Validation of the Short Form-36 Health Survey Supported With Isokinetic Strength Testing After Sport Knee Injury

Dusa Marn-Vukadinovic and Helena Jamnik

Context: Valid patient-based outcome instruments are necessary for comprehensive patient care that focuses on all aspects of health, from impairments to participation restrictions. Objective: To validate the Slovenian translation of Medical Outcome Survey (MOS) Short Form Health Survey (SF-36) and to assess relations among various knee measurements, activity tested with Oxford Knee Score (OKS) and health-related quality of life as estimated with SF-36 domains. Design: Descriptive validation study. Setting: Isokinetic laboratory in outpatient rehabilitation unit. Participants: 101 subjects after unilateral sport knee injury. Interventions: All subjects completed the SF-36 and OKS, and isokinetic knee-muscle strength output at 60°/s was determined in 78 participants. Within a 3-d period, 43 subjects completed the SF-36 and OKS questionnaires again. Main Outcome Measures: Reliability testing included internal consistency and test–retest reliability. Correlations between SF-36 subscales and OKS were calculated to assess construct validity, and correlation between SF-36 subscales and muscle strength was calculated to assess concurrent validity. Results: Chronbach α was above .78 for all SF-36 subscales. ICCs ranged from .80 to .93. The correlation between OKS and the physical-functioning subscale, showing convergent construct validity, was higher (r = .83, P < .01) than between OKS and mental health (r = .50, P < .01), showing divergent construct validity. Knee-extensor weakness negatively correlated with physical-functioning (r = −.59, P < .01) and social-functioning (r = −.43, P < .01) subscales. Conclusions: The Slovenian translation of the SF-36 is a reliable and valuable tool. The relationships between knee-muscle strength and activity and between knee-muscle strength and SF-36 subscales in patients after sport knee injury were established.

Keywords: knee joint, accidents, rehabilitation, evaluation, dynamometry

Throughout the history of medicine clinicians have developed numerous methods of assessing patients’ body structure and function. Medical treatment nowadays requires insight into the health consequences as perceived by the patients in addition to the assessment of impairments by clinician-based measures. However, during the
rehabilitation process, assessment of activities such as ambulation and participation in social life is also important. Frameworks such as the World Health Organization’s (WHO) International Classification of Functioning, Disability, and Health (ICF)\(^1\) are helpful in understanding the impact of health conditions (injury, illness, disease) on body structures and body functions, activity, and participation that are influenced by environmental and personal factors. In addition, these frameworks illustrate how health conditions influence areas such as health-related quality of life (HRQoL) that are often overlooked in traditional patient evaluation. For example, an anterior cruciate ligament tear may be associated with a disability as a result of inability to run, which can affect perceived physical, mental, and social well-being. This disability can manifest itself by negatively affecting a promising soccer career or hampering an avid runner’s social activity. Patient-based specific instruments are necessary to evaluate particular components of activity limitations and participation restrictions, based on disease, activity, or goal of the measurement.\(^2\)

In clinical practice we usually rely on clinician-based outcomes, which refer to body structure and function. Muscle weakness represents impaired knee function. Isokinetic dynamometry is an accurate and objective tool to assess individual muscle strength in comparison with the identical healthy muscle,\(^3\) and it is especially useful in assessing slight changes in muscle strength. The method has been recognized as a reliable and valuable test of knee-muscle weakness in healthy people and patients after knee injury.\(^4\)–\(^6\)

Functional tests or activity grading enables us to assess the influence of muscle weakness, range of motion, stability, muscle atrophy, or pain on personal activity. One can express activity restrictions in everyday life with a self-report instrument on a scoring scale and on an activity-grading scale.\(^7\) Many specific self-administered questionnaires showing the impact of an impaired or injured knee on subjective walking ability have been described,\(^8,\)\(^9\) and the Oxford Knee Score (OKS) is one of the simplest and shortest. It was evaluated in patients with knee-related injuries\(^10\) and includes questions related to knee symptoms and activity limitations during daily living and ambulation.

