RESEARCH ARTICLE   **Open Access**

Translation, cross-cultural adaptation, and validation of the Italian language Forgotten Joint Score-12 (FJS-12) as an outcome measure for total knee arthroplasty in an Italian population

Valerio Sansone[1,2] , Peter Fennema[3], Rachel C. Applefield[2], Stefano Marchina[2], Raffaella Ronco[4], Walter Pascale[1,2] and Valerio Pascale[1,2*]

Abstract

Background: With ever-increasing numbers of patients undergoing elective total knee arthroplasty (TKA) and as TKA is performed in increasingly younger patients, patient demands and expectations have also increased. With improved patient outcomes, new PROMs with heightened discriminatory power in well-performing patients are needed. The present study aimed to translate and validate the Italian version of the Forgotten Joint Score (FJS-12) as a tool for evaluating pre-operative through longitudinal post-operative outcomes in an Italian population.

Methods: In this prospective study, patients with unilateral osteoarthritis, undergoing TKA surgery between May 2015 and December 2017 were recruited to participate in the study. The FJS-12 and WOMAC were collected pre-operatively and at six and 12 months post-operatively. According to the COSMIN checklist, reliability, internal consistency, validity, responsiveness, effect size, and ceiling effects and floor effects were evaluated.

Results: One hundred twenty patients completed the study, 66 of which participated in the evaluation of test-retest reliability. Good test-retest reliability was found (ICC = 0.90). The FJS-12 also showed excellent internal consistency (Cronbach α = 0.81). Construct validity with the WOMAC, as a measure of the Pearson correlation coefficient, was moderate (r = 0.45 pre-operatively; r = 0.46 at 6 months and r = 0.42 at 12 months post-operatively). From six to 12 months, the change was slightly greater for the WOMAC than for the FJS-12 patients (effect size d = 0.94; d = 0.75, respectively). At 12-months follow-up, the ceiling effects reflecting the maximum score were 12% for the FJS-12 and 6% for the WOMAC; however, scores within 10% of the maximum score were comprised 30% of the FJS-12 scores and 59% for the WOMAC.

(Continued on next page)

* Correspondence: valerio.pascale@unimi.it
[1]Department of Orthopaedics, Università degli Studi di Milano, Via Festa del Perdono, 7, 20122 Milan, MI, Italy
[2]IRCCS Istituto Ortopedico Galeazzi, Via Riccardo Galeazzi, 4, 20161 Milan, MI, Italy
Full list of author information is available at the end of the article

Sansone *et al. BMC Musculoskeletal Disorders* (2020) 21:23

Page 2 of 9

(Continued from previous page)

Conclusion: The Italian FJS-12 demonstrated strong measurement properties in terms of reliability, internal consistency, and construct validity in TKA patients. Furthermore, a more detailed look at ceiling effects shows a superior discriminatory capacity when compared to the WOMAC at 12-months follow-up, particularly in better-performing patients.

Trial registration: clinicaltrials.gov NCT03805490. Registered 18 January 2019 (retrospectively registered).

Keywords: Patient reported outcomes, Total knee arthroplasty, Validation, Translation, Forgotten Joint Score, Arthroplasty, Ceiling effect

Background

Worldwide, the number of total knee arthroplasties (TKAs) is steadily increasing at annual rates ranging from 5.3% in France to 17% in Portugal [1]. In Italy, annual growth in TKA procedures has been reported to be from 6.6% [2] to 12.8% [1]. Remarkably, the number of younger TKA patients has been increasing rapidly [3]; in some populations up to a 3-fold has been seen over the last twenty years [4]. Consequently, increased patient expectations and improved outcomes have been reported [5]. Therefore, instruments which seek to measure outcomes after TKA must adapt to meet the changing expectations [6, 7].

