

# Cross-cultural Validation of the Norwegian Version of the Banff Patellofemoral Instability Instrument 2.0

Trine Hysing-Dahl,<sup>\*†‡</sup> PT, MSc, Liv Heide Magnussen,<sup>§</sup> PT, Prof., Anne Gro Heyn Faleide,<sup>†‡</sup> PT, PhD, Asle Birkeland Kjellsen,<sup>||</sup> MD, Ingunn Fleten Mo,<sup>†</sup> PT, MSc, Per Arne Skarstein Waaler,<sup>||</sup> MD, Renate Mundal,<sup>||</sup> PT, BSc, and Eivind Inderhaug,<sup>‡||</sup> MD, MPH, PhD

*Investigation performed at the Department of Orthopaedic Surgery, Haralds plass Deaconess Hospital, Bergen, Norway*

**Background:** The Banff Patellofemoral Instability Instrument (BPfII) 2.0 is a disease-specific quality of life questionnaire for patients with patellofemoral instability. While good psychometric properties have been demonstrated, the data lack cross-cultural validity, construct validity, and an established measurement error.

**Purpose:** To (1) translate and cross-culturally adapt the BPfII 2.0 to the Norwegian version (BPfII 2.0–No) and (2) examine the psychometric properties of the Norwegian version.

**Study Design:** Cohort study (diagnosis); Level of evidence, 3.

**Methods:** The BPfII 2.0 was translated according to international guidelines. A cohort of 100 patients surgically treated for recurrent patellofemoral instability completed the BPfII 2.0–No, related outcome measures (Norwich Patellar Instability Score, International Knee Documentation Committee Subjective Knee Form 2000, Knee injury and Osteoarthritis Outcome Score, and Tampa Scale of Kinesiophobia), and functional tests (Y-Balance Test–Lower Quarter, single-leg hop tests, and knee extension strength) before and/or 6 months after surgery. We evaluated the face and content validity, internal consistency (Cronbach  $\alpha$ ), test-retest reliability (intraclass correlation coefficient [ICC]), measurement error (SEM and smallest detectable change at the individual [SDC<sub>ind</sub>] and group levels [SDC<sub>group</sub>]). Construct validity was assessed by testing 9 hypotheses on the correlation between the BPfII 2.0–No and the outcome measures/functional tests (Pearson  $r$ ).

**Results:** The BPfII 2.0–No had good face and content validity. Internal consistency was excellent ( $\alpha = .95$ ), and no floor or ceiling effects were found. Test-retest reliability was high (ICC<sub>2,1</sub> = 0.87; 95% CI, 0.77–0.93), and measurement error was low (SEM = 7.1). The SDC<sub>ind</sub> was 19.7 points and the SDC<sub>group</sub> was 2.8 points. Eight of the 9 hypotheses regarding construct validity were confirmed.

**Conclusion:** The BPfII 2.0–No was found to be valid and reliable. This study adds further knowledge on the measurement properties of the BPfII 2.0 that can be used internationally.

**Keywords:** Banff Patellofemoral Instability Instrument 2.0; COSMIN; patellar instability; quality of life; reliability; validity

Recurrent lateral dislocation of the patella is a disabling disorder that causes pain and reduces the quality of life.<sup>11</sup> The etiology is diverse but often includes deviant knee anatomy that predisposes the patella to lateral

dislocation. Patients experiencing recurrent dislocations are advised to consider surgical treatments that address anatomic risk factors for dislocation and medial retinaculum reinforcement procedures.<sup>18,43</sup> Disease-specific quality of life measurements are commonly used to monitor patients' progression after treatment and determine the success or failure of surgical interventions.<sup>15</sup> Until recently, there was a lack of validated outcome measures for patients with recurrent patellofemoral instability.<sup>11</sup>

The Orthopaedic Journal of Sports Medicine, 11(5), 23259671231168881  
DOI: 10.1177/23259671231168881  
© The Author(s) 2023

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

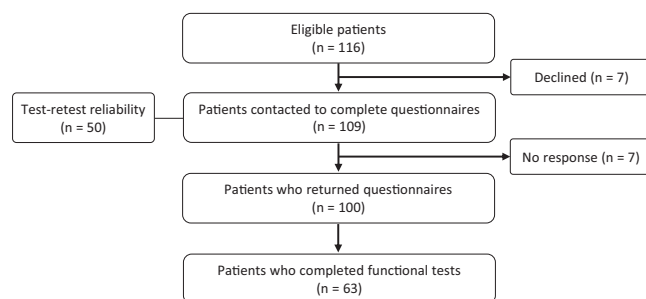
This has reduced clinicians' ability to assess clinical interventions and has limited their ability to understand the daily subjective limitations of these patients. The Banff Patellofemoral Instability Instrument (BPII) was developed to fill this gap in measuring quality of life in patients with patellofemoral instability.<sup>11</sup> The questionnaire comprises 5 domains, covering key aspects of quality of life, including symptoms/physical complaints, work- and school-related concerns, recreational activity, and sport participation/competition.<sup>15</sup> It was originally modified from a questionnaire for patients with anterior cruciate ligament (ACL) deficiency (the ACL-quality of life questionnaire).<sup>11</sup> The first version of the BPII included 32 items. After principal components analysis and item reduction, a shortened 23-item BPII 2.0 was introduced<sup>15</sup> in 2016. Both versions have demonstrated good measurement properties in surgically and nonsurgically treated patients,<sup>10-12,15,17</sup> adolescents,<sup>16,17</sup> and adults. Each of the 23 items of the BPII 2.0 is equally weighted and answered on a visual analog scale. The final score is calculated as a mean of the scores from all answered items (range, 0-100), where a higher score reflects a higher quality of life.<sup>15</sup>

