, there are no recommendations on or standards for constructing hypotheses and for the number of hypotheses.

6 CONCLUSIONS AND CLINICAL IMPLICATIONS

There is a great need to specifically develop and apply outcome measures for evaluating knee function in a pediatric population with knee disorders. In the present thesis, outcome measures originally designed for adults have been evaluated for use in children, and new instruments have been developed. The present work highlights the importance of using instruments that are standardized for the particular study population when measuring knee muscle strength and when evaluating medio-lateral knee position. It also provides a new PRO measure, the KOOS-Child, for measuring knee function and knee-related QoL in children with knee disorders. The following paragraphs present more specific conclusions drawn from the studies.

STUDY I

- The SMC, reliably measured knee muscle strength in children and adults. However, the disagreement found between the two instruments, more in the pediatric group, warrants care in standardizing measuring positions for different body size and demands caution in comparing one's muscle strength measured using different methods.

STUDY II

- According to the reliability results of the Single-limb mini squat test and the Q-angle, we believe the Single-limb mini squat test can be used in a pediatric population however, the Q-angle needs further investigation before its use can be justified.
- The smallest detectable change of the Q-angle when used in children was 4°- 5°, indicating that the difference of < 5° found between age and sex may not be clinically relevant.
- No association could be found between the Single-limb mini squat test and the Q-angle indicating that two different concepts being measured.

STUDIES III and IV

- The KOOS is not well understood by children, especially younger children. Modifications of the KOOS were made based on qualitative feedback from the children, resulting in a new pediatric version, the KOOS-Child.
- The KOOS-Child demonstrates good psychometric properties and we recommend that it be used in clinical and research practice, when evaluating knee function and knee-related QoL in children with knee disorders.
- The KOOS-Child was developed in Swedish and English versions and can be used to monitor both individuals and groups over a short or a longer time-period.

7 FUTURE PERSPECTIVES

In the future, there is a need for prospective cohort studies to clarify long-term health in children with various knee disorders. To do this properly, outcome measures with good psychometric properties when used in the specific population need to be used. The studies in this thesis extend the knowledge and understanding of outcome measures when treating children with knee disorders in clinical settings and in research. However, this is just the beginning.

In future studies, it would be of interest to further evaluate the performance of KOOS-Child in children with different knee disorders. Normative values for healthy children and children with different knee diagnoses need to be established. The responsiveness and interpretability of the KOOS-Child also need to be further investigated for children with different knee diagnoses and for different treatment algorithms. Future studies will also be needed to help resolve possible issues with long-term follow up using the KOOS-Child – namely, the transition into adolescence and adulthood.

Furthermore, there is a need for additional studies evaluating existing outcome measures and, if necessary, developing new ones, such as a standardized battery of functional performance tests to evaluate knee function – specifically designed for pediatric populations with knee disorders. Future research in this area could potentially yield significant health benefits and new clinical guidelines for treating children with knee disorders.

8 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all people who have inspired, supported and encouraged me during these years and made this thesis a reality. Especially, I would like to thank:

All children and their families, friends and colleagues who participated in the studies, for sharing your time.

Associate Professor, **Eva Weidenhielm Broström**, main supervisor and friend. You amaze me with your never-ending energy, optimism and new research ideas. You have supported me along the way, through easy and hard times. Thank you for encouraging me to become a PhD-student and for always believing in me.

Professor, **Maura D Iversen**, co-supervisor, for invaluable support and for generously sharing your expertise in qualitative methodology and research in general. Thank you for critically reviewing my work, answering my long lists of questions and for constructive discussions. I have learned so much from you.

Professor, **Ewa Roos**, co-supervisor, for always asking me the right questions and sharing your amazing expertise about development and evaluation of patient-reported outcome measures and research methodology. Thank you for the support along the way and for your patience with my never-ending meeting agendas.

Per-Mats Janarv, PhD, Pediatric orthopedic surgeon and co-supervisor, for encouraging and supporting me. Thank you for introducing me to children's orthopedics and for generously sharing your knowledge.

Professor, **Suzanne Werner**, co-supervisor, for sharing your expertise in sports medicine and for critically reviewing my work.

Karolinska Institutet and the **Department of Women's and Children's health,** for giving me the opportunity to become a PhD-student.

Agneta Cuypers, external mentor, for always being there. Your support has made this thesis a reality. I'm very grateful for our long conversations about research and life. I hope that we will continue to have our long talks even when I have finished my PhD-studies.

Associate Professor, **Lanie Gutierrez Farewik**, co-author, colleague and friend, for all the support, constructive discussions and linguistic help.

Associate Professor, **Åsa Bartonek**, co-author, colleague and friend, for your support and for sharing your knowledge in research. Thank you for always being helpful, for your valuable comments and for critically reviewing my work.

Mia André, PhD and Head of the Pediatric Section for Physiotherapy, Karolinska University Hospital, Solna, for believing in me and giving me the best working possibilities to combine research and clinical work. Thank you for always being available and supportive. You always have the right things to say at the right time and your advices, about work and life in general, has meant a lot to me.

Pia Lundell, physiotherapist, co-author and friend, for your important contribution and for generously sharing your knowledge in sports medicine.

My friends, former and present colleagues and roommates at the Motion Analysis Laboratory, **Marie Eriksson, Elin Lööf, Josefine Naili Eriksson, Kicki Löwing,** and **Ruoli Wang**. These years have been full of joy and laughter but sometimes also frustration and hard times. With you around, work has been so much easier. You are the best! Especially I would like to thank, **Mikael Reimeringer** for the invaluable support with computer related issues, **Cecilia Lidbeck,** friend and colleague, for sharing you knowledge, for support and encouragement both in clinical and research work and excellent mentorship, and **Anna-Clara Esbjörnsson**, dear friend and colleague, for your honesty, invaluable support and for great friendship, without your support I'm not sure I would have reached this far!

Friends and colleagues, **Stefan Gantelius, Johan von Heideken** and **Marie Askenberger,** Pediatric orthopedic surgeons and PhD-students, for all our research discussions, for sharing your expertise in clinical work, contribution in recruiting patients, but most of all for the joy and talks about life in general during our PhD-student pub nights. Especially I would like to thank **Eva Bengtsson Moström**, friend, co-author and colleague, for your dear friendship, great contribution and all the encouragement and support along the way!

No one mentioned, no one forgotten - all my fantastic **colleagues at the physiotherapist department**, Astrid Lindgren Children's hospital. My former colleagues, for teaching me almost everything I know about pediatric physiotherapy and my present colleagues and friends, for making me love going to work and for encouragement and support along the way. I'm really looking forward to join the group again and to get to know all my new colleagues.

Research colleagues on the 7th floor, at the Department of Women's and Children's Health, for inspiring discussions during seminars and lunch-breaks.

None-Marie Kemp for always being very helpful and having the answer to every possible question there is and **Astrid Häggblad** for all the help regarding the administrative issues being a PhD-student.

Christina Björk, medical secretary at the Orthopedic department, Astrid Lindgren Children's hospital, for your rigorous and invaluable work in transcribing all the interviews.

Anna Eriksson, for your valuable contribution translating all the comments from the interviews to make the discussions in our research group possible.

All the **physiotherapy clinics** helping out with the recruitment of patients. A special thank to **Anna Hellsten** at the Sport och Rehab, Vasastan, for your interest and valuable contribution.

Elisabeth Berg, for great statistical help during the years.