Historically, successful interventions have been measured by concrete, surgeon-centred assessments such as range of motion, joint stability, and implant survival. In recent years, however, patient-reported outcome measurements (PROMs) have become more common and provide a new tool for evaluating the everyday life of patients after TKA [8]. The ability to discriminate between patients with good or excellent outcome from patients with poorer outcomes can be determined by implementing one of several validated scores, such as the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [9] and the Knee Society Score [10]; however, a considerable ceiling effect has been detected among these scores [11, 12], revealing a weakness in the capacity to differentiate between good and excellent scores. As joint arthroplasty has evolved, outcomes have continued to improve, and as a consequence patient demands and expectations have also increased. With improved patient outcomes, new PROMs with heightened discriminatory power in well-performing patients are needed [7]. In 2012, Behrend et al. introduced the FJS-12 as a means of assessing the patient's degree of awareness of their prosthetic joint; it has since been proven to have a lower ceiling effect than other orthopaedic scores [13].

The objective of the present study is to cross-culturally adapt the FJS-12 for use in the Italian population and to test the validity, reliability, and floor and ceiling effects in a group of consecutive patients, assessed from pre-operative presentation through 1-year follow up. To our knowledge, no prior study validating an Italian version of the FJS-12 has been published.

Methods

Validation study

A prospective, observational study was performed to evaluate the reliability, validity and responsiveness of the Italian version of the FJS-12. First, the original FJS-12 [13] was translated using the translation and back-translation method [14]. Subsequently, the translated Italian version of the FJS-12 was validated in patients who received a primary TKA. Our inclusion criteria were the presence of unilateral knee osteoarthritis (Kellgren Lawerence scale of III-IV) requiring a TKA, no previous surgery of the lower limb, and fluency in Italian reading and comprehension. Exclusion criteria deemed ineligible any patient with a history of a previous joint prosthesis of the lower limb, previous operation of the affected knee, previous major cardiac events, current use of a walking aid, or lack of the informed consent. Based on the health status questionnaire guidelines which recommend having at least 10 patients per question [15], the sample size estimation.

For the evaluation of test-retest reliability, only the first 66 consecutive patients were involved. These patients completed the preoperative FJS-12 questionnaire twice, 2 weeks apart, prior to surgery. Our study protocol adhered strictly to the Declaration of Helsinki (EN ISO 14155:1 e EN ISO 14155: 2) and to the guidelines of Good Clinical Practice and was approved by our local ethics committee in May 2015.

Assessment instruments

The Forgotten Joint Score

The FJS-12 was first described in the literature by Behrend et al. in 2012 [13]. It is a PROM for patients who have undergone a TKA; it introduces the concept of a "forgotten" joint as the ideal objective to pursue in prosthetic surgery. Although other

scores have been validated and shown to be reliable, they do not have adequate responsiveness and they show a high ceiling/floor effect. The FJS-12 has shown optimal discriminatory capacity and a reduced ceiling effect compared to other scores [13]. As a result of its effectiveness, the FJS-12 has since been translated in numerous languages such as but not limited to French [16], German [17], Dutch [18], and Chinese [19].

The original FJS-12 is composed of 12 items, measuring the patient's ability to forget the presence of an artificial joint in their daily life. For each item, there is a five-points Likert scale response. The raw results are converted to a 0–100 points scale. Highest score corresponds to good outcome with the patient not aware about the presence of the prosthesis. In the case of more than 4 answers are missing, the score should not be used (Appendix) [13].

WOMAC
The WOMAC is a clinical orthopaedic score that was first reported in 1986. The questionnaire is comprised of 24 total questions, 5 questions on pain, 2 on stiffness, and 17 on function. For each item there is a five-points Likert scale response. Total final WOMAC scores can range from 0, the lowest functional status level, to 100, the best functional status level [9].

The WOMAC score has been translated in 65 different languages, and at present is one of the most prevalently used scores in clinical orthopaedic research. In a recent review, carried out by analysing 76 articles from 22 different countries, the WOMAC once again demonstrated excellent validity and internal consistency (> 0.90) according to the consensus-based standards for the selection of health measurement instruments (COSMIN) criteria [20]. For this reason, the WOMAC was chosen as the comparison questionnaire in the present study.