Further psychometric testing is required to build greater scientific soundness for the BPII 2.0. In particular, cross-cultural validity and hypothesis testing have only been performed in 1 other cohort<sup>4</sup> than the original development study.<sup>15</sup> Furthermore, measurement error should be established with a recommended method.<sup>6</sup> Although a few translations have been made,<sup>4,27,39</sup> there is currently no translated and validated Norwegian version available.

The present study aimed to provide Norwegian clinicians with a tool to evaluate quality of life in patients with recurrent patellofemoral instability. Second, this study also aimed to further expand knowledge on the validity of the BPII 2.0 by examining content validity, internal consistency, test-retest reliability, measurement error, and construct validity. We hypothesized that a Norwegian version of the BPII 2.0 (BPII 2.0-No) would be valid and reliable in patients with recurrent patellofemoral instability.

## METHODS

Before enrollment and data collection, all patients gave their written, informed consent, and the study protocol received ethics committee approval. This study was performed in 2 stages. First, the BPII 2.0 was translated and cross-culturally adapted. Second, the BPII 2.0-No was



**Figure 1.** Flowchart of patient participation.

examined for measurement properties in a prospective cohort of patients before and/or 6 months after surgery for patellofemoral instability, including a 2-week test-retest interval at the 6-month follow-up.

## Translation and Cross-cultural Adaption

The BPII 2.0 was translated and cross-culturally adapted into Norwegian according to the guidelines described by Beaton et al.<sup>3</sup> This process involved an expert committee of 3 orthopaedic surgeons (E.I., A.B.K., P.A.S.W.), 4 physical therapists (T.H.-D., A.G.H.F., I.F.M., R.M.) specialized in orthopaedic physical therapy, a researcher (L.H.M.) with extensive experience in clinimetrics research methodology, a teacher specialized in the Norwegian language, and 2 back-translators, both native speakers of English. The expert committee had close contact with the author of the original version throughout the process. To ensure readability for adolescents, a 12-year-old child completed the prefinal version of the questionnaire and commented on difficult wording.

## Examination of Measurement Properties

Patients undergoing surgical treatment for recurrent patellar dislocation were prospectively recruited from 2 Norwegian orthopaedic centers from January 2021 to September 2022. Patients were eligible for participation if they were aged  $\geq 13$  years at the time of surgery, fluent in Norwegian, and able to understand and complete the questionnaires. Patients with concomitant knee ligament injuries were excluded. A total of 100 patients met the inclusion

\*Address correspondence to T. Hysing-Dahl, PT, MSc, Haraldsplass Deaconess Hospital, V/Avdeling for Rehabiliteringstjenester, Postboks 6165, Bergen, 5892, Norway (email: trine.hysing-dahl@haraldsplass.no) (Twitter: @HysingDahl).

<sup>†</sup>Haraldsplass Deaconess Hospital, Bergen, Norway.

<sup>‡</sup>University of Bergen, Bergen, Norway.

<sup>§</sup>Western Norway University of Applied Science, Bergen, Norway.

<sup>||</sup>Haukeland University Hospital, Bergen, Norway.

Final revision submitted January 12, 2023; accepted February 9, 2023.

One or more of the authors has declared the following potential conflict of interest or source of funding: Open access funding was provided by the University of Bergen. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from the Regional Committee for Medical and Health Research Ethics of Mid-Norway (reference No. 185067).

criteria and gave their consent for participation in the study. See Figure 1 for the enrollment flowchart.

In addition to the BPII 2.0–No, all patients completed a battery of patient-reported outcome measures (PROMs)—including the Norwich Patellar Instability Score (NPI), International Knee Documentation Committee (IKDC) Subjective Knee Form 2000, Knee injury and Osteoarthritis Outcome Score (KOOS), and Tampa Scale of Kinesiophobia (TSK). Patients also underwent functional testing—including the Y-Balance Test–Lower Quarter (YBT-LQ), single-leg hop tests, and knee extension strength—preoperatively and/or 6 months after surgery.

The NPI was developed to assess patient-perceived symptoms of patellofemoral instability during activity. It includes 19 questions that are rated on a 5-point Likert scale, with options from never to always.<sup>32</sup> The sum score ranges from 0 to 250 and is presented as a percentage, where a higher score indicates higher instability.<sup>34</sup> The NPI has demonstrated good measurement properties in several domains,<sup>32–34</sup> including adequate construct validity,<sup>10,32,34</sup> high internal consistency, and responsiveness.<sup>32</sup> The score has recently been translated into Norwegian (validation process completed; provided by the responsible investigator (T.H.-D.).