In addition, generic patient self-report instruments measure activity limitations and participation restrictions across different diagnostic groups and disabilities. They do not target any specific population. Authors have used these tools such as Medical Outcome Survey (MOS) Short Form Health Survey (SF-36) after knee injury or knee surgery to measure performance and health.\(^11\)–\(^14\) The SF-36 measures HRQoL in a healthy population and in patients to compare the impact of diseases on overall well-being.\(^15\) Since the SF-36 was made available in 1990,\(^16\) it has been described as a reliable and valid tool for numerous populations of patients in different countries\(^17\) and as a widely accepted measure in validation studies and in patient care. It has been used in studies of patients after knee surgery\(^14,\)\(^18\) and after knee injuries.\(^19\) Versions 1.0 and 2.0 of the SF-36 and shorter survey forms, translated into many languages, constitute the International Quality of Life Assessment project.\(^16\) The official Slovenian translation of the SF-36, version 1.0, has been available since 2002.

The primary aim of this study was to determine reliability, validity, and other practical considerations of the Slovenian translation of the SF-36. Secondarily we aimed to compare clinician-based knee measurements (range of motion, muscle atrophy, and muscle strength) with patient-based perceived activity or HRQoL
as estimated by the SF-36 domains in patients after sport knee injury during the rehabilitation process.

**Methods**

This study used a descriptive validation design. The main outcome measures were the physical-health and mental-health subscale scores of the SF-36, the OKS total score, knee-flexion range of motion, knee-extension deficit in range of motion, normalized thigh-muscle atrophy, and normalized peak-torque deficit values at 60°/s of the knee flexors and extensors. Participants were successively assigned to the Biodex, non-Biodex, and test–retest groups.

**Participants**

Patients who required prolonged physical therapy after conservative or surgical treatment of an injured knee came to the outpatient rehabilitation unit. We included all participants who fulfilled the following inclusion criteria in a 2-year period: Only subjects after unilateral ligament and/or menisci and/or chondral knee injury including patellar posttraumatic disorders and without other pathology were selected. All patients had been injured in recreational sports and had a preinjury Tegner Activity Scale score\(^2\) of ≥4. However, we excluded subjects with bone fracture, tumor, or joint inflammation or with severely restricted knee range of motion at the end of the outpatient program. In total, 101 patients were included. They gave informed consent to participate in the study. Our institutional research board approved the study protocol.

**Procedure**

**Interventions.** Before the assessment began, the rehabilitation medicine specialist who was the leader of the standard team of experts for isokinetic dynamometry did the examination, prescription of rehabilitation program, and follow-up. She recorded clinical characteristics such as inflammation, pain, tenderness, range-of-motion restriction, instability, muscle atrophy, muscle weakness, and ambulation capacity. The difference of both thigh circumferences at 10 cm proximal to the superior edge of the patella determined muscle atrophy. The outpatient rehabilitation program consisted of kinesiotherapy, stimulation of proprioception, gait restoration, and pathology-specific symptomatic physical therapy. Duration and intensity of a particular treatment modality depended on each subject’s abilities. The program checked range of knee motion, muscle atrophy, and isokinetic knee-muscle strength. Most programs ended more than 12 weeks after injury or operation. At the same time, all participants successively completed the OKS and SF-36 questionnaires. By the conclusion of the program patients were usually fit for full weight bearing and muscle strengthening and safe measurement of knee-muscle strength.

**Patient Assignment.** Patients were divided into 2 groups. Seventy-eight subjects who were able and agreed to perform the isokinetic dynamometric measurements were assigned to the Biodex group. Twenty-three subjects who were not fit to undergo isokinetic dynamometric measurement represented the non-Biodex group.
Within a 3-day interval all subjects from the non-Biodex group and 20 subjects from the Biodex group completed the OKS and SF-36 questionnaires twice as shown in Figure 1. There was no clinical change in the patients’ condition within these 3 days. Hence, in a 1-year period 43 participants answered the questionnaires twice to provide data for test–retest reliability testing. The results of 78 subjects from the Biodex group were used in validity testing. We checked all 101 baseline and 43 repeated questionnaires for floor and ceiling effects.

Isokinetic Dynamometry. The Biodex dynamometer provided muscle-strength data through isokinetic knee movements, and we followed the knee protocol and recommendations in the manufacturer’s manual.21 Subjects performed 4 series of 5 consecutive reciprocal concentric contractions of the knee extensors and flexors. They began on the healthy limb and for each velocity with extension first. At first the uninvolved limb performed a set of 5 movements at a velocity of 60°/s, followed by—for the purpose of better collaboration control—a set at 120°/s. At the beginning of each set, the patient did a warm-up with 3 to 5 submaximal and 1 maximal contraction in both directions to become familiar with the measurement or the new velocity. Patients rested for 1 to 2 minutes between the series of repetitions. The same procedure continued on the impaired side, with equal verbal encouragement.