Translation and adaptation of the Forgotten Joint Score
In the present study, the official Italian translation of the FJS-12 was not used as it was not available at the time our study began. Nevertheless, our translation of the FJS-12 questionnaire from English to Italian was executed according to an internationally accepted method.

The initial translation was performed independently by two native Italian speaking orthopaedic surgeons, who were informed about the score, together with a professional translator, native Italian speaker, who was uninformed about the score. The first version was obtained by a consensus among the three translators. The backwards translation into English was performed independently by two native English speakers with a medical background. The English version

hence obtained was deemed valid in a consensus meeting between all of the involved translators. The Italian version was then tested on a sample of 20 subjects with knee osteoarthritis to evaluate the comprehensibility and simplicity of the writing.

Subsequently, the validation of the translated FJS-12 score was conducted. In the validation phase, 217 patients who were to undergo a TKA were enrolled. After signing an informed consent, the patients completed the FJS-12 and WOMAC scores preoperatively, at 6 months and at 12 months post-operatively during the routine follow-up visits. The questionnaires were never filled out in the presence of the principal or sub-investigators. Patients who for geographic reasons could not complete the questionnaires at our clinic during the follow-up visits, completed the questionnaires by mail.

Reliability
Reliability is defined as the degree to which the measurement is free from measurement error [21]. In order to assess the reliability, internal consistency, test-retest reliability, and measurement error were calculated.

Internal consistency
Internal consistency is determined by the degree of inter-relatedness among questionnaire items [21]. A Cronbach's α of greater than 0.7 was considered sufficient [15].

Test-retest reliability
Test-retest reliability refers to the extent to which results from the same patient, with a static health status, remain unchanged over time [21]. According to the COSMIN manual recommendation, the retest was performed 2 weeks after the first test was administered to avoid relevant changes in health condition. Intraclass correlation coefficients (ICC) were calculated for all patients to ascertain any change of the knee condition since the first evaluation. Sufficient test-retest reliability was assumed for an ICC greater than 0.7 [15].

Measurement error
Measurement error is the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured [21]. The standard error of measurement (SEM) was calculated using the formula $SD * \sqrt{1 - ICC}$ (SD = standard deviation) [15]. The smallest detectable change (SDC) reflects the smallest individual change in a score that can be interpreted as a real change. It was calculated with the following formula: $SEM \times 1.96 \times \sqrt{2}$ [15].

Validity

The validity of the FJS-12 was measured via construct validity, which is the degree to which the scores of a questionnaire are consistent with hypotheses that the questionnaire validly measures the construct to be measured [21]. For a Pearson correlation coefficient $0.3 < r < 0.7$ moderate correlation was assumed. For an $r < 0.3$ the correlation was considered poor. To facilitate this calculation, the questionnaire scores were divided into 10 classes before realizing the evaluation.

Responsiveness

The capacity of a questionnaire to detect a change over time in the construct which is to be measured is referred to as the responsiveness [21]. It was calculated by measuring the change between pre-operative presentation and six-month follow-up, and between the six-months and 12-month follow-ups. This change is reported as effect size (ES) for the mean change in terms of Cohen's d.

Furthermore, the SDC was used as a secondary measure of responsiveness. When the SDC is smaller than the minimal important change (MIC), a positive rating for responsiveness is assumed [15]. The MIC calculation was simply half of the SD [22].

Statistical analysis

Statistical analysis was executed using the software SPSS (IMB SPSS Statistics 21, SPSS INC, Chicago, Illinois). Descriptive data are given as mean ± SD, unless otherwise stated. Ceiling effect was examined by considering the percentage of patients that achieved the maximum score (100) as well as when patients reached a score within 5 and 10% of the maximum achievable score for the FJS-12 and the WOMAC [23]. To be comparable with previous studies, if less than 15% of patients achieved the maximum score, ceiling effects were considered acceptable [15, 24].