Caroline Källmin, for your invaluable support during one part of this journey. I think of your advices often and will always remember and try my best to follow them.

Mathias Karlsson, my brother-in-law for drawing the fantastic illustrations to the KOOS-Child and in the thesis. I love them!

All my wonderful friends outside work, you mean so much to me. Especially I would like to thank **Pia Linghede** and **Lynn Andersson** for always being there, through easy and hard times, for sharing great times, just us or together with our families, and for bringing me back to earth when work is eating me up.

Inger Karlsson, my dear mother-in-law, for always being there for the family and me. For giving life perspective and for taking the best care of Ville whenever needed!

My grandfather, **Sten Johansson**, for your enthusiasm, support and never-ending interest in my work and studies.

My parents, **Eva** and **Åke Örtqvist**. You are my greatest role-models in life. Thank you for your endless love and support. This journey has not always been easy, but you have always been there by my side believing in me and giving me strength. I love you and you are the best!

My sisters, **Lisa** and her family **Daniel** and **Agnes,** and **Anne** and her boyfriend **Gustaf**, for being the best family there is. Lisa and Anne, you are my best friends and I am so grateful for all the happy memories we share, for all support and the joy you bring to my life. Thank you for always being there, through easy and hard times. I love you.

And finally to **Kecke,** my soulmate and love of my life. You always encourage and support me. Thank you for your immense patience and for giving me perspective and reminding me of the important things in life. You complete me!

To **Ville**, our son and miracle, you are the love, happiness and joy of our life. You mean the world to me!

The work of this thesis was supported by Karolinska Institutet (KID-funding), grants from Karolinska University hospital (FoU), Promobilia foundation, Sunnerdahl foundation, Norrbacka Eugenia Foundation, Foundation Frimurarna barnhuset in Stockholm, Trygg Hansa and Sällskapet Barnavård.

9 REFERENCES

1. Caine D, Caine C, Maffulli N. Incidence and distribution of pediatric sport-related injuries. Clin J Sport Med 2006;16:500-13.
2. Maffulli N, Longo UG, Gougoulias N, Caine D, Denaro V. Sport injuries: a review of outcomes. Br Med Bull 2011;97:47-80.
3. Roos EM. Joint injury causes knee osteoarthritis in young adults. Curr Opin Rheumatol 2005;17:195-200.
4. von Porat A, Roos EM, Roos H. High prevalence of osteoarthritis 14 years after an anterior cruciate ligament tear in male soccer players: a study of radiographic and patient relevant outcomes. Ann Rheum Dis 2004;63:269-73.
5. Caine DJ, Golightly YM. Osteoarthritis as an outcome of paediatric sport: an epidemiological perspective. Br J Sports Med 2011;45:298-303.
6. Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. Am J Sports Med 2007;35:1756-69.
7. Vitale MG, Levy DE, Johnson MG, Gelijns AC, Moskowitz AJ, Roye BP et al. Assessment of quality of life in adolescent patients with orthopaedic problems: are adult measures appropriate? J Pediatr Orthop 2001;21:622-8.
8. Hinton RY, Sharma KM. Acute and recurrent patellar instability in the young athlete. Orthop Clin North Am 2003;34:385-96.
9. Scott W. Surgery of the Knee, Chapter I: Anatomy. Philadelphia, USA: Churchill Livningston Elsevier; 2006.
10. Caine D, DiFiori J, Maffulli N. Physeal injuries in children's and youth sports: reasons for concern? Br J Sports Med 2006;40:749-60.
11. Shea KG, Grimm NL, Ewing CK, Aoki SK. Youth sports anterior cruciate ligament and knee injury epidemiology: who is getting injured? In what sports? When? Clin Sports Med 2011;30:691-706.
12. Tanner JM, Davies PS. Clinical longitudinal standards for height and height velocity for North American children. J Pediatr 1985;107:317-29.
13. Socialstyrelsen. Skador bland barn i Sverige. 2011:1-179.
14. Louw QA, Manilall J, Grimmer KA. Epidemiology of knee injuries among adolescents: a systematic review. Br J Sports Med 2008;42:2-10.
15. Parkkari J, Pasanen K, Mattila VM, Kannus P, Rimpela A. The risk for a cruciate ligament injury of the knee in adolescents and young adults: a population-based cohort study of 46 500 people with a 9 year follow-up. Br J Sports Med 2008;42:422-6.
16. Houghton KM. Review for the generalist: evaluation of anterior knee pain. Pediatr Rheum online J 2007;5:8.
17. Fairbank JC, Pynsent PB, van Poortvliet JA, Phillips H. Mechanical factors in the incidence of knee pain in adolescents and young adults. J Bone Joint Surg Br 1984;66:685-93.
18. de Lucena GL, dos Santos Gomes C, Guerra RO. Prevalence and associated factors of Osgood-Schlatter syndrome in a population-based sample of Brazilian adolescents. Am J Sports Med 2011;39:415-20.
19. Witvrouw E, Lysens R, Bellemans J, Cambier D, Vanderstraeten G. Intrinsic risk factors for the development of anterior knee pain in an athletic population. A two-year prospective study. Am J Sports Med 2000;28:480-9.

20. Vahasarja V, Kinnuen P, Serlo W. Arthroscopy of the acute traumatic knee in children. Prospective study of 138 cases. Acta Orthop Scand 1993;64:580-2.
21. Nietosvaara Y, Aalto K, Kallio PE. Acute patellar dislocation in children: incidence and associated osteochondral fractures. J PediatrOrthop 1994;14:513-5.
22. Moksnes H, Engebretsen L, Risberg MA. The current evidence for treatment of ACL injuries in children is low: a systematic review. J Bone Joint Surg Am 2012;94:1112-9.
23. Caine D, Maffulli N, Caine C. Epidemiology of injury in child and adolescent sports: injury rates, risk factors, and prevention. Clin Sports Med 2008;27:19-50.
24. Atkin DM, Fithian DC, Marangi KS, Stone ML, Dobson BE, Mendelsohn C. Characteristics of patients with primary acute lateral patellar dislocation and their recovery within the first 6 months of injury. Am J Sports Med 2000;28:472-9.
25. Beasley LS, Vidal AF. Traumatic patellar dislocation in children and adolescents: treatment update and literature review. Curr Opin Pediatr 2004;16:29-36.
26. Fithian DC, Paxton EW, Stone ML, Silva P, Davis D, Elias DA, et al. Epidemiology and natural history of acute patellar dislocation. Am J Sports Med 2004;32:1114-21.
27. Nikku R, Nietosvaara Y, Aalto K, Kallio PE. Operative treatment of primary patellar dislocation does not improve medium-term outcome: A 7-year follow-up report and risk analysis of 127 randomized patients. Acta Orthop 2005;76:699-704.
28. Emery CA, Meeuwisse WH. The effectiveness of a neuromuscular prevention strategy to reduce injuries in youth soccer: a cluster-randomised controlled trial. Br J Sports Med 2010;44:555-62.
29. Hagglund M, Atroshi I, Wagner P, Walden M. Superior compliance with a neuromuscular training programme is associated with fewer ACL injuries and fewer acute knee injuries in female adolescent football players: secondary analysis of an RCT. Br J Sports Med 2013;47:974-9.
30. Walden M, Atroshi I, Magnusson H, Wagner P, Hagglund M. Prevention of acute knee injuries in adolescent female football players: cluster randomised controlled trial. BMJ 2012;344:e3042.
31. DiStefano LJ, Blackburn JT, Marshall SW, Guskiewicz KM, Garrett WE, Padua DA. Effects of an age-specific anterior cruciate ligament injury prevention program on lower extremity biomechanics in children. Am J Sports Med 2011;39:949-57.
32. Conaghan PG. Update on osteoarthritis part 1: current concepts and the relation to exercise. Br J Sports Med 2002;36:330-3.
33. Maffulli N, Longo UG, Gougoulias N, Loppini M, Denaro V. Long-term health outcomes of youth sports injuries. Br J Sports Med 2010;44:21-5.
34. Oiestad BE, Holm I, Engebretsen L, Aune AK, Gunderson R, Risberg MA. The prevalence of patellofemoral osteoarthritis 12 years after anterior cruciate ligament reconstruction. Knee Surg, Sports Traumatol, Arthrosc 2013;21:942-9.
35. Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med 2000;133:635-46.
36. Lohmander LS, Ostenberg A, Englund M, Roos H. High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. Arthritis Rheum 2004;50:3145-52.
37. Bout-Tabaku S, Best TM. The adolescent knee and risk for osteoarthritis - an opportunity or responsibility for sport medicine physicians? Curr Sports Med Rep 2010;9:329-31.