The IKDC is a knee-specific, patient-reported tool, comprising 18 questions across 3 domains—symptoms, physical activity, and function.<sup>13</sup> One sum score, ranging from 0 to 100, is made, with a higher score indicating a better function.<sup>13</sup> The IKDC has demonstrated good psychometric properties for patients with mixed knee pathologies and injuries<sup>1</sup> and is validated in patients with patellar instability.<sup>25</sup>

The KOOS was developed to assess patients' opinions about their knee function and associated problems. It comprises 5 subscales—Pain, Symptoms, Activities of Daily Living, function in Sport and Recreation, and knee-related Quality of Life (QOL). Scores from each subscale range from 0 (lowest function) to 100 (highest function).<sup>28</sup> The KOOS was developed for patients with knee injuries and/or osteoarthritis but is frequently used in patients with patellofemoral instability. The questionnaire has demonstrated satisfactory psychometric properties for a variety of knee conditions.<sup>28</sup>

The TSK measures fear of movement in patients with low back pain,<sup>20</sup> but it has also been used to measure fear of reinjuries in patients after medial patellofemoral ligament reconstruction.<sup>31</sup> The sum of the 13 items included provides a score from 0 to 52, with a higher score indicating more fear of movement.<sup>20</sup> The Norwegian version of the TSK is validated for patients with sciatica.<sup>9</sup> It is widely used to assess kinesiophobia after knee injuries such as ACL injuries.<sup>2,14,24,37</sup>

The YBT-LQ evaluates knee stability and 1-leg dynamic balance in 3 directions—anterior, posteromedial, and posterolateral.<sup>26</sup> Reach distance is normalized to leg length, which is measured from the anterior superior iliac spine to the most distal portion of the medial malleolus. Standing on 1 leg, patients reach as far as possible in each direction without losing their balance, and the mean reach distance of 3 attempts is recorded in centimeters. Results are presented as a composite score of all 3 directions (as

a percentage).<sup>26,30</sup> The YBT-LQ has proven to be a reliable test for impaired balance symmetry and potentially increased risk for injury.<sup>30</sup>

Single-leg hop tests evaluate the function, dynamic strength, and lower extremity muscle power.<sup>29,41</sup> They comprise 4 tasks—single-leg hop for distance (in cm), triple hop for distance (in cm), triple crossover hop for distance (in cm), and 6-m timed hop (in s). Results are presented as a limb symmetry index (in %), calculated as  $\text{surgical limb/uninvolved limb} \times 100$  for each test individually and as a sum score (all 4 tests combined).<sup>23</sup> Single-leg hop tests are reliable and valid performance tests for patients with knee injuries.<sup>19</sup>

Knee extension strength (peak torque in N·m) was measured with an isokinetic device (Biodex system 4 dynamometer; Biodex Medical Systems) using a standardized protocol of 5 repetitions at 60 deg/s. Isokinetic strength tests are considered the gold standard for measuring muscle strength<sup>7</sup> and are reliable and valid outcome measures after a knee injury.<sup>35,38</sup>

## Data Evaluation and Statistical Analysis

The study sample size was determined according to recommendations from Terwee et al,<sup>36</sup> suggesting a minimum of 50 patients for assessing construct validity, reliability, and floor or ceiling effects, and a minimum of 100 patients for assessing internal consistency.<sup>6</sup> SPSS 26.0 (IBM Corp) was used for data analyses, which included descriptive statistics, testing of normality, examination of internal consistency, test-retest reliability, Bland-Altman plots, hypothesis testing (significance level,  $P \leq .05$ ), and floor and ceiling effects. Continuous variables were reported as means and standard deviations, and categorical variables were reported as absolute values and relative frequencies. The measurement error was calculated in Microsoft Excel 2016.

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) were followed when examining the measurement properties of the BPII 2.0–No.<sup>21,36</sup> These guidelines provide definitions and criteria for evaluation of the quality of a questionnaire's validity and reliability. Face validity, the degree to which a questionnaire looks as though it is an adequate reflection of the construct,<sup>22</sup> and cultural adaptation of the Norwegian version were assessed by the expert committee. As no difficulties were encountered, no changes were made to the final version that was used in the validation process. Content validity, the degree to which the content of the BPII 2.0 is an adequate reflection of the construct to be measured,<sup>22</sup> was assessed in a subgroup of 10 patients who tested the prefinal version. "Think aloud" interviewing<sup>42</sup> was applied when completing the BPII 2.0–No followed by questions about patients' interpretation of each item, item relevance, any ambiguous wording, and overall importance for their quality of life.

As no gold standard exists for measuring quality of life in patients with patellofemoral instability, construct validity was assessed by forming and testing hypotheses<sup>6</sup>

TABLE 1  
The 9 Study Hypotheses Regarding Expected Associations Between the BPII 2.0–No and Measures of Knee Function<sup>a</sup>

Hypothesis
1. There would be a medium to large negative correlation ( $-0.30 < r < -1.0$ ) with the NPI
2. There would be a large negative correlation ( $-0.50 < r < -1.0$ ) with the TSK
3. There would be a large correlation ( $0.50 < r < 1.0$ ) with the IKDC
4. There would be a large correlation ( $0.50 < r < 1.0$ ) with the KOOS–Pain
5. There would be a large correlation ( $0.50 < r < 1.0$ ) with the KOOS–Symptoms
6. There would be a large correlation ( $0.50 < r < 1.0$ ) with the KOOS–ADL
7. There would be a large correlation ( $0.50 < r < 1.0$ ) with the KOOS–Sport/Rec
8. There would be a large correlation ( $0.50 < r < 1.0$ ) with the KOOS–QOL
9. There would be a small to medium correlation ( $0.10 < r < 0.50$ ) with the functional tests

<sup>a</sup>ADL, Activities of Daily Living; BPII 2.0–No, 23-item Banff Patellofemoral Instability Instrument–Norwegian version; IKDC, International Knee Documentation Committee Subjective Knee Form 2000; KOOS, Knee injury and Osteoarthritis Outcome Score; NPI, Norwich Patellar Instability Score; QOL, Quality of Life; Sport/Rec, Sport and Recreation; TSK, Tampa Scale of Kinesiophobia.