The dynamometer graphically and numerically registered knee-muscle strength as extensor or flexor torques in tenths of an Nm through the whole motion. The software identified and listed peak-torque values in each direction and at each velocity for each knee. It also calculated a peak-torque deficit (PTD) for an identical series of repetitions for knee extensors and flexors according to the following formula: 

$$PTD = 100 - \left( \frac{\text{peak torque of involved knee}}{\text{peak torque of uninvolved knee}} \right) \times 100.$$  

We expressed muscle weakness as PTD at 60°/s for extensors and flexors in further statistical analysis.

OKS. The OKS is a self-report patient-based knee-specific questionnaire. It does not consider any clinical findings. There are 12 items on a 5-point rating scale, with scores for each answer ranging from 0 to 4. The sum of item scores gives the total score of the questionnaire. The highest possible score is 48, representing a normal knee, and 0 represents the worst possible status. As a knee-specific scale it is sensitive and responsive to functional change.22 Unfortunately, it is not sufficiently knee-specific, because it is heavily influenced by more proximal pathology.22

SF-36. The SF-36 is a self-administered patient-based questionnaire containing 35 items plus 1 extra item about personal perception of general health at the moment.

Figure 1 — Patient assignment.
Thirty-five items cover 8 of the most frequently measured health concepts, presented in 8 subscales: physical functioning (PF), role limitations resulting from physical health problems (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role limitations resulting from emotional problems (RE), and mental health (MH). The first 4 subscales refer more to physical health, and the other 4 subscales, to mental health. In addition, PF, RP, and BP contribute mostly to physical health phenomena, and MH and RE contribute mostly to mental health. VT, SF, and GH contribute to both physical health and mental health phenomena. Each subscale consists of at least 2 items (BP, SF) and maximum 10 items (PF). Items in 2 subscales (RP, RE) have only 2 possible answers. The scoring system is described in the SF-36 user’s manual. The lowest score for each subscale and component is 0, and the highest is 100. A higher score indicates more favorable health status. Scores for the 8 dimensions are usually presented as a profile.

**Statistical Analysis**

We determined the floor effect (percentage of the lowest possible score) and ceiling effect (percentage of the highest possible score) for the scores of the 8 subscales and 2 components of the SF-36 and for the OKS overall score in 144 questionnaires completed by 101 subjects.

Reliability testing included internal consistency and test–retest reliability. Chronbach $\alpha$ calculation showed internal consistency assessed in 101 baseline questionnaires. Test–retest reliability was assessed in the sample of 43 patients by ICC$_{1,1}$—single-score intraclass correlation coefficient for a 1-way model. These participants repeatedly answered the OKS and SF-36 questionnaires at 3-day intervals.

Construct-validity testing included verification of the theoretical relationship between the SF-36 subscale scores and external criteria (OKS) with the Spearman correlation coefficient. Comparison between the OKS and the analogous subscale score of the SF-36 (PF) showed convergent construct validity, and correlations between the OKS and MH score of the SF-36 showed divergent construct validity.

The Spearman correlation coefficient demonstrated concurrent validity by comparing the scores of the SF-36 domains with knee PTD in the group of 78 patients who had completed isokinetic dynamometry. Analysis showed correlations between knee range-of-motion values and the scores of SF-36 domains and between thigh-muscle atrophy and the SF-36 in the same group of patients. Muscle atrophy (MA) was expressed as a relative circumference difference: $MA = [(\text{circumference of uninvolved thigh} - \text{circumference of involved thigh})/\text{circumference of uninvolved thigh}] \times 100$. Baseline questionnaire scores served also to test validity of the SF-36. STATISTICA 6 software (StatSoft, Inc, Tulsa, OK) analyzed data with statistical significance level set at $P < .01$.

**Results**

**Patient Characteristics**

Table 1 summarizes patient characteristics with regard to the subsample purpose. Their mean age was 40 years (SD = 14, range = 16–72). According to the Tegner Activity Scale, their activity level before injury ranged from 4 to 7. Most of the
101 participants were rated at 5 or 6 (n = 70). Participants’ injuries had occurred in recreational sports, mostly downhill skiing. A pure anterior cruciate ligament injury was the most common health condition (n = 26). Some patients (n = 23) sustained a combination of anterior cruciate ligament and other knee-ligament ruptures or menisci tear. Other diagnoses were pure menisci tear (n = 14) or in combination with other knee lesions (n = 11), patellar disorders (n = 15), and others (n = 23) such as medial collateral ligament lesion, femoral chondral defect, muscle or tendon rupture, popliteal cyst formation, cicatrices, or undiagnosed knee disorders (distortion or contusion). Most patients had undergone surgery (n = 64) such as arthroscopy for anterior cruciate ligament reconstruction (n = 25) or partial meniscectomy (n = 22).