38. Hinton RY, Rivera VR, Pautz MJ, Sponseller PD. Ligamentous laxity of the knee during childhood and adolescence. J Pediatr Orthop 2008;28:184-7.
39. Ageberg E, Bennell KL, Hunt MA, Simic M, Roos EM, Creaby MW. Validity and inter-rater reliability of medio-lateral knee motion observed during a single-limb mini squat. BMC Musculoskelet Disord 2010;11:265-73
40. Livingston LA. The quadriceps angle: a review of the literature. J Orthop Sports Phys Ther 1998;28:105-9.
41. Jones, Stratton G. Muscle function assessment in children. Acta Paediatr 2000;89:753-61.
42. Noyes FR, Barber-Westin SD, Fleckenstein C, Walsh C, West J. The drop-jump screening test: difference in lower limb control by gender and effect of neuromuscular training in female athletes. Am J Sports Med 2005;33:197-207.
43. Gustavsson A, Neeter C, Thomee P, Silbernagel KG, Augustsson J, Thomee R, et al. A test battery for evaluating hop performance in patients with an ACL injury and patients who have undergone ACL reconstruction. Knee Surg, Sports Traumatol, Arthrosc 2006;14:778-88.
44. Reid A, Birmingham TB, Stratford PW, Alcock GK, Giffin JR. Hop testing provides a reliable and valid outcome measure during rehabilitation after anterior cruciate ligament reconstruction. Phys. Ther. 2007;87:337-49.
45. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. BMJ 2010;340:464-7
46. Speight J, Barendse SM. FDA guidance on patient reported outcomes. BMJ 2010;340:c2921
47. Snyder-Mackler L, Fitzgerald GK, Bartolozzi AR, Ciccotti MG. The relationship between passive joint laxity and functional outcome after anterior cruciate ligament injury. Am J Sports Med 1997;25:191-5.
48. Greenberger HB, Paterno MV. Relationship of knee extensor strength and hopping test performance in the assessment of lower extremity function. J Orthop Sports Phys Ther 1995;22:202-6.
49. Holm I, Fredriksen P, Fosdahl M, Vollestad N. A normative sample of isotonic and isokinetic muscle strength measurements in children 7 to 12 years of age. Acta Paediatr 2008;97:602-7.
50. Ostenberg A, Roos E, Ekdahl C, Roos H. Isokinetic knee extensor strength and functional performance in healthy female soccer players. Scand J Med Sci Sports 1998;8:257-64.
51. Bremander AB, Dahl LL, Roos EM. Validity and reliability of functional performance tests in meniscectomized patients with or without knee osteoarthritis. Scand J Med Sci Sports 2007;17:120-7.
52. Ageberg E, Thomee R, Neeter C, Silbernagel KG, Roos EM. Muscle strength and functional performance in patients with anterior cruciate ligament injury treated with training and surgical reconstruction or training only: a two to five-year followup. Arthritis Rheum 2008;59:1773-9.
53. Hewett TE, Myer GD, Ford KR, Heidt RS, Colosimo AJ, McLean SG, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. Am J Sports Med 2005;33:492-501.
54. Beenakker EA, Maurits NM, Fock JM, Brouwer OF, van der Hoeven JH. Functional ability and muscle force in healthy children and ambulant Duchenne muscular dystrophy patients. Eur J Paediatr Neurol 2005;9:387-93.
55. Eek MN, Tranberg R, Beckung E. Muscle strength and kinetic gait pattern in children with bilateral spastic CP. Gait & posture 2011;33:333-7.

56. Harbo T, Brincks J, Andersen H. Maximal isokinetic and isometric muscle strength of major muscle groups related to age, body mass, height, and sex in 178 healthy subjects. Eur J Appl Phys 2012;112:267-75.
57. Eek MN, Kroksmark AK, Beckung E. Isometric muscle torque in children 5 to 15 years of age: normative data. Arch Phys Med Rehabil 2006;87:1091-9.
58. Taylor NF, Dodd KJ, Graham HK. Test-retest reliability of hand-held dynamometric strength testing in young people with cerebral palsy. Arch Phys Med Rehabil 2004;85:77-80.
59. Lu TW, Hsu HC, Chang LY, Chen HL. Enhancing the examiner's resisting force improves the reliability of manual muscle strength measurements: comparison of a new device with hand-held dynamometry. J Rehabil Med 2007;39:679-84.
60. Wikholm JB, Bohannon RW. Hand-Held Dynamometer Measurements: Tester strength makes a difference. J Orthop Sports Phys Ther 1991;13:191-8.
61. Tsiros MD, Grimshaw PN, Shield AJ, Buckley JD. The Biodex isokinetic dynamometer for knee strength assessment in children: advantages and limitations. Work 2011;39:161-7.
62. Lund H, Sondergaard K, Zachariassen T, Christensen R, Bulow P, Henriksen M, et al. Learning effect of isokinetic measurements in healthy subjects, and reliability and comparability of Biodex and Lido dynamometers. Clin Physiol Funct Imaging 2005;25:75-82.
63. Maffiuletti NA, Bizzini M, Desbrosses K, Babault N, Munzinger U. Reliability of knee extension and flexion measurements using the Con-Trex isokinetic dynamometer. Clin Physiol Funct Imaging 2007;27:346-53.
64. De Ste CM, Deighan M, Armstrong N. Assessment and interpretation of isokinetic muscle strength during growth and maturation. Sports Med 2003;33:727-43.
65. Örtqvist M, Gutierrez-Farewik E, Farewik M, Jansson A, Bartonek, Broström E. Reliability of a new instrument for measuring plantarflexor muscle strength Arch Phys Med Rehabil 2007;88:1164-70.
66. Andriacchi TP. Dynamics of knee malalignment. Orthop Clin North Am 1994;25:395-403.
67. Hunter DJ, Sharma L, Skaife T. Alignment and osteoarthritis of the knee. J Bone Joint Surg Am 2009;91 Suppl 1:85-9.
68. Rauh MJ, Koepsell TD, Rivara FP, Rice SG, Margherita AJ. Quadriceps angle and risk of injury among high school cross-country runners. J Orthop Sports Phys Ther 2007;37:725-33.
69. Tomsich DA, Nitz AJ, Threlkeld AJ, Shapiro R. Patellofemoral alignment: reliability. J Orthop Sports Phys Ther 1996;23:200-8.
70. Trulsson A, Roos EM, Ageberg E, Garwicz M. Relationships between postural orientation and self reported function, hop performance and muscle power in subjects with anterior cruciate ligament injury. BMC Musculoskelet Disord 2010;11:143-51.
71. Barber-Westin SD, Smith ST, Campbell T, Noyes FR. The Drop-Jump Video Screening Test: Retention of Improvement in Neuromuscular Control in Female Volleyball Players. J Strength Cond Res 2010;24:3055-62.
72. Smith TO, Hunt NJ, Donell ST. The reliability and validity of the Q-angle: a systematic review. Knee Surg Sports Traumatol Arthrosc 2008;16:1068-79.
73. Woodland LH, Francis RS. Parameters and comparisons of the quadriceps angle of college-aged men and women in the supine and standing positions. Am J Sports Med 1992;20:208-11.
74. Bayraktar B, Yucesir I, Ozturk A, Cakmak AK, Taskara N, Kale A, et al. Change of quadriceps angle values with age and activity. Saudi Med J 2004;25:756-60.