(Table 1). The predefined hypotheses were based on former validation studies<sup>10</sup> on the BPII 2.0, findings from previous translations,<sup>4,39</sup> and clinical experience. We expected PROMs that measure similar constructs, particularly the KOOS and IKDC, to have large positive correlations with the BPII 2.0–No. Measures of patellofemoral instability (NPI) and kinesiophobia (TSK) were expected to have large negative correlations with the BPII 2.0–No because of the inverse nature of the scales. Further, as knee function is assumed to affect quality of life, we also included hypotheses about associations between functional tests (YBT-LQ and hop tests) and the BPII 2.0–No. We expected functional tests to have a small to medium positive correlation with the BPII 2.0–No, as functional tests only address the physical dimension of quality of life. All correlations were investigated using the Pearson correlation coefficient ( $r$ ), where 0.10–0.29 was considered small, 0.30–0.49 medium, and 0.50–1 large.<sup>5</sup>

Internal consistency was assessed by the Cronbach alpha coefficient ( $\alpha$ ), where .70 is acceptable,  $\geq .80$  is preferable, and  $> .95$  might indicate item redundancy.<sup>36</sup> We also examined the floor and ceiling effects, defined as  $>15\%$  of participants having the minimum or maximum score. Test-retest reliability was examined in a subgroup of 50 patients who completed the BPII 2.0–No at both 2 weeks before the 6-month follow up and at a 6-month follow-up. Reliability was calculated using the intraclass correlation coefficient ( $ICC_{2,1}$ ) with 95% CI based on 2-way

TABLE 2  
Baseline Patient Characteristics (N = 100)<sup>a</sup>

Characteristic	Value
Age at surgery, y	22.7 $\pm$ 6.4
Female sex	71 (71)
Years since the first dislocation	7.1 $\pm$ 6.1
Bilateral problems	53 (53)
BMI, kg/m <sup>2</sup>	25.6 $\pm$ 5.4
Surgical procedure <sup>b</sup>	
MPFL-R	21 (22.1)
Combined surgery	74 (77.9)

<sup>a</sup>Data are reported as mean  $\pm$  SD or n (%). BMI, body mass index, MPFL-R, medial patellofemoral ligament reconstruction.

<sup>b</sup>Data for 5 patients are missing because of postponed surgery.

random, single measures with absolute agreement.<sup>36</sup> An ICC of 0.70–0.89 indicates a high correlation and 0.90–1 indicates a very high correlation.<sup>6</sup> The standard error of measurement was calculated from the mean of the variance between tests.<sup>36</sup> A 95% CI of standard error of measurement was made to suggest the limits of measurement error ( $1.96 \times SEM$ ). To express the smallest change score (with  $P \leq .05$ ) that can be interpreted as a real change and not measurement error, the smallest detectable change (SDC) at the individual level ( $SDC_{ind}$ ) was calculated based on the SEM ( $1.96 \times \sqrt{2} \times SEM$ ). The SDC on the group level ( $SDC_{group}$ ) was calculated as  $SDC_{ind}/\sqrt{n}$ , where  $n$  represents the number of patients returning the BPII 2 weeks before the six months follow-up (N = 50).<sup>36</sup> A Bland-Altman plot was used to evaluate the limits of agreement (LoA).<sup>6</sup>

## RESULTS

### Patient Characteristics

The 100 study patients had a mean age of 22.7  $\pm$  6.4 years, and 71% were women. The mean time from first dislocation to surgery was 7.1  $\pm$  6.1 years (range, 0–27 years) and 53% of patients had bilateral problems. Patient characteristics are described in Table 2.

### Data Quality

Overall, patients were able to complete the BPII 2.0–No without assistance. Only 1 patient had 1 missing item. The mean BPII 2.0 score was 42.7  $\pm$  17.3 before surgery and 65.4  $\pm$  20.2 at the 6-month follow-up. There were no floor or ceiling effects for the overall score, as none of the patients had either the lowest possible score (0) or the highest score (100) at the 6-month follow-up.