Results

Between May 2015 and December 2017, 217 patients were enrolled in our study. Ninety-seven patients were excluded for missing one of the two follow-up time points. Of the remaining 120 patients that could be included in our study, 39 (32.5%) were men and 81 (67.5%) were women. The mean age was 70.0 ± 8.4 years (range 44–86). The mean interval for the 6-month post-operative follow-up was 6.5 ± 0.8 months (range 5–8 months); and the mean interval for the final follow-up at 1 year was 12.5 ± 1.2 months (range 10–15). The FJS-12 and WOMAC pre-operative, 6-month follow-up, and 12-month follow-up outcomes are summarized in Table 1. To evaluate the test-retest

Table 1 Mean scores for FJS-12 and WOMAC at all time points

	Pre-op	6 m follow-up	12 m follow-up
FJS-12	24.5 ± 16.6	60.5 ± 25.0	73.1 ± 23.4
WOMAC	41.1 ± 16.8	74.0 ± 18.4	85.6 ± 15.7
1-way ANOVA	$p < 0.001$	$p < 0.001$	$p < 0.001$

reliability, the first 66 patients were included and the time span between the first and second pre-operative questionnaire was 14 days ±2 days.
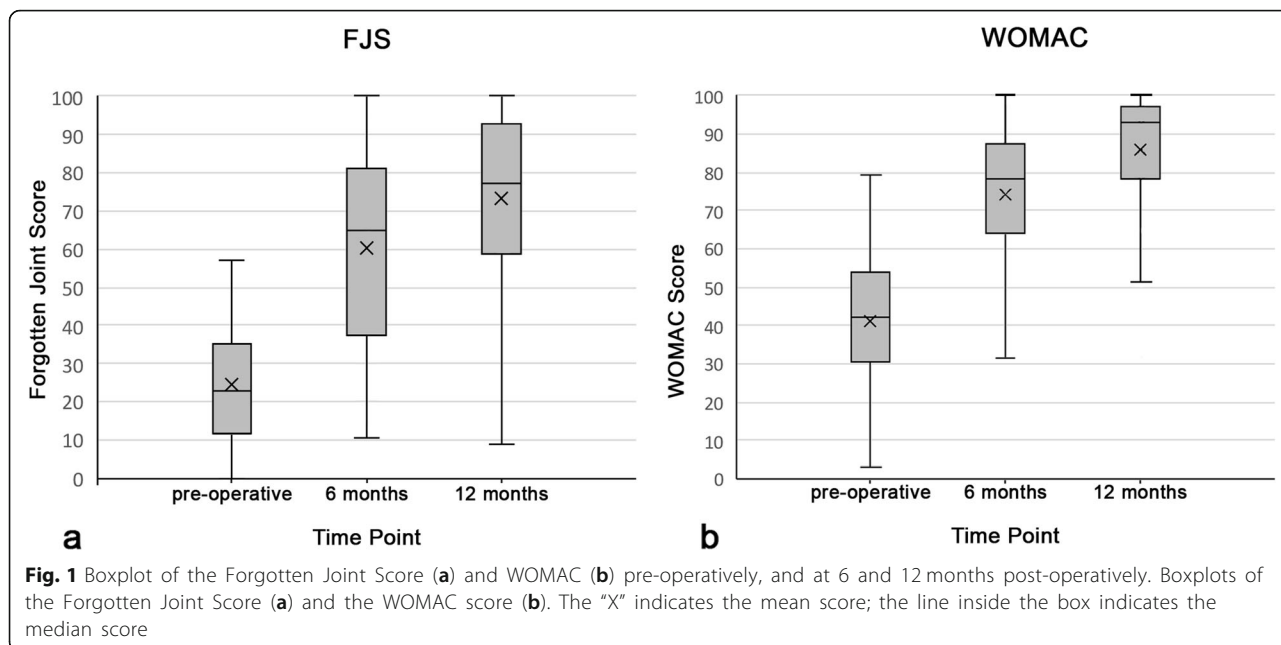
14 total complications were recorded. There were 12 cases (10%) of anaemia, 1 case (0.8%) of acute pulmonary oedema and atrial fibrillation, and 1 case (0.8%) of hypoglycaemia.