75. Shultz SJ, Nguyen AD, Schmitz RJ. Differences in lower extremity anatomical and postural characteristics in males and females between maturation groups. J Orthop Sports Phys Ther 2008;38:137-49.
76. Roos EM, Ostenberg A, Roos H, Ekdahl C, Lohmander LS. Long-term outcome of meniscectomy: symptoms, function, and performance tests in patients with or without radiographic osteoarthritis compared to matched controls. Osteoarthritis Cartilage 2001;9:316-24.
77. Streiner D, Norman GR. Health measurement scales - a practical guide to their development and use, Fourth ed. New York: Oxford University press 2008.
78. The EuroQol group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
79. Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res 2010;19:875-86.
80. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS) - development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88-96.
81. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the international knee documentation committee subjective knee form. Am J Sports Med 2001;29:600-13.
82. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res 1985:43-9.
83. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med 1982;10:150-4.
84. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). Arthritis Care Res 2011;63 Suppl 11:S208-28.
85. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation of a Swedish version. Scand J Med Sci Sports 1998;8:439-48.
86. Roos EM, Engelhart L, Ranstam J, Anderson AF, Irrgang JJ, Marx RG, et al. ICRS Recommendation document: Patient-reported outcome instruments for use in patients with articular cartilage defects. Cartilage 2011;2:153-72.
87. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17-27.
88. Iversen MD, Lee B, Connell P, Andersen J, Anderson AF, Kocher MS. Validity and comprehensibility of the International Knee Documentation Committee Subjective Knee Evaluation form in Children. Scand J Med Sci Sports 2010;20:87-95.
89. Kocher MS, Smith JT, Iversen MD, Brustowicz K, Ogunwole O, Andersen J, et al. Reliability, validity, and responsiveness of a modified International Knee Documentation Committee Subjective Knee Form (Pedi-IKDC) in children with knee disorders. Am J Sports Med 2011;39:933-9.
90. Trama A, Dieci M. Quality of life in clinical trials for children. Eur J Clin Pharmacol 2011;67 Suppl 1:41-7.

91. Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. Health Qual Life Outcomes 2007;5:1-13.
92. Bruton A, Conway H, Holgate ST. Reliability; What is it, and how is it measured? Phys Ther 2000;86:94-9.
93. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
94. De Vet H, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. New York: Cambride University Press; 2013.
95. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539-49.
96. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737-45.
97. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420-8.
98. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
99. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil 1998;12:187-99.
100. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC medical research methodology 2010;10:22-30.
101. DeMaio TJ, Rothgeb JM. Cognitive interviewing techniques: in the lab and in the field. In: Schwarz N, Sudman S, eds. Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research. San Francisco, CA: Jossey-Bass Publishers; 1995:177-95.
102. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. J AdvNurs 2003;42:57-63.
103. Willis GB. Cognitive interviewing: A tool for improving questionnaire design. Thousand Oaks: Sage Publications; 2005.
104. Goldbloom RB, Kim RK, Hodder-Malloy C, et al. Design and reliability of pediatric HealthQuiz: preliminary report of a comprehensive, computerized, self-administered child health assessment. Clin Pediatr 1999;38:645-54.
105. Irwin DE, Varni JW, Yeatts K, DeWalt DA. Cognitive interviewing methodology in the development of a pediatric item bank: a patient reported outcomes measurement information system (PROMIS) study. Health Qual Life Outcomes 2009;7:3-13.
106. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349-62.
107. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407-15.
108. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Exp Rev Pharmacoeconomics Outcomes Res 2011;11:171-84.

109. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Clin Epidemiol 2002;55:900-8.
110. Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 2010;63:760-6.
111. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed: Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
112. Wang D, Jones MH, Khair MM, Miniaci A. Patient-reported outcome measures for the knee. J Knee Surg 2010;23:137-51.
113. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.
114. Lydick E, Epstein RS. Interpretation of quality of life changes. Qual Life Res 1993;2:221-6.
115. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol 2010;63:524-34.
116. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life outcomes 2006;4:54.
117. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res 2001;10:571-8.
118. Kovacs FM, Abraira V, Royuela A, et al. Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. BMC musculoskelet Disord 2008;9:43-52.
119. Terwee CB, Roorda LD, Knol DL, De Boer MR, De Vet HC. Linking measurement error to minimal important change of patient-reported outcomes. J Clin Epidemiol 2009;62:1062-7.
120. Tourangeau R, Rips LJ, Rasinski K. The psychology of survey response. Cambridge: Cambridge University Press; 2002.
121. Willis GB, Royston P, Bercini D. The use of verbal report methods in the development and testing of survey questionnaires. Applied Cognitive Psychology 1991;5:251-67.
122. Maynard DW, Houtkoop-Steenstra H, Schaeffer NC, van der Zouwen J. Standardization and tacit knowledge: Interaction and practice in the survey interview. New York: John Wiley & Sons; 2002.
123. Oksenberg L, Canell C, Kalton G. New strategies for pretesting survey questions. Journal of official statistics 1991;7:349-65.
124. Tourangeau R. Cognitive sciences and survey methods. In: Jabine T, Straf M, Tanur J, Tourangeau R, eds. Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines. Washington DC: National Academy Press; 1984:73-100.
125. Andersson GB, Ruperto N, Berg S, Hagelberg S, Jonsson NO, Magnusson B, et al. The Swedish version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ). Clin Exp Rheumatol 2001;19:146-50.
126. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. Qual Life Res 2010;19:887-97.