### Measurement Properties

The expert committee agreed that the BPII 2.0–No had good face validity. Further, support for good content validity was found as interviewed patients reported (1) a high

TABLE 3  
Descriptive Statistics at 6-Month Follow-up on Measurements Used in Hypothesis Testing (n = 72)<sup>a</sup>

Hypothesis	Measurement	Mean ± SD (Range)	Correlation Analysis <sup>b</sup>	
			r	P
1	NPI	10.5 ± 11.2 (0–39.2)	–0.485	.001
2	TSK <sup>c</sup>	27.7 ± 6.6 (15–44)	–0.579	.001
3	IKDC <sup>d</sup>	69.1 ± 15.0 (35.6–100)	0.564	.001
4	KOOS–Pain <sup>d</sup>	83.4 ± 13.7 (42–100)	0.579	.001
5	KOOS–Symptoms <sup>d</sup>	77.6 ± 13.6 (50–100)	0.448	.001
6	KOOS–ADL <sup>d</sup>	93.6 ± 7.6 (72–100)	0.472	.001
7	KOOS–Sport/Rec <sup>d</sup>	64.6 ± 23.1 (10–100)	0.547	.001
8	KOOS–QOL <sup>d</sup>	58.7 ± 20.7 (6–100)	0.723	.001
9	YBT–LQ, reach distance, % <sup>e</sup>	73.1 ± 8.8 (52.7–100.4)	–0.031	NS
9	Hop test, LSI% <sup>f</sup>	90.8 ± 15.3 (38–124)	–0.109	NS
9	PT extension 60 deg/s, N·m <sup>g</sup>	91.1 ± 41.9 (13–184.5)	0.206	NS
9	PT flexion 60 deg/s, N·m <sup>g</sup>	65.5 ± 24.1 (19.9–134.5)	–0.005	NS

<sup>a</sup>ADL, Activities of Daily Living; IKDC, International Knee Documentation Committee Subjective Knee Form 2000; Sport/Rec, Sport and Recreation; KOOS, Knee injury and Osteoarthritis Outcome Score; LSI, limb symmetry index; NPI, Norwich Patellar Instability Score; NS, not significant; PT, peak torque; QOL, Quality of Life; TSK, Tampa Scale of Kinesiophobia; YBT–LQ, Y–Balance Test–Lower Quarter.

<sup>b</sup>For correlation with the 6-month follow-up BPII 2.0–No.

<sup>c</sup>2 missing TSK questionnaires.

<sup>d</sup>2 missing KOOS and IKDC questionnaires.

<sup>e</sup>62 patients completed the YBT–LQ.

<sup>f</sup>55 patients completed hop tests.

<sup>g</sup>60 patients completed the isokinetic strength tests.

relevance of included items, (2) no missing key aspects of their knee-related QOL, and (3) comprehensible instructions and questions of the BPII 2.0–No. Further, the patients interpreted the instructions and questions as intended.

**Construct Validity.** The BPII 2.0–No displayed medium negative correlations with the NPI ( $r = -0.485$ ) and a large negative correlation with the TSK ( $r = -0.579$ ). Further, large correlations between the BPII 2.0–No and the IKDC ( $r = 0.564$ ), and medium to large correlations with all KOOS subscales ( $r = 0.448$  to  $0.723$ ) were evident. There was no correlation between physical tests (YBT–LQ, hop tests, and strength tests) and the BPII 2.0–No. In total, 8 of the 9 predefined hypotheses were confirmed (Table 3).

Internal consistency was excellent for the total BPII 2.0 questionnaire, with a Cronbach  $\alpha$  of .95 ( $n = 99$ ). The  $\alpha$  value varied from .95 to .96 if any of the items were deleted. The test-retest reliability of the BPII 2.0–No was high, with an ICC<sub>2,1</sub> of 0.87 (0.77–0.93) (Table 4). The SEM was 7.1, indicating that a change in score for 1 individual must exceed 19.7 points (SDC<sub>ind</sub>) and on the group level must exceed 2.8 (SDC<sub>group</sub>) points to be interpreted as a true change (exceeding measurement error). A graphic presentation of the LoA is presented in a Bland–Altman plot (Figure 2). The upper limit was 16.8 and the lower limit was –21.1 points.

## DISCUSSION

The BPII 2.0 was, as the first disease-specific PROM, successfully translated and cross-culturally adapted for use in

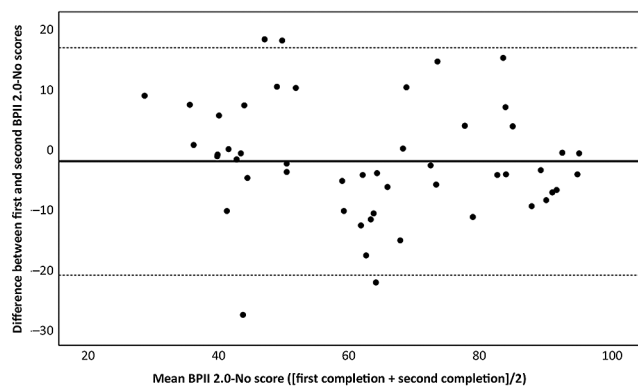
TABLE 4  
Test-Retest Reliability, Measurement Error, and SDC of the BPII 2.0–No (n = 50)<sup>a</sup>

Variable	Value
BPII 2.0–No, first administration, mean ± SD	62 ± 18.5
BPII 2.0–No, second administration, mean ± SD	64.9 ± 20.6
Mean difference	2.9
ICC <sub>2,1</sub> (95% CI)	0.87 (0.77–0.93)
SEM	7.10
1.96 × SEM	13.91
SDC <sub>ind</sub>	19.67
SDC <sub>group</sub>	2.78

<sup>a</sup>BPII 2.0–No, Banff Patellofemoral Instability Instrument–Norwegian version; ICC, intraclass correlation coefficient; SDC, smallest detectable change; ind, individual.