A Cronbach's α of 0.81 (95% CI: 0.76–0.86, $p < 0.001$) indicated excellent internal consistency for the FJS-12. As this value is greater than 0.7, sufficient internal consistency was assumed. The test-retest reliability was also good for all patients, with an ICC of 0.90 (95% CI: 0.86–0.92, $p < 0.001$) The SEM for the total FJS-12 was 5.25; therefore, the SDC, reflecting the smallest individual change in a score that can be considered as real change, was 14.55.

The Pearson correlation coefficient between the FJS-12 and WOMAC was 0.45 ($p < 0.001$) at pre-op, $r = 0.46$ ($p < 0.001$) at 6 months, and $r = 0.42$ ($p < 0.001$) 12 m. Therefore, construct validity showed a moderate correlation between the two functional scores at all time points.

To measure the performance of the FJS-12 over time, the data was analysed at all time-intervals after surgery (Fig. 1). The results are detailed in Table 1. Preoperatively, the mean FJS-12 score was 24.4, and at 6 months the mean was 60.5 (Table 1), resulting, as expected, in a very large ES for the time span from pre-operative to six-month follow-up (Cohen's d = 1.84). The WOMAC also demonstrated a very large ES (Cohen's d = 1.29), with mean pre-op scores of 41.1 and mean 6-month follow-up scores of 74.0. For the interval between the six month and one-year follow-up, a medium effect size (d = 0.75) was found for FJS-12 and a high effect size (d = 0.94) for WOMAC. Lastly, as the pre-operative MIC was 8.3, while the pre-operative SDC was 14.55, according to Terwee et al. [15], no positive rating for responsiveness could be given as SDC was greater than MIC. The MICs for both the FJS-12 and WOMAC at all three time points can be found in Table 2.

When the maximum score of 100 was used as the cut-off for measuring the ceiling effect at the 12-month follow-up, the ceiling effect was less than 15%, and therefore acceptable, for both the FJS-12 (12%) and WOMAC (6%). However, when plotted, a visible

Sansone *et al. BMC Musculoskeletal Disorders* (2020) 21:23

Page 5 of 9

Fig. 1 Boxplot of the Forgotten Joint Score (**a**) and WOMAC (**b**) pre-operatively, and at 6 and 12 months post-operatively. Boxplots of the Forgotten Joint Score (**a**) and the WOMAC score (**b**). The "X" indicates the mean score; the line inside the box indicates the median score

"ceiling" can be seen in both scores (Fig. 2), although the FJS-12 succeeds in distributing the scores more evenly than the WOMAC. Plotting a bisector, reveals that most subjects have a lower FJS-12 than WOMAC score, confirming the hypothesis of FJS-12 having a lower ceiling effect. Table 3 shows the ceiling effects at three different cut-off scores. At 12 months post-operatively, more than a third of WOMAC scores were above 95, compared to 19% of the FJS-12 scores. This disparity greatens further when the cut-off is lowered to scores greater than 90, which comprised more than half (59%) of the WOMAC scores and only 30% of the FJS-12 scores.

Based on the pre-op scores, both the FJS-12 (5%) and WOMAC (0%) were in the acceptable range (of < 15%) for floor effects. Furthermore, at 12 months follow-up, 0 patients recorded the minimum score.