127. Jobe J, Tourangeau R, Smith A. Contributions of survey research to the understanding of memory. Applied Cognitive Psychology 1993;7:567-84.
128. Lessler JT, Sirken MG. Laboratory-based research on the cognitive aspects of survey methodology: The goals and methods of the national center of health statistics study. Milbank Memorial Fund Quarterly/ Health and Society 1985;63 565-81.
129. Insall J, Falvo KA, Wise DW. Chondromalacia Patellae. A prospective study. J Bone Joint Surg Am 1976;58:1-8.
130. Schmitt J, Di Fabio RP. The validity of prospective and retrospective global change criterion measures. Arch Phys Med Rehabil 2005;86:2270-6.
131. Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley & Sons1986.
132. de Vet HC, Bouter LM, Bezemer PD, Beurskens AJ. Reproducibility and responsiveness of evaluative outcome measures. Theoretical considerations illustrated by an empirical example. Int J Technol Assess Health Care 2001;17:479-87.
133. de Vet HC, Ader HJ, Terwee CB, Pouwer F. Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. Qual Life Res 2005;14:1203-18.
134. Norman GR, Streiner DL. PDQ Statistics. Toronto: BC Decker Inc; 2003.
135. Juniper EF, Jaeschke R. How to develop and validate a new health-related quality of life instrument. Quality of Life and Pharmacoeconomics in Clinical Trials. Spilker B. Lippincott-Raven Philadelphia1996.
136. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102-9.
137. Slobogean GP, Mulpuri K, Reilly CW. The International Knee Documentation Committee Subjective Evaluation Form in a preadolescent population: pilot normative data. Am J Sports Med 2008;36:129-32.
138. Kramer LC, Denegar CR, Buckley WE, Hertel J. Factors associated with anterior cruciate ligament injury: history in female athletes. J Sports Med Phys Fitness 2007;47:446-54.
139. Loudon JK, Jenkins W, Loudon KL. The relationship between static posture and ACL injury in female athletes. J OrthopSports PhysTher 1996;24:91-7.
140. Augustsson J, Thomee R, Karlsson J. Ability of a new hop test to determine functional deficits after anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 2004;12:350-6.
141. Augustsson J, Thomee R, Linden C, Folkesson M, Tranberg R, Karlsson J. Single-leg hop testing following fatiguing exercise: reliability and biomechanical analysis. Scand J Med Sci Sports 2006;16:111-20.
142. Willis GB, DeMaio TJ, Harris-Kojetin B. Is the bandwagon headed to the methodological promised land? Evaluation of the validity of cognitive interviewing techniques. In: Sirken M, Herrmann D, Schechter S, Schwarz N, Tanur J, Tourangeau R, eds. Cognition and Survey Research. New York Wiley; 1999.
143. Angst F. The new COSMIN guidelines confront traditional concepts of responsiveness. BMC Med Res Methodol 2011;11:152-58.
144. Moksnes H, Engebretsen L, Eitzen I, Risberg MA. Functional outcomes following a non-operative treatment algorithm for anterior cruciate ligament injuries in skeletally immature children 12 years and younger. A prospective cohort with 2 years follow-up. Br J Sports Med 2013;47:488-94.

145. Moksnes H. Functional and radiological outcomes following a non-operative treatment algorithm after ACL injuries in skeletally immature children, Ph.D. Thesis, Norwegian School of Sport Sciences, 2013.
146. Granan LP, Forssblad M, Lind M, Engebretsen L. The Scandinavian ACL registries 2004-2007: baseline epidemiology. Acta Orthop 2009;80:563-7.
147. Paradowski PT, Bergman S, Sunden-Lundius A, Lohmander LS, Roos EM. Knee complaints vary with age and gender in the adult population. Population-based reference data for the Knee injury and Osteoarthritis Outcome Score (KOOS). BMC Musculoskelet Disord 2006;7:38-46.
148. Medical Subject Headings (MeSH), U.S National Library of Medicine (NLM), (Jan 2014) http://www.nlm.nih.gov/mesh/meshhome.html
149. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-68.
150. Blair J, Srinath KP. A note on sample size for behavior coding pretest. Field methods 2008;20:85-95.
151. Sirken M, Herrmann D, Schechter S, Schwarz N, Tanur J, Tourangeau R. Cognition and survey research. New York: Wiley; 1999.
152. Patton MQ. Qualitative evaluation and research methods. Newbury Park: Sage; 1994.
153. Sandelowski M. Sample size in qualitative research. Res Nurs Health 1995;18:179-83.
154. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002;11:193-205.
155. Valderas JM, Ferrer M, Mendivil J, et al. Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. Value Health 2008;11:700-8.
156. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing: Washington, DC: American Educational Research Association; 1999.
157. Conrad KJ, Smith EV. International conference on objective measurement: applications of Rasch analysis in health care. Medical care 2004;42:1-6.
158. World Health Organization (WHO) International Classification of Functioning, Disability and Health (ICF) Geneva, 2001.
159. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. J Rehabil Med 2005;37:212-8.
160. Lollar DJ, Simeonsson RJ. Diagnosis to function: classification for children and youths. J Dev Behav Pediatr 2005;26:323-30.

10 APPENDICES

A. *KOOS-Child User's Guide*

B. *KOOS-Child LK2.0 (final version) in Swedish*

C. *KOOS-Child LK2.0 (final version) in English*

D. *EQ-5D-Y* [79, 126] (shortened and modified for use in Study IV)

E. *Child Health Assessment Questionnaire (CHAQ)*[125] (shortened and modified for use in Study IV)

F. *Visual Analogue Scale (VAS) formatted questions of knee function* (specifically designed for Study IV)

G. *Global Perceived Effect (GPE) Scores* (specifically designed for Study IV)

A

A User's Guide to:

The Knee injury and Osteoarthritis Outcome Score for children

KOOS-Child

KOOS-Child is a patient-reported outcome measure employing five-item Likert scales. KOOS-Child covers 5 dimensions (subscales): Pain, Symptoms (titled “Knee problems” in the KOOS-Child), Difficulty during daily activities (ADL), Function in sport and play (Sports/Play) and knee-related Quality of Life (QOL).

The psychometric testing of the preliminary version of the KOOS-Child (LK 1.0) [1] has been completed and a final version KOOS-Child (LK 2.0) is now available [2]. 9 items (P5, P7, P8b, A4, A6, A8, A9, A11, A15) in the 1.0 version have been deleted due to high ceiling effects. Thus, 39 items have been kept in the 2.0 version. Since items have been deleted, but no items have been added, you can always calculate KOOS-Child scores from the 1.0 version. Please note that there are two KOOS-Child scoring files (excel) available, one for use when data was collected using the preliminary KOOS-Child questionnaire form (LK 1.0) and one for use when data was collected with the final KOOS-Child (LK 2.0) questionnaire form. When the scoring file for KOOS Child LK 1.0 is used the result is automatically converted to the LK 2.0 scores.

The structure of the 5 KOOS-Child subscales as well as the numbering of the items has been kept in the final version (LK 2.0). The subscales are scored separately as previously: Pain (8 items); Symptoms (7 items); ADL (11 items); Sport/Rec (7 items); and QOL (6 items). Standardized answer options are given (5 Likert boxes) and each question gets a score from 0 to 4, where 0 indicates no problem. The five scores are calculated as the sum of the items included, in accordance with score calculations of the KOOS score***.*** Raw scores are then transformed to a 0-100 scale, with zero representing extreme knee problems and 100 representing no knee problems, as common in orthopedic scales. Scores between 0 and 100 represent the percentage of total possible score achieved. An aggregate score is not calculated since it is regarded desirable to analyze and interpret the different dimensions separately.

KOOS-Child (LK 2.0) Scoring instructions

Assign the following scores to the boxes:

None	A little	Some	A lot	Extreme
0	1	2	3	4

Missing data: If a mark is placed outside a box, the closest box is chosen. If two boxes are marked, that which indicated the more severe problems is chosen.