Norwegian-speaking patients with recurrent patellofemoral instability. The present study indicated that the BPII 2.0–No is relevant and comprehensible, and overall it holds acceptable standards when used to assess these patients. Construct validity was acceptable and good test-retest reliability was demonstrated. The SDC was 19.7 points, indicating that changes in scores in 1 individual need to exceed this number to be interpreted as a “true” change.

When analyzing the data quality of the BPII 2.0, no floor or ceiling effects were found. This is in line with findings from the original publication and other language validation studies,<sup>4,15</sup> indicating that the questionnaire can capture changes at both ends of the score range.<sup>36</sup> Support



**Figure 2.** Bland-Altman plot displaying limits of agreement ( $n = 50$ ). BPII 2.0–No, Banff Patellofemoral Instability Instrument–Norwegian version.

for good construct validity was found, as 8 of the 9 predefined hypotheses were confirmed. The BPII 2.0 showed, as presumed, the largest associations with the KOOS-QOL ( $r = 0.723$ ). This finding is comparable with the Dutch adaptation of the first version of the BPII.<sup>39</sup> The moderate negative association with perceived symptoms of patellar instability (NPI) ( $r = -0.485$ ) seen in the present study also corresponds to results from 2 other studies<sup>4,10</sup> on the BPII 2.0 and the BPII. As assumed, a large negative association between the BPII 2.0 and kinesiophobia measured by the TSK was found ( $r = -0.579$ ). It was expected that fear of reinjuries because of increased movement and physical activity could have a negative impact on quality of life. Shams et al.<sup>31</sup> also reported higher TSK scores in patients who had undergone medial patellofemoral ligament reconstruction compared with healthy adults.

It is reasonable to assume that impaired knee function can have an impact on a person's quality of life through restrictions in daily life and sport/recreational activities, and this is supported by the present large association ( $r = 0.56$ ) between the BPII 2.0 and self-perceived knee function (IKDC). Counterintuitively, this was not the case when knee function was measured using functional tests in the current cohort. As this was the first study to assess the associations between the BPII 2.0 and functional tests, there were no studies to inform what associations to expect. Nonetheless, it was assumed that measured functional performance would have a small, statistically significant association with quality of life. This was not supported in the present study, as no association between the functional tests and the BPII 2.0 was present. A possible explanation for this is that the selected functional tests do not capture aspects of functioning that are relevant for quality of life for patients with patellofemoral instability. Another explanation is that there may be a mismatch between the patient's perception of physical ability and actual performance.

The internal consistency of the BPII 2.0–No was excellent ( $\alpha = .95$ ) and in line with values reported by Lafave et al<sup>15</sup> at the 6-month follow-up after surgery. This indicates that the questionnaire measures 1 single construct. At the same time, the Cronbach  $\alpha$  did not change

significantly when items were deleted, indicating item redundancy. This was further supported by high interitem correlation values (Supplemental Table S1, available separately). Several of the interitem correlations approached the limit of 0.7 described by de Vet et al,<sup>6</sup> indicating that several items capture the same aspect of quality of life. Hence, a further item reduction may be needed. The first version of the BPII underwent principal component analysis,<sup>15</sup> leading to a shorter version (BPII 2.0), from which the Norwegian version is translated. The principal component analysis is, however, a data reduction method, computed without regard to any underlying structure caused by latent variables. Studies performing exploratory factor analysis would therefore be valuable to determine the dimensionality of the questionnaire.<sup>6</sup>

The present high test-retest reliability is comparable to previous results.<sup>4,11,15-17</sup> The 2-week interval between completions should be long enough to prevent recall and short enough to minimize the risk of changes in the patient's condition.<sup>36</sup> The patients' quality of life was expected to be stable in this relatively short time. Our patients are not completely comparable to those of the study by Lafave et al,<sup>15</sup> as they investigated test-retest reliability before surgery, while we tested test-retest reliability 6 months after surgery. Consequently, in their study, the mean scores on the BPII 2.0 were lower than those in the present study.

This is the first study to report the SEM of the BPII 2.0 with a sufficient method according to recommendations from the COSMIN.<sup>6</sup> Interestingly, our SEM value was considerably higher than previously reported (7.10 vs 2.64 and 2.13).<sup>15,17</sup> Possible explanations for this discrepancy may be the method used to obtain the SEM value. While the SEM in this study is derived from the mean of the variance between tests, the other 2 values are calculated from the standard deviations, a method de Vet et al.<sup>6</sup> warn against using because it does not take into account systematic differences. In addition, the insufficient sample size in the study by Lafave et al<sup>17</sup> might affect their results. This is also the first study to report the 95% LoA and SDC for the BPII 2.0, thereby providing more detailed information on the smallest change in score that can be interpreted as a "real" change above measurement error in 1 individual ( $SDC_{ind}$ ).<sup>36</sup> The  $SDC_{ind}$  estimated in the present study indicates that only a change of  $>19.7$  points on the BPII 2.0 can be considered as a "real" within-person change for patients with recurrent patellofemoral instability. For the BPII 2.0, there is no established minimal important change assessed with, for example, a receiver operating characteristic curve, as recommended by de Vet et al.<sup>6</sup> Consequently, interpretation of changes in the BPII 2.0 must be done with this in mind, and future studies should address this limitation.