At all three time points, question 12, was left unanswered by a large number of patients (21.7% at pre-op, 14.2% at 6 months post-op, and 7.5% at 12 months post-op). Question 7 was left unanswered by 2.5% at pre-op and 6 months post-op and 0.8% at 12 months

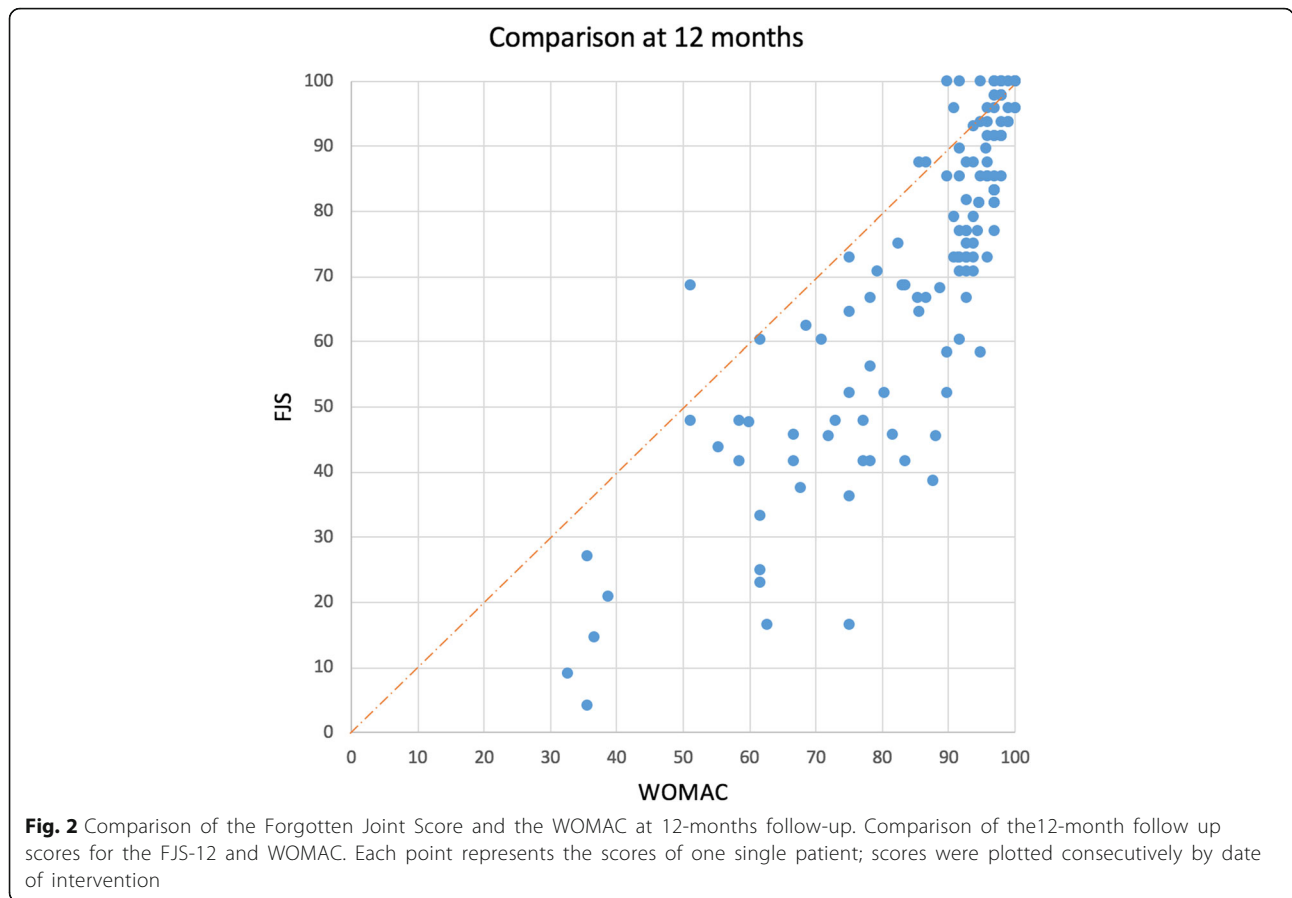post-op. Finally, question 4 was left unanswered by 3.3% of patients only at pre-op.