**Collected Data**

Structural and functional clinical characteristics were recorded for the group of 78 patients who completed dynamometric measurement. Knee-extensor PTD varied from 0.4% to 88.6%, with a median of 30.9%. The variance of knee-extensor peak-
torque values in a set of 5 repetitions was more than 20% in 5 participants (6%). Individual score variance of knee-flexor peak-torque values in a set was more than 20% in 19 participants (24%). In some of the cases with high variance the involved knee-flexor peak torque was higher than that of the uninvolved one. The median knee-flexor PTD was 14.1, and it ranged from –82.6% (which represents stronger flexors of the involved knee) to 69.5% (when the flexors of the uninvolved knee were stronger than the flexors of the involved knee). Most patients could fully extend their knee joints and flex to 140°. There was 1 subject with a knee-extension deficit of 15° and 13 subjects with 5° or 10° deficits. Two participants could perform knee flexion of only 80°, 1 of 90°, and the others from 110° to 155° of knee flexion. Median normalized thigh-circumference deficit was 2 cm (range 0–9.3 cm).

Participants answered both questionnaires on their own, except for 5 who did it by phone, which took 10 to 15 minutes. Table 2 presents the characteristics of the score distribution on the SF-36 subscales and components.

Floor and ceiling effects were evaluated in 101 baseline SF-36 and OKS questionnaires and the 43 retest questionnaires. Subjects achieved neither the maximum nor the minimum possible score on PC of the SF-36. One questionnaire showed the maximum possible score on MC, but that was not the case for subscales RP and RE, which had the highest or the lowest scores in more than 30% of questionnaires in our study (Table 2). Two patients achieved the highest possible overall score of 48 on the OKS, and no subject had the lowest possible score on it.

**Reliability**

The mean values of the Chronbach $\alpha$ coefficient were .93 for the PC score and .92 for MC. Alphas ranged from .78 for the GH subscale score of the SF-36 to .93 for the PF subscale (Table 3).

According to the ICC, the test–retest reliability of the SF-36 domains ranged from .80 for RP to .93 for VT. The ICCs were equal to or above .90 for the PC and MC of the SF-36 and in the subscales PF, GH, and VT. ICCs for the other subscales were above .80 (Table 3).

**Construct Validity**

Correlation between the OKS and PC values of the SF-36 (Table 4) was significant and high ($r = .80, P < .01$), with high correlations between the OKS and PF ($r = .83, P < .01$) and between the OKS and BP ($r = .78, P < .01$). On the other hand, weaker correlations were found between the OKS and subscales representing mental health in the SF-36, namely, RE ($r = .48, P < .01$) and MH ($r = .50, P < .01$).

**Concurrent Validity**

The correlations between GH or MH and any kind of knee function (range of motion, PTD, or MA) were too low to be significant. Knee range of motion or MA did not correlate with the PC or MC scores of the SF-36. There were weak yet significant negative correlations between knee-muscle weakness and the PC score. The PF, BP, and SF scores significantly negatively correlated with PTD values at 60°/s of knee flexion or extension. The VT score showed a significant negative correlation
Table 2  Score Distribution for the SF-36 Domains