The prospective inclusion of patients in the present study represents a strength. The high standard deviations found in the present analysis are in line with previous work<sup>4,15</sup> and indicate that a heterogeneous population was evaluated, as is the case with patellar instability patients.<sup>40</sup> As the Norwegian-speaking population is relatively homogeneous, we assume that the inclusion of 86% of all patients undergoing surgical treatment for recurrent patellar dislocation in 2 orthopaedic units constitutes a representative,

unselected, cohort of this group of patients in Norway. Moreover, mapping quality of life is a reasonable tool to utilize when evaluating the effect of surgery.

### Limitations

There are some limitations to consider. This study was conducted in 2 orthopaedic units; however, a multicenter study would increase the scientific value of the results. Few cross-cultural validation studies of the BPII 2.0 to other languages exist; thus, comparison with other studies is limited. Although the IKDC is validated in patients with patellofemoral instability,<sup>25</sup> demonstrating sufficient reliability, the presence of a substantial ceiling effect may compromise the validity of the questionnaire in this population.<sup>25</sup> The existing Norwegian version of the IKDC used in the present study has not undergone a recommended assessment of measurement properties.<sup>8</sup> The KOOS questionnaire is frequently used in patients with patellofemoral instability but has not been validated in this population.<sup>12</sup> In addition, a validated Norwegian version of KOOS has not been published.<sup>8</sup> Even though the TSK measures aspects (fear of movement) that may be relevant for patients with patellofemoral instability, no information on validity and reliability in this patient group exists. Although there is limited information on the measurement properties of the Norwegian versions of the questionnaires, they are in extensive use and well accepted in research and clinical communities. The functional tests are also in extensive use after knee injuries; however, their reliability and validity in patients with patellofemoral instability have not been established. Therefore, limitations in the comparative use of these instruments should be acknowledged.

### CONCLUSION

The Norwegian version of BPII 2.0 was successfully translated and cross-culturally adapted into Norwegian. The support for content validity in the present study indicates that the items of the BPII 2.0–No reflect relevant and important aspects of quality of life in patients with patellofemoral instability. The scale has good construct validity and reproducibility. This study also highlights the need to perform an exploratory factor analysis to establish the factor structure of the BPII 2.0. The present study adds to the growing evidence on the validity and reliability of the BPII 2.0 in accordance with the COSMIN guidelines.

### ACKNOWLEDGMENT

The authors thank Laurie A. Hiemstra, MD, PhD, at the Banff Sport Medicine Foundation, for her collaboration during the translation process.

### REFERENCES

1. Anderson AF, Irrgang JJ, Kocher MS, Mann BJ, Harrast JJ. The International Knee Documentation Committee Subjective Knee Evaluation Form: normative data. *Am J Sports Med.* 2006;34(1):128-135.
2. Ardern CL, Taylor NF, Feller JA, Whitehead TS, Webster KE. Psychological responses matter in returning to preinjury level of sport after anterior cruciate ligament reconstruction surgery. *Am J Sports Med.* 2013;41(7):1549-1558.
3. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000;25(24):3186-3191.
4. Becher C, Attal R, Balcarek P, et al. Successful adaptation of the Banff Patella Instability Instrument (BPII) 2.0 into German. *Knee Surg Sports Traumatol Arthrosc.* 2018;26(9):2679-2684.
5. Cohen J. *Statistical Power Analysis for the Behavioural Sciences.* 2nd ed. Lawrence Erlbaum Associates; 1988.
6. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide.* Cambridge University Press; 2011.
7. Drouin JM, Valovich-mcLeod TC, Shultz SJ, Gansneder BM, Perrin DH. Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. *Eur J Appl Physiol.* 2004;91(1):22-29.
8. Faleide AGH, Inderhaug E, Vervaat W, et al. Anterior cruciate ligament-return to sport after injury scale: validation of the Norwegian language version. *Knee Surg Sports Traumatol Arthrosc.* 2020; 28(8):2634-2643.
9. Haugen AJ, Grøvre L, Keller A, Grotle M. Cross-cultural adaptation and validation of the Norwegian version of the Tampa scale for kinesiophobia. *Spine (Phila Pa 1976).* 2008;33(17):e595-e601.
10. Hiemstra LA, Kerslake S, Lafave M, Mohtadi NG. Concurrent validation of the Banff Patella Instability Instrument to the Norwich Patellar Instability Score and the Kujala Score in patients with patellofemoral instability. *Orthop J Sports Med.* 2016;4(5):2325967116646085.
11. Hiemstra LA, Kerslake S, Lafave MR, et al. Initial validity and reliability of the Banff Patella Instability Instrument. *Am J Sports Med.* 2013;41(7):1629-1635.
12. Hiemstra LA, Page JL, Kerslake S. Patient-reported outcome measures for patellofemoral instability: a critical review. *Curr Rev Musculoskelet Med.* 2019;12(2):124-137.
13. 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(5):600-613.
14. Kvist J, Ek A, Sporrstedt K, Good L. Fear of re-injury: a hindrance for returning to sports after anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc.* 2005;13(5):393-397.
15. Lafave MR, Hiemstra L, Kerslake S. Factor analysis and item reduction of the Banff Patella Instability Instrument (BPII): introduction of BPII 2.0. *Am J Sports Med.* 2016;44(8):2081-2086.
16. Lafave MR, Hiemstra LA, Kerslake S. Validity, reliability, and responsiveness of the Banff Patellar Instability Instrument (BPII) in an adolescent population. *J Pediatr Orthop.* 2018;38(10):e629-e633.
17. Lafave MR, Hiemstra LA, Parikh SN, Peterson D, Kerslake S. Validity and reliability of the Banff Patellofemoral Instability Instrument 2.0 in an adolescent population. *J Pediatr Orthop.* 2020;40(2):e103-e108.
18. Liu JN, Steinhaus ME, Kalbian IL, et al. Patellar instability management: a survey of the International Patellofemoral Study Group. *Am J Sports Med.* 2018;46(13):3299-3306.
19. Logerstedt D, Grindem H, Lynch A, et al. Single-legged hop tests as predictors of self-reported knee function after anterior cruciate ligament reconstruction: the Delaware-Oslo ACL cohort study. *Am J Sports Med.* 2012;40(10):2348-2356.
20. Miller RP, Kori SH, Todd DD. The Tampa Scale: a measure of kinesiophobia. *Clin J Pain.* 1991;7(1):51.
21. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:22.