Discussion

In order for patients to "forget" their artificial knee, in addition to being free of pain, it must offer an acceptable range of movement and provide stability during daily activities. The original FJS-12 is a PROM that comprises all three of these parameters in the assessment of the post-operative outcome.

This is the first study to validate the Italian version of the FJS-12 using the COSMIN checklist [21]. Sufficient validity and reliability in the Italian version of the FJS-12 were confirmed. Moderate correlation was found between the FJS-12 and the WOMAC at all time points. According to Terwee and colleagues [15], the FJS-12 did not achieve a positive rating for responsiveness as the MIC (8.3) was higher than the SDC (17.7). However, for FJS-12, effect size was very large between pre-op and six months (1.84), and from 6 months to 12 months there was a medium effect size (d = 0.75). Furthermore, Although the FJS-12 was originally intended to measure post-operative outcomes, our study shows that it can also be used across the patient's entire surgical pathway, evaluating pre-operative and longitudinal post-operative outcomes.

The mean age of 70 years for patients in our study is comparable to the average age in other knee arthroplasty studies [13, 25]. The substantial number of patients (21.7% at pre-op and 7.5% at 12-month

Table 2 Minimal important change (MIC) for FJS-12 and WOMAC at all time points

	Pre-op	6 m follow-up	12 m follow-up
FJS-12	8.3	12.5	11.7
WOMAC	8.4	9.2	7.8

Fig. 2 Comparison of the Forgotten Joint Score and the WOMAC at 12-months follow-up. Comparison of the12-month follow up scores for the FJS-12 and WOMAC. Each point represents the scores of one single patient; scores were plotted consecutively by date of intervention

follow-up) who left question 12, regarding sports, un-answered has been noted in other studies of FJS-12, such as in the French study where 40% of patients left question 12 unanswered. In future versions of the questionnaire, it would be wise to consider rephrasing this question to improve the applicability for Italians and increase the likelihood of subjects to complete this item.

The original FJS-12 publication found a Cronbach's α of 0.95, and a positive rating for internal consistency was given [13]. Unlike the original FJS-12 study, in the present study, test-retest reliability was evaluated pre-operatively, and excellent internal consistency was demonstrated via a Cronbach's α of 0.81. Likewise, two other validation studies, one Danish [26] and one German [17], both found a high level of internal

Table 3 Ceiling effects at 12-months follow-up for the FJS-12 and WOMAC

Maximum score (100)		score of > 95%		score of > 90%	
FJS-12	WOMAC	FJS-12	WOMAC	FJS-12	WOMAC
12%	6%	19%	34%	30%	59%

consistency (Cronbach's α = 0.96 and Cronbach's α = 0.95, respectively). Several other studies have investigated test-retest reliability via calculation of the ICC, all finding high internal consistency.[17,18,19,26] In the present study, good test-retest reliability was confirmed with an ICC of 0.90 [15]; this value was comparable to the German study's ICC of 0.80 [17], the Dutch study's ICC of 0.94 [18], and the Chinese study's ICC of 0.97 [19], and the Danish study's ICC of 0.91 [26].

Similar to the original FJS-12 [17], it was assumed that the WOMAC, which is widely used in Italy, would be a good tool with which to compare the FJS-12. However, unlike the high correlation between FJS-12 and WOMAC seen in previous studies previous studies [13, 27], we found only a moderate correlation between the two functional scores. The differences in the phrasing of the items among the two questionnaires makes it unsurprising that only a modest correlation would be found.