As long as at least 50% of the subscale items are answered for each subscale, a mean score can be calculated. If more than 50% of the subscales items are omitted, the response is considered invalid and no subscale score is calculated. For the subscale pain this means that 4 items must be answered, for symptoms 4 items, for ADL 6 items, for Sport/Play 4 items and for QOL 3 items must be answered to calculate a subscale score. Subscale scores are independent and can be reported for any number of the individual subscales, i.e. if a particular subscale is not considered valid (for example the subscale Sport/Play 2 weeks after ACL reconstruction) the results from the other subscale can be reported at this time point.

Guidelines on applicability of subscales and items: It is important to determine whether or not each subscale is relevant at the time point chosen, considering the specific study population. For example, Difficulty with Sports/Play function may not be relevant to assess 2 weeks post-operatively.

Pain and ADL subscales: If a subject avoids an activity (e.g. twisting/pivoting or going up or down stairs) due to doctor's order or because the subject has chosen to avoid the activity, the subject should be instructed to choose "(4) Extreme" for those items.

Sports/Play subscale: The same as above. Also, if a subject does not normally engage in an activity (e.g. running or jumping), the subject should be instructed to leave the item blank.

Score calculation: Apply the mean of the observed items within the subscale (e.g. KOOS-Child Pain), divide with 4, and multiply with 100; when this number is then subtracted from 100 you have the KOOS-Child subscale estimate for that particular cross-sectional assessment of the individual patient. Manual scoring formulas as well as excel formulas are provided below. Excel spreadsheets are available at koos.nu For manual calculations, please use the formulas provided below for each subscale:

1. PAIN $100 - \frac{\text{Mean Score (P1−P9)}\times 100}{4} = KOOS - child\ Pain$

2. SYMPTOMS $100 - \frac{\text{Mean Score (S1−S7)}\times 100}{4} = KOOS - child\ Symptoms$

3. ADL $100 - \frac{\text{Mean Score (A1−A17)}\times 100}{4} = KOOS - child\ ADL$

4. SPORT/PLAY $100 - \frac{\text{Mean Score (SP1−SP7)}\times 100}{4} = KOOS - child\ Sport/Play$

5. QOL $100 - \frac{\text{Mean Score (Q1−QN6)}\times 100}{4} = KOOS - child\ QOL$

KOOS-Child (LK 2.0) Excel scoring files

Please note that there are TWO KOOS-Child scoring files (excel) available, one for use when data was collected using the preliminary KOOS-Child questionnaire form (LK 1.0) and one for use when data was collected with the final KOOS-Child (LK 2.0) questionnaire form. When the scoring file for KOOS-Child LK 1.0 is used the result is automatically converted to the LK 2.0 scores. Excel spreadsheets with formulas to calculate the five subscale scores are available from www.koos.nu. If you for any reason prefer to use your own spreadsheets, the excel formulas for KOOS-Child LK 2.0 are given below.

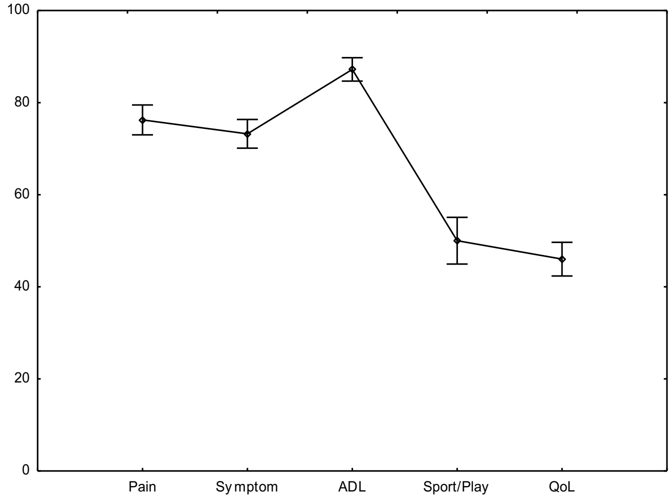
Excel formulation: When the raw data has been entered in the order the items occur in the KOOS questionnaires available from koos.nu, these excel formulations can be copy-pasted directly into an English version of an excel spreadsheet to automatically calculate the five sub score scales. Please note that it has been considered that the items in the subscale symptoms appear first in the questionnaire.

KOOS-Child Pain:	=100-AVERAGE(I2:P2)/4*100
KOOS-Child Symptoms:	=100-AVERAGE(B2:H2)/4*100
KOOS-Child ADL:	=100-AVERAGE(Q2:AA2)/4*100
KOOS-Child Sport/Play:	=100-AVERAGE(AB2:AH2)/4*100
KOOS-Child QOL:	=100-AVERAGE(AI2:AN2)/4*100

KOOS-Child Profile

To visualize differences in the five different KOOS-Child sub scores and change between different administrations of the KOOS-Child (e.g. pre-treatment to post-treatment), KOOS-Child Profiles can be plotted.

The profile below shows the mean subscale scores ±95 % Confidence Intervals for a cohort of 115 children (56% girls, aged 7-16) suffering from a knee injury (e.g. ACL injury, Patellar dislocation, Osgood-Schlatter lesion) and treated in primary and secondary care at different time points from injury/onset.



Reference

1. Örtqvist M, Roos EM, Brostrom EW, Janarv P-M, Iversen MD.
Development of the Knee Injury and Osteoarthritis Outcome Score for Children (KOOS-Child): Comprehensibility and Content Validity. Acta Orthop 2012;83:666-73.

2. Örtqvist M, Iversen MD, Janarv P-M, Brostrom EW, Roos EM.
Psychometric properties of the Knee injury and Osteoarthritis Outcome Score for Children (KOOS-Child) in children with knee disorders. Submitted October 2013.

B

KOOS-Child knäformulär

DATUM: ___________________ PERSONNUMMER: ____________________

NAMN: ____

INSTRUKTIONER

De här frågorna handlar om hur ditt skadade knä påverkar dig. Svara på varje fråga genom att kryssa för det alternativ du tycker är bäst (endast ett alternativ per fråga). Om du är osäker, kryssa ändå för det alternativ som känns riktigast.

KNÄPROBLEM

S1. Hur ofta har knät varit svullet de senaste 7 dagarna?
Aldrig Sällan Ibland Ofta Alltid

S2. Hur ofta har du hört något ljud från knät de senaste 7 dagarna?
Aldrig Sällan Ibland Ofta Alltid

S3. Hur ofta har ditt knä hakat upp sig (fastnat) de senaste 7 dagarna?
Aldrig Sällan Ibland Ofta Alltid

S4. Hur ofta har du, utan hjälp, kunnat sträcka knät helt de senaste 7 dagarna?
Alltid Ofta Ibland Sällan Aldrig

S5. Hur ofta har du, utan hjälp, kunnat böja knät helt de senaste 7 dagarna?
Alltid Ofta Ibland Sällan Aldrig

S6. Hur svårt har du haft att röra på knät när du vaknat på morgonen de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

S7. Hur svårt har du haft att röra på knät om du varit stilla en stund senare under dagen de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

P1. Hur ofta har du haft ont i knät den senaste månaden?
Aldrig Sällan Ibland Ofta Hela tiden

HUR ONT

Hur ont har du haft i knät när du har gjort följande aktiviteter de **senaste 7 dagarna**?
Kryssa för det bästa svarsalternativet för varje fråga

	Inte ont	Lite ont	Ganska ont	Mycket ont	Extremt ont
P2. Snurra/vrida på det skadade knät när du går/står/springer					
P3. Sträcka fullt på ditt skadade knä					
P4. Böja fullt på ditt skadade knä					
P6a. Gå upp för trappor					
P6b. Gå ner för trappor					
P8a. Sitta med ditt skadade knä böjt					
P9. Stå på båda benen, oberoende av hur länge					