<table>
<thead>
<tr>
<th>Domain</th>
<th>Baseline score, n = 101</th>
<th>Test–Retest Reliability, n = 43</th>
<th>Dynamometry, n = 78</th>
<th>Floor (ceiling) effect, n = 144</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1st assessment</td>
<td>2nd assessment</td>
<td></td>
</tr>
<tr>
<td>Physical functioning</td>
<td>52.3 (27.5)</td>
<td>47.9 (28.1)</td>
<td>53.9 (28.3)</td>
<td>1.4 (1.4)</td>
</tr>
<tr>
<td>Role physical</td>
<td>31.7 (38.6)</td>
<td>26.2 (34.9)</td>
<td>33.0 (39.4)</td>
<td>52.8 (15.3)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>49.0 (26.0)</td>
<td>45.5 (26.3)</td>
<td>50.7 (26.6)</td>
<td>0.7 (4.2)</td>
</tr>
<tr>
<td>General health</td>
<td>67.2 (20.9)</td>
<td>67.8 (22.1)</td>
<td>67.5 (21.3)</td>
<td>0 (6.9)</td>
</tr>
<tr>
<td>Vitality</td>
<td>59.9 (20.4)</td>
<td>57.0 (22.7)</td>
<td>60.6 (19.9)</td>
<td>0 (0.7)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>62.5 (29.8)</td>
<td>59.6 (31.4)</td>
<td>63.3 (30.2)</td>
<td>4.2 (20.1)</td>
</tr>
<tr>
<td>Role emotional</td>
<td>68.3 (39.8)</td>
<td>62.0 (40.9)</td>
<td>70.5 (39.8)</td>
<td>21.5 (51.4)</td>
</tr>
<tr>
<td>Mental health</td>
<td>69.0 (18.7)</td>
<td>71.7 (19.3)</td>
<td>68.6 (18.3)</td>
<td>0 (1.4)</td>
</tr>
<tr>
<td>Physical component</td>
<td>55.7 (21.5)</td>
<td>53.2 (22.1)</td>
<td>56.4 (19.9)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Mental component</td>
<td>64.8 (19.3)</td>
<td>63.2 (24.9)</td>
<td>65.1 (19.2)</td>
<td>0 (0.7)</td>
</tr>
</tbody>
</table>

Reported as mean (SD) and percentage scoring at the floor effect (the lowest possible score) and at the ceiling effect (the highest possible score).

a The values represent the percentage of questionnaires with the lowest possible score (floor effect) or the highest possible score (ceiling effect) of different parts of the SF-36.
Validation of the SF-36 and Strength Testing

with extensor PTD. Table 4 shows all these results. MA was moderately correlated with knee-extensor PTD ($r = .63, P < .01$).

**Discussion**

In this study we established relationships between objective clinician-based measurements and subjective patients’ perception of their own disability and health after sport knee injury. Application of the WHO-ICF model in clinical reasoning helps us understand components that determine the impact of structural and functional clinical findings on perceived activity and participation limitation, as well as on perceived health. Participants in this study estimated their HRQoL through the SF-36 scores. The Slovenian translation of this tool has not been validated until now, and we are unaware of any validation study of the SF-36 after knee injury sustained in recreational sport.

We decided to use the unofficial Slovenian translation of the OKS (not yet validated) and compare it with the official Slovenian translation of the SF-36 in the validation process. This is because of the lack of a well-accepted and standardized system of assessing activity after sport knee injury. The use of a generic scale and disability-specific scale gives us more complex information about personal functioning and has been recommended for assessing patient outcomes.

The Tegner Activity Scale is suitable for recording preinjury activity level. It rates people receiving a disability pension with 0 and competitive sportsmen with 10 points. Participants in our study were active in sports (a Tegner Activity Scale score of $\geq 4$) before sport soft-tissue knee injury. The scale was a measure of sample homogenization and took into account the patients’ preinjury and desired postrehabilitation activity levels.

We normalized thigh circumference and muscle torque to avoid interpersonal anthropometric differences. Values of the uninvolved limb served as control values. Normalization of thigh-circumference values is not common. PTD, which considers both legs of the same person, was calculated for knee extensors and flexors.

### Table 3 Results of Reliability Testing of the SF-36 Subscales and Components

<table>
<thead>
<tr>
<th>Domain</th>
<th>Chronbach $\alpha$ (n = 101)</th>
<th>ICC$_{1,1}$a (n = 43)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>.93 (.92–.93)</td>
<td>.90</td>
</tr>
<tr>
<td>Role physical</td>
<td>.85 (.78–.84)</td>
<td>.80</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>.89 (.80–.95)</td>
<td>.81</td>
</tr>
<tr>
<td>General health</td>
<td>.78 (.69–.78)</td>
<td>.92</td>
</tr>
<tr>
<td>Vitality</td>
<td>.82 (.76–.78)</td>
<td>.93</td>
</tr>
<tr>
<td>Social functioning</td>
<td>.91 (.87–.88)</td>
<td>.87</td>
</tr>
<tr>
<td>Role emotional</td>
<td>.82 (.61–.83)</td>
<td>.88</td>
</tr>
<tr>
<td>Mental health</td>
<td>.86 (.81–.86)</td>
<td>.93</td>
</tr>
<tr>
<td>Physical component</td>
<td>.93 (.92–.93)</td>
<td>.95</td>
</tr>
<tr>
<td>Mental component</td>
<td>.92 (.91–.93)</td>
<td>.95</td>
</tr>
</tbody>
</table>