Between pre-operative presentation and 6 months post-operatively, the FJS-12 demonstrated a very large effect size (Cohen's d = 1.84) that was slightly larger than the effect size found by the WOMAC

Sansone *et al. BMC Musculoskeletal Disorders*   (2020) 21:23

Page 7 of 9

(Cohen's d = 1.29). However, for the interval between 6 months and 12 months post-operatively only a medium effect size was found (Cohen's d = 0.75), and this change was slightly less than the change detected via the WOMAC (Cohen's d = 0.94). Our results differ from Hamilton et al.'s 2017 study in which they found the FJS-12 to be more responsive in the latter period of follow-up when compared to another Orthopaedic score; however, their effect size (Cohen's d = 0.12) was much smaller than in our study.

According to Terwee et al. [15], no positive rating for responsiveness could be given for the Italian version of the FJS-12 as our MIC (8.3) was less than our SDC (17.7). Our result was comparable with Bauman et al.'s findings (MIC = 10.9; SDC = 13.1), while Ingelsrud et al. [28] found a higher value for MIC [14].

The WOMAC is one of the most commonly used Orthopaedic scores and it has been pivotal in the measurement of arthroplasty outcomes. Similar to several other orthopaedic scores, the WOMAC is valid and reliable, however, these scores may suffer from high ceiling effects [29, 30] and may therefore overlook relevant changes in post-operative patient function over time, especially among patients with better outcomes. Therefore, the development of more sensitive tools for assessing arthroplasty outcomes is of growing interest [31, 32]. Moreover, as the number of younger patients with increased expectations is growing, instruments with which to measure outcomes must adapt to better distinguish between good and excellent results. The original FJS-12 paper [13] as well as several other validations studies [17, 18, 27] found no relevant ceiling effects. In our study of 120 patients, acceptably low ceiling and floor effects were found. Our ceiling effects at 12-months follow-up (12%) were similar to those found in the original FJS-12 paper (9.2%). The floor effect was 0% at 12 months, which was also similar to Behrend's 3.3% [13].

Based solely on the maximum score (100), the FJS-12 showed a higher ceiling effect than WOMAC. However, the plotted scores reveal a more complex view of the ceiling effect. Nearly twice as many WOMAC scores were contained within the ≥90% of the maximum score; hence, FJS-12 was actually more effective at stratifying the scores in the upper ranges. We would recommend that future papers seeking to measure the ceiling and floor effects of orthopaedic questionnaires examine the effects graphically as a ceiling or a floor effect may be present even when the maximum score is not reached.

In one recent publication, the FJS-12 was used in a retrospective cohort study of Italian patients who had undergone unicompartmental knee arthroplasty (UKA) [33]. Although our TKA patients cannot be directly compared to the UKA patients in Zambianchi et al.'s study, the present study's pre-operative and post-operative results were similar to their results; however, their pre-operative scores were on average slightly lower and their post-operative scores were on average slightly greater than our scores [34]. Such differences in baseline scores may also reflect different indications for surgery or differences in health status in the general populations.

In the two other studies that included pre-operative data [35, 36], our data (24.5 ± 16.6) was more similar to the data of the French population (24 ± 16) than the British population (11.5 ± 11.6). Likewise, considering the one-year follow-up data, our results (73.1 ± 23.4) were very similar to the French study's results (70 ± 27) and quite different from the British results (45.7 ± 31.2). Cross-cultural variability may have affected the disparities among the results [37].

Our prospective study included 120 TKA patients. The principal limitation in our study was our high drop-out rate. Our study is strengthened by the decision to include only TKA patients with primary knee osteoarthritis to create a more homogenous sample. Furthermore, none of the patient's knees had been previously operated for any other pathology. Finally, this is one of the first studies to evaluate the FJS-12 from pre-operative presentation through post-operative follow-ups [35, 36]. Although the FJS-12 was developed as a post-operative measurement tool, if future studies also collected pre-operative data, it would provide reference point to which post-operative outcomes could be compared, providing a more complete understanding of the success of joint arthroplasty. As low floor effects were found at pre-operative presentation, the FJS-12 may be suitable for use pre-operatively as well as post-operatively.

Conclusion

In summary, the FJS-12 demonstrated good test-retest reliability and was moderately