SVÅRIGHETER VID VARDAGSAKTIVITETER

A1. Hur svårt har du haft att gå ner för trappor de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A2. Hur svårt har du haft att gå upp för trappor de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A3. Hur svårt har du haft att resa dig från en stol de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A5. Hur svårt har du haft att böja dig ned och plocka upp något från golvet de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A7. Hur svårt har du haft att gå i/ur en bil de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A10. Hur svårt har du haft att resa dig från sängen de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A12. Hur svårt har du haft att ändra läge på knät när du har legat i sängen de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A13. Hur svårt har du haft att gå i/ur badkaret/duschen de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A14. Hur svårt har du haft att sitta på en stol med ditt skadade knä böjt de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A16. Hur svårt har du haft att bära tunga väskor, ryggsäck eller liknande de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

A17. Hur svårt har du haft att bädda sängen, städa ditt rum, plocka i/ur diskmaskin eller liknande de senaste 7 dagarna?
Inte alls svårt Lite Ganska Mycket Extremt svårt

SVÅRIGHETER VID LEK OCH IDROTT

SP1. Hur svårt har du haft att gå ner på huk när du har lekt eller idrottat under de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	
SP2. Hur svårt har du haft att springa när du har lekt eller idrottat under de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	
SP3. Hur svårt har du haft att hoppa när du har lekt eller idrottat under de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	
SP4. Hur svårt har du haft att snurra/vrida på det skadade knät när du har lekt eller idrottat under de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	

SP5. Hur svårt har du haft att sitta på knä under de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	
SPN6. Hur svårt har du haft att hålla balansen när du har gått/sprungit på ojämn mark de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	
SPN7. Hur svårt har du haft att vara med på sportaktiviteter på grund av din knäskada under de senaste 7 dagarna? Inte alls svårt Lite Ganska Mycket Extremt svårt	

HUR HAR DIN KNÄSKADA PÅVERKAT DITT LIV?

Q1. Hur ofta tänker du på ditt skadade knä?
Aldrig Sällan Ibland Ofta Hela tiden

Q2. Hur mycket har du ändrat ditt sätt att leva på grund av ditt skadade knä?
Inte alls Lite Ganska Mycket Väldigt mycket

Q3. Hur mycket kan du lita på ditt skadade knä?
Helt och hållet Mycket Ganska Lite Inte alls

Q4. Hur mycket problem har du med ditt skadade knä över huvudtaget?
Inga alls Små Mitt i mellan Stora Mycket stora

QN5. Hur svårt har du haft att ta dig till eller runt i skolan (gå i trappor, öppna dörrar, bära böcker, vara med på rasten) på grund av ditt skadade knä?
Inte alls svårt Lite Ganska Mycket Extremt svårt

QN6. Hur svårt har du haft att göra saker med vänner på grund av ditt skadade knä?
Inte alls svårt Lite Ganska Mycket Extremt svårt

Tack för att Du tagit dig tid att besvara samtliga frågor!

C

KOOS-Child KNEE SURVEY

Today's date: ___________________ Date of birth: ___________________

Name: ____

INSTRUCTIONS

These questions collect information about how your injured knee affects you. Answer every question by ticking the appropriate box, only one box for each question. If you are unsure about how to answer a question, please select the best answer you can.

KNEE PROBLEMS

S1. During the past 7 days, how often has your knee been swollen?

Never	Rarely	Sometimes	Often	Always

S2. During the past 7 days, how often has your knee made any noise/sounds?

Never	Rarely	Sometimes	Often	Always

S3. During the past 7 days, how often did your knee get stuck?

Never	Rarely	Sometimes	Often	Always

S4. During the past 7 days, how often have you been able to fully straighten your knee on your own?

Always	Often	Sometimes	Rarely	Never

S5. During the past 7, days how often have you been able to fully bend your knee on your own?

Always	Often	Sometimes	Rarely	Never

S6. During the past 7 days, how much difficulty have you had moving your knee just after waking up in the morning?

No difficulty	A little	Some	A lot	Extreme difficulty

S7. During the past 7 days, how much difficulty have you had later in the day moving your knee after being sedentary for a while?

None	A little	Some	A lot	Extreme

P1. During the past month, how often have you experienced knee pain?

Never	Rarely	Sometimes	Often	All the time

HOW PAINFUL

How much knee pain have you experienced **in the past 7 days** during the following activities? Check the best answer for each item

	No pain	A little pain	Some pain	A lot of pain	Extreme pain
P2. Twisting/pivoting on your injured knee when walking/standing/running					
P3. Fully straightening your injured knee					
P4. Fully bending your injured knee					
P6a. Walking up stairs					
P6b. Walking down stairs					
P8a. Sitting with your injured knee bent					
P9. Standing upright on both legs for any amount of time					

DIFFICULTY DURING DAILY ACTIVITIES

A1. During the past 7 days, how much difficulty have you had walking down stairs?
No difficulty A little Some A lot Extreme difficulty

A2. During the past 7 days, how much difficulty have you had walking up stairs?
No difficulty A little Some A lot Extreme difficulty

A3. During the past 7 days, how much difficulty have you had standing up from a chair?
No difficulty A little Some A lot Extreme difficulty

A5. During the past 7 days, how much difficulty have you had to bend down and pick up an object from the floor?
No difficulty A little Some A lot Extreme difficulty

A7. During the past 7 days, how much difficulty have you had getting in to/out of a car?
No difficulty A little Some A lot Extreme difficulty

A10. During the past 7 days, how much difficulty have you had to get out of bed?
No difficulty A little Some A lot Extreme difficulty

A12. During the past 7 days, how much difficulty have you had to change knee position when lying in bed?
No difficulty A little Some A lot Extreme difficulty

A13. During the past 7 days, how much difficulty have you had getting in to/out of the bathtub/shower?
No difficulty A little Some A lot Extreme difficulty

A14. During the past 7 days, how much difficulty have you had to sit in a chair with your injured knee bent?
No difficulty A little Some A lot Extreme difficulty

A16. During the past 7 days, how much difficulty have you had to carry heavy bags /backpacks etc?
No difficulty A little Some A lot Extreme difficulty

A17. During the past 7 days, how much difficulty have you had to do light chores such as cleaning your room, filling/emptying the dishwasher, making your bed, etc?
No difficulty A little Some A lot Extreme difficulty

DIFFICULTY DURING SPORTS AND PLAYING

SP1. During the past 7 days, how much difficulty have you had to squat down during play or sports activities? No difficulty A little Some A lot Extreme difficulty	
SP2. During the past 7 days, how much difficulty have you had to run during play or sports activities? No difficulty A little Some A lot Extreme difficulty	
SP3. During the past 7 days, how much difficulty have you had to jump during play or sports activities? No difficulty A little Some A lot Extreme difficulty	
SP4. During the past 7 days, how much difficulty have you had to twist/pivot because of your injured knee during play or sports activities? No difficulty A little Some A lot Extreme difficulty	

SP5. During the past 7 days, how much difficulty have you had to kneel because of your injured knee? No difficulty A little Some A lot Extreme difficulty	
SPN6. During the past 7 days, how much difficulty have you had to keep your balance when walking /running on uneven ground? No difficulty A little Some A lot Extreme difficulty	
SPN7. During the past 7 days, how much difficulty have you had playing sports because of your injured knee? No difficulty A little Some A lot Extreme difficulty	

HOW HAS YOUR INJURY AFFECTED YOUR LIFE?