*a* Single-score intraclass correlation coefficient for a 1-way model.
Table 4  Correlations Between Domains of the SF-36 and Oxford Knee Score (OKS) or Knee Functions to Show Validity Estimates for the SF-36

<table>
<thead>
<tr>
<th>Domain of the SF-36</th>
<th>OKS (n = 101)</th>
<th>Limited knee extension (n = 78)</th>
<th>Limited knee flexion (n = 78)</th>
<th>PTD</th>
<th>Knee extensors (n = 78)</th>
<th>Knee flexors (n = 78)</th>
<th>MA (n = 78)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>.83</td>
<td>−.12</td>
<td>.22</td>
<td></td>
<td>−.59</td>
<td>−.38</td>
<td>−.26</td>
</tr>
<tr>
<td>Role physical</td>
<td>.64</td>
<td>.07</td>
<td>.15</td>
<td></td>
<td>−.21</td>
<td>−.26</td>
<td>.05</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>.78</td>
<td>−.11</td>
<td>.18</td>
<td></td>
<td>−.40</td>
<td>−.29</td>
<td>−.12</td>
</tr>
<tr>
<td>General health</td>
<td>.49</td>
<td>.02</td>
<td>.15</td>
<td></td>
<td>−.22</td>
<td>−.13</td>
<td>−.22</td>
</tr>
<tr>
<td>Vitality</td>
<td>.62</td>
<td>−.09</td>
<td>−.02</td>
<td></td>
<td>−.32</td>
<td>−.21</td>
<td>−.18</td>
</tr>
<tr>
<td>Social functioning</td>
<td>.70</td>
<td>−.04</td>
<td>.07</td>
<td></td>
<td>−.43</td>
<td>−.28</td>
<td>−.17</td>
</tr>
<tr>
<td>Role emotional</td>
<td>.48</td>
<td>−.12</td>
<td>.05</td>
<td></td>
<td>−.22</td>
<td>−.27</td>
<td>−.18</td>
</tr>
<tr>
<td>Mental health</td>
<td>.50</td>
<td>−.05</td>
<td>.00</td>
<td></td>
<td>−.22</td>
<td>−.11</td>
<td>−.17</td>
</tr>
<tr>
<td>Physical component</td>
<td>.80</td>
<td>−.04</td>
<td>.22</td>
<td></td>
<td>−.48</td>
<td>−.30</td>
<td>−.24</td>
</tr>
<tr>
<td>Mental component</td>
<td>.66</td>
<td>.03</td>
<td>.05</td>
<td></td>
<td>−.36</td>
<td>−.20</td>
<td>−.22</td>
</tr>
</tbody>
</table>

PTD, peak torque deficit; MA, muscle atrophy.
Normalization by comparing the involved with the uninvolved side of the same person is one of the most recommended ways to express individual muscle strength\(^4\) and even performance.\(^26\) Despite the reliable protocol of isokinetic dynamometry, we have to consider high variability of peak torque in a series of movements toward flexion in our study. In about one-quarter of the participants, variability of flexor peak-torque values was extremely high, making it hard to reach valid conclusions about their relation to perceived activity. A variance of peak torque in a series of repetitions that is not higher than 20% has been recommended to make valid conclusions.\(^4\) Patients with weak knee muscles after sport injury often fail to perform maximal effort during knee flexion immediately after maximal effort of the extensors.

The SF-36 is a reliable and valuable tool for use in patients with injured knees. The instrument is ready to be applied in clinical and research trials in Slovenia for international comparison of results in this patient population, as well as for comparison with different treatments or impairments. Recently, we have observed growing interest in instruments of activity and participation assessment in our country, but psychometric validation studies are rare.\(^27,28\) Researchers have proposed a checklist of instrument properties in reporting clinometric characteristics of a questionnaire that intends to facilitate future instrument selection.\(^29\) According to this checklist, we should explore several instrument features of the SF-36: floor and ceiling effects, internal consistency, test–retest reliability, construct convergent and construct divergent validity, time to administer the scale, interpretability, appropriateness of the scoring method, and concurrent validity.