22. 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(7):737-745.
23. Noyes FR, Barber SD, Mangine RE. Abnormal lower limb symmetry determined by function hop tests after anterior cruciate ligament rupture. *Am J Sports Med*. 1991;19(5):513-518.
24. Paterno MV, Flynn K, Thomas S, Schmitt LC. Self-reported fear predicts functional performance and second ACL injury after ACL reconstruction and return to sport: a pilot study. *Sports Health*. 2018;10(3):228-233.
25. Paxton EW, Fithian DC, Stone ML, Silva P. The reliability and validity of knee-specific and general health instruments in assessing acute patellar dislocation outcomes. *Am J Sports Med*. 2003;31(4):487-492.
26. Plisky PJ, Gorman PP, Butler RJ, et al. The reliability of an instrumented device for measuring components of the star excursion balance test. *N Am J Sports Phys Ther*. 2009;4(2):92-99.
27. Rhatomy S, Pontoh LA, Phatama KY, et al. The Banff Patellar Instability Instrument: validity and reliability of an Indonesian version. *Eur J Orthop Surg Traumatol*. 2023;33(3):617-622. doi:10.1007/s00590-022-03336-6
28. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynonn BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sports Phys Ther*. 1998;28(2):88-96.
29. Saper MG, Fantozzi P, Bompadre V, Racicot M, Schmale GA. Return-to-sport testing after medial patellofemoral ligament reconstruction in adolescent athletes. *Orthop J Sports Med*. 2019;7(3):2325967119828953.
30. Shaffer SW, Teyhen DS, Lorenson CL, et al. Y-balance test: a reliability study involving multiple raters. *Mil Med*. 2013;178(11):1264-1270.
31. Shams K, DiCesare CA, Grawe BM, et al. Biomechanical and functional outcomes after medial patellofemoral ligament reconstruction: a pilot study. *Orthop J Sports Med*. 2019;7(2):2325967119825854.
32. Smith TO, Chester R, Hunt N, et al. The Norwich Patellar Instability Score: validity, internal consistency and responsiveness for people conservatively-managed following first-time patellar dislocation. *Knee*. 2016;23(2):256-260.
33. Smith TO, Choudhury A, Navratil R, Hing CB. Psychometric properties of the Norwich Patellar Instability Score in people with recurrent patellar dislocation. *Knee*. 2019;26(6):1192-1197.
34. Smith TO, Donell ST, Clark A, et al. The development, validation and internal consistency of the Norwich Patellar Instability (NPI) score. *Knee Surg Sports Traumatol Arthrosc*. 2014;22(2):324-335.
35. Sole G, Hamrén J, Milosavljevic S, Nicholson H, Sullivan SJ. Test-retest reliability of isokinetic knee extension and flexion. *Arch Phys Med Rehabil*. 2007;88(5):626-631.
36. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
37. Tripp D, Stanish W, Ebel-Lam A, Brewer B, Birchard J. Fear of reinjury, negative affect, and catastrophizing predicting return to sport in recreational athletes with anterior cruciate ligament injuries at 1 year postsurgery. *Sport Exerc Perform Psychol*. 2011;1:74-81.
38. Undheim MB, Cosgrave C, King E, et al. Isokinetic muscle strength and readiness to return to sport following anterior cruciate ligament reconstruction: is there an association? A systematic review and a protocol recommendation. *Br J Sports Med*. 2015;49(20):1305-1310.
39. Van Sambeeck JD, Van de Groes SA, Koeter S. Dutch translation and validation of the Norwich Patellar Instability Score and Banff Patella Instability Instrument in patients after surgery for patellar instability. *Acta Orthop Belg*. 2020;86(3):470-481.
40. Weber AE, Nathani A, Dines JS, et al. An algorithmic approach to the management of recurrent lateral patellar dislocation. *J Bone Joint Surg Am*. 2016;98(5):417-427.
41. White AE, Chatterji R, Zaman SU, et al. Development of a return to play checklist following patellar instability surgery: a Delphi-based consensus. *Knee Surg Sports Traumatol Arthrosc*. 2020;28(3):806-815.
42. Willis GB, Artino AR Jr. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *J Grad Med Educ*. 2013;5(3):353-356.
43. Zimmerer A, Sobau C, Balcarek P. Recent developments in evaluation and treatment of lateral patellar instability. *J Exp Orthop*. 2018;5(1):3.