Q1. How often do you think about your knee problem?

Never   Rarely   Sometimes   Often   All the time

Q2. How much have you changed your lifestyle because of your injured knee?

Not at all   A little   Some   A lot   Very much

Q3. How much do you trust your injured knee?

Not at all   A little   Some   A lot   Completely

Q4. Overall, how much difficulty do you have with your injured knee?

No difficulty   A little   Some   A lot   Extreme difficulty

QN5. How much difficulty have you had getting to school or walking around in school (climbing stairs, opening doors, carrying books, participating during recess) because of your injured knee?

No difficulty   A little   Some   A lot   Extreme difficulty

QN6. How much difficulty have you had to do things with friends because of your injured knee?

No difficulty   A little   Some   A lot   Extreme difficulty

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

D

Shortened and modified version of the EQ-5D-Y[79, 126] used in Study IV

EQ-5D-Y

Hur är din hälsa IDAG?

Sätt ett kryss i den ruta som bäst beskriver din hälsa IDAG

Kunna röra sig

- Jag har inte svårt att gå ❑
- Jag har lite svårt att gå ❑
- Jag har mycket svårt att gå ❑

Ta hand om mig själv

- Jag har inte svårt att tvätta mig eller klä på mig själv ❑
- Jag har lite svårt att tvätta mig eller klä på mig själv ❑
- Jag har mycket svårt att tvätta mig eller klä på mig själv ❑

Göra vanliga aktiviteter (*till exempel gå i skolan, sport-och fritidsaktiviteter, lek, göra saker med familj eller kompisar*)

- Jag har inte svårt att göra mina vanliga aktiviteter ❑
- Jag har lite svårt att göra mina vanliga aktiviteter ❑
- Jag har mycket svårt att göra mina vanliga aktiviteter ❑

Ha ont eller ha besvär

- Jag har inte ont eller några besvär ❑
- Jag har lite ont eller lite besvär ❑
- Jag har mycket ont eller mycket besvär ❑

Känna sig orolig, ledsen eller olycklig

- Jag är inte orolig, ledsen eller olycklig ❑
- Jag är lite orolig, ledsen eller olycklig ❑
- Jag är mycket orolig, ledsen eller olycklig ❑

E

Shortened and modified version of the Child Health Assessment Questionnaire (CHAQ)[125] used in Study IV

CHAQ formulär för Barn/Tonåringar

Med detta formulär vill vi försöka få en uppfattning om hur sjukdomen påverkar Dig i det dagliga livet. I de följandefrågorna markeras med ett kryss det svar som bäst beskriver Din förmåga **UNDER DEN GÅNGNA VECKAN. NOTERA ENDAST SVÅRIGHETER SOM BEROR PÅ SJUKDOM.** Den sista kolumnen, ”inte aktuellt” ska användas när Du är för liten att klara den uppgiften. När det däremot gäller en aktivitet som Du inte klarar p.g.a. sjukdom, markera i kolumnen ”omöjligt att utföra”. Det finns även frågor som inte rör ditt knä på denna sida. Svara även på dessa.

	Utan svårighet	Med viss svårighet	Med stor svårighet	Omöjlig att utföra	Inte aktuellt
KOD	0	1	2	3	4
RESA SIG					
Hur klarar Du att:					
- Resa Dig från en låg stol eller från golvet?					
- Kliva ur sängen ?					
GÅ					
Hur Klarar Du att:					
- Gå utomhus på plan mark ?					
- Gå uppför 5 trappsteg ?					

GREPPFUNKTION					
Hur klarar DU att:					
- Skriva med penna/kladda med krita ?					
- Öppna en bildörr ?					
- Öppna en burk som redan varit öppnad ?					
- Vrida på och stänga av en vattenkran ?					
- Trycka ned ett dörrhandtag ?					
AKTIVITETER					
- Springa och leka ?					

SMÄRTA:
Vi vill också veta hur mycket värk Du haft p. g. a. Din sjukdom. Hur mycket ont har Du haft p.g.a. Din sjukdom UNDER DEN GÅNGNA VECKA?
Sätt ett X på linjen nedan för att visa hur ont Du har haft.

Ingen värk ____ Mycket svår värk

F

Visual Analogue Scale formatted questions of knee function – specifically designed for Study IV

Hur har din knäskada påverkat ditt liv?

Sätt ett **X** på linjen nedan för att visa hur du upplever att din knäskada påverkat dig.

Inte alls -- Extremt

Hur stora knäproblem (svullnad, att röra på knät mm) tycker du att du har?

Sätt ett kryss **X** på linjen nedan för att visa hur stora knäproblem du har.

Inga alls -- Mycket stora

Hur upplever du att ditt knä fungerar i vardagsaktiviteter (trappgång, resa dig mm)?

Sätt ett kryss **X** på linjen nedan för att visa hur du upplever din knäfunktion.

Mycket bra -- Mycket dåligt

Hur upplever du att ditt knä fungerar när du leker eller idrottar?

Sätt ett kryss **X** på linjen nedan för att visa hur du upplever din knäfunktion.

Mycket bra -- Mycket dåligt

Hur ont har du i ditt knä?

Sätt ett kryss **X** på linjen nedan för att visa hur ont du har.

Inte ont -- Extremt ont

G

Global Perceived Effect (GPE) scores – specifically designed for Study IV

Generic score:

Hur mår du i ditt knä nu, jämfört med förra gången, för 3 månader sedan?

Mycket bättre	Bättre	Lite bättre	Samma	Lite sämre	Sämre	Mycket sämre
☐	☐	☐	☐	☐	☐	☐

Subscale-specific scores:

Hur ont har du nu, jämfört med förra gången, för 3 månader sedan?

Mycket mindre	Mindre	Lite mindre	Samma	Lite mer	Mer	Mycket mer
☐	☐	☐	☐	☐	☐	☐

Hur upplever du dina knäproblem (svullnad, stelhet, nedsatt rörlighet) nu, jämfört med förra gången, för 3 månader sedan?

Mycket mindre	Mindre	Lite mindre	Samma	Lite mer	Mer	Mycket mer
☐	☐	☐	☐	☐	☐	☐

Hur klarar du att utföra vardagsaktiviteter (sitta, stå, gå, trappor, övriga förflyttningar, klä på dig, hjälpa till hemma) nu, jämfört med förra gången, för 3 månader sedan?

Mycket bättre	Bättre	Lite bättre	Samma	Lite sämre	Sämre	Mycket sämre
☐	☐	☐	☐	☐	☐	☐

Hur klarar du att utföra lek- och idrottsaktiviteter (springa, hoppa, sitta på huk, sitta på knä, vrida/snurra på skadat knät) nu, jämfört med förra gången, för 3 månader sedan?

Mycket bättre	Bättre	Lite bättre	Samma	Lite sämre	Sämre	Mycket sämre	Utför ej
☐	☐	☐	☐	☐	☐	☐	☐

Hur mår du över huvudtaget om du tänker på sådant som har med knät att göra (lita på knät, livsstil, hur ofta du tänker på knät, etc.) nu, jämfört med förra gången, för 3 månader sedan?

Mycket bättre	Bättre	Lite bättre	Samma	Lite sämre	Sämre	Mycket sämre
☐	☐	☐	☐	☐	☐	☐