The lack of floor and ceiling effects is a major consideration regarding the responsiveness of the SF-36. Our participants never achieved the best or the worst possible score on the PC subscale of the SF-36, and we found a low percentage of questionnaires with the lowest or the highest possible score on MC and on 6 out of 8 subscales, namely PF, BP, GH, VT, SF, and MH. This is well within recommendations of <30% of the sample with the highest or the lowest score.\(^20\) Considering the subscales RP and RE, the lowest or highest scores occurred in more than 50% because these are subscales with only 2 possible answers on all items. That has been improved in SF-36 version 2.0.\(^15\) Marx et al, who applied the SF-36, version 1, in their study, noticed the occurrence of the highest score values in all subscales.\(^30\)

Our results demonstrate very strong internal consistency of the SF-36 subscales and components in patients after sports knee injury using the Chronbach \(\alpha\) coefficient. Hence, all items of the SF-36 strongly contributed to the measurement. A Chronbach \(\alpha\) coefficient of >.70 was considered acceptable\(^20\) for patients enrolled in a clinical trial. The GH subscale had the poorest internal consistency in this study, similar to previous research.\(^31\) Adequate group comparison is possible when Chronbach \(\alpha\) coefficient is equal to or exceeds .50–.70.\(^31\) Patients with sport knee injury achieved Chronbach \(\alpha\) coefficient value for PC, MC, PF, and SF of .90 and higher in our study. They met the recommended standard for assessment of individual patients, because a reliability of .90 is regarded as adequate for individual comparison, for which higher measurement standards are needed than those used in group comparison.\(^31\)

Excellent repeatability across the second administration of the SF-36 confirms strong reliability of all SF-36 components and subscales, with ICCs ranging from
.80 for the RP to .95 for each component of the SF-36. There was no need to ask all subjects to complete questionnaires twice because we felt that the sample of 43 patients who did was suitable for determining test–retest reliability. Values of ≥.75 are considered to indicate good test–retest reliability. These findings suggest that the SF-36 can be used reliably in active patients after sport knee injury.

We aimed to determine construct validity of the SF-36 by comparing its components and subscales with the OKS. The 10-item PF subscale is very useful in rehabilitation because it reflects activity level according to the ICF. Strong correlation ($r \geq .80$) between knee-specific activity measurement (OKS) and the PF or PC indicated that they measure the same activity dimension, showing convergent construct validity. PC, PF, and BP were the strongest qualifiers of activity after sport knee injury. The PF subscale of the SF-36 has already been recommended as a core qualifier to assess physical performance of neurological or orthopedic patients in research and clinical practice. The correlation between the OKS and PF was significantly stronger than correlation between the OKS and MH or between the OKS and RE (both about $r = .50$). MH and RE represent mental health, and their low correlations with the OKS show divergent construct validity. Moderate correlation ($r = .70$) between the OKS and SF indicates an important relationship between athletes’ activity and their participation in social life.

Besides psychometric qualities, the SF-36 addresses some practical considerations that are also important during instrument evaluation and selection. The SF-36 is a simple and short questionnaire for obtaining necessary data in the rehabilitation process. The time needed to complete the 36 items was about 10 minutes in our study, which meets practical standards for survey completion. Self-administration also saves experts time when collecting useful data and ensures a focus on patient preferences rather than on the disease. On the other hand, the need for comprehensive scoring instructions and special software for the SF-36 could prevent its routine use in some less well-funded institutions.

At the end of the rehabilitation process participants rated their HRQoL on average (PC = 51.6 ± 23.7, MC = 65.2 ± 21.7) similarly to other athletes with knee disorder (PC = 46.1 ± 10.2, MC = 53.1 ± 8.7). The mean subscale scores for the group with sport knee injury are well below the means for the U.S. general population, ranging from 84.2 ± 23.3 (PF) to 60.9 ± 21.0 (VT). One of the aims of our study was to identify functional deficits that underlie activity limitations and restrictions in participation or HRQoL. Correlations between MA and the SF-36 component scores were too low to be statistically significant with the given sample size, although there was moderate correlation between MA and extensor PTD. We did not find any significant correlations between knee range of motion and the SF-36 subscales, either, suggesting that perception of HRQoL is not associated with most clinician-based measures. It seems that there is some misunderstanding between experts and patients about good treatment results. Focusing on clinician-based measures as an indicator of HRQoL may be incorrect because they do not reflect patient expectations. Patient-based tools have remained the main indicator of HRQoL. It has been shown that clinical findings often fail to correlate with activity scores, and it has been suggested that separate scores for symptoms, subjective functions, and objective results be presented.

Weak knee muscles are associated with reduced activity in patients after knee sport injury. We observed the highest negative correlation between knee-muscle
isokinetic strength and PF score ($r = -0.59$ for extensors and $r = -0.38$ for flexors, $P < .01$), showing a close relationship between objective knee-muscle function and perceived walking (and related daily activity) ability determined with the PF subscale score of SF-36. We can say that the PF or OKS score in this patient group provides some information about knee-muscle weakness, and it is less labor intensive for a physician when patients complete the questionnaire before examination. Great variability of flexor peak torques in one-fourth of subjects could influence a weaker relationship between flexor strength and perceived disability than in the case of the relationship between knee-extensor strength and PF. Marx et al noticed that activity level was the most important prognostic factor after sport injury and suggested finding a relationship with decreased function.

Our study identified interesting relationships between knee-muscle strength and some aspects of athletes’ HRQoL, showing that patients with weak knee extensors have lower perceived physical and mental HRQoL. In fact, knee-muscle weakness was related to low HRQoL expectations regarding physical functioning, bodily pain, vitality, and social functioning in people after sport knee injury. We determined moderate relations between knee-muscle strength (flexors and extensors) and the PC score, while associations of knee PTD to the BP, VT, or SF scores were negative and weak yet significant. Subjects who are active in recreation might focus more on physical performance and participation in social life. Lack of correlation between knee-muscle strength and the RP score seemed to show that work restrictions tended to have less impact on athletes’ HRQoL than social participation. Similarly, injured adolescent athletes demonstrated that injury could affect areas of health beyond the physical domains—for example, social functioning. We can assume that correlation between BP score and perceived activity score indicates that some patients’ activities are restricted if they feel pain. It seems that the impact of pain on athletes’ perceived activity is greater than that of knee-muscle weakness. In our study, MH and GH were not related to depressed body function such as range of motion, strength, or MA. Physicians have the impression that impairments such as pain or clinical objective findings like joint range of motion, muscle atrophy, and weakness are main determinants of personal activity, participation, and health. Similarly, injured adolescent athletes demonstrated that injury could affect areas of health beyond the physical domains—for example, social functioning. A rather weak correlation has been found between clinician and patient rating of severity of knee condition in athletic patients with knee disorder.35

Our study has limitations such as a lack of in-depth responsiveness testing of the SF-36 in patients after soft-tissue sport knee injury for the purpose of measuring outcomes, follow-up, and rehabilitation planning. Another potential limitation is that we did not use an additional mental-health-specific instrument to determine whether it correlated with MH more than with PF. However, adequate construct validity of the SF-36 was confirmed through divergent construct validity. In addition, the use of an older “version 1” of the SF-36 questionnaire in this study reflects in poor responsiveness of some SF-36 subscales. Quality of life is a multidimensional and patient-specific construct. We did not search for the effect of personal and environmental factors on activity, participation, and HRQoL expectations. Subjective variables such as self-efficacy beliefs or patient-derived subjective assessment of symptoms and function have been found to be more important in patients’ satisfaction with the outcome than objective findings.
The results of our study are reliable and valuable for patients who are active in recreational sports and have sustained knee ligament, meniscus, or cartilage injury. We speculate that the SF-36 would be adequate for patients with other kinds of knee disabilities or for a more sedentary group of patients.

Conclusions

We evaluated the clinical use of the SF-36 for patients after sport knee injury. The Slovenian translation of the SF-36 remains a highly reliable and valuable tool to test activity, participation, and HRQoL, and it provides meaningful, patient-based information. The overall PC and MC and most subscale scores were shown to be sensitive to recreational athletes with knee injuries. The study demonstrated a strong relationship between perceived activity and bodily pain, and activity was moderately associated with athletes’ participation in social life. We found that PC, PF, and BP scores of the SF-36 in this patient group are good qualifiers of activity, and they provide some information about knee-muscle weakness. We established a poor relationship between knee-muscle strength and athlete social functioning. Physical and mental health were not related to depressed body function in terms of knee range of motion, knee-muscle strength, or muscle atrophy after sport knee injury.

The SF-36 allows an international comparison of results in patients after sport knee injury, as well as comparison of different treatments or impairments in clinical and research trials.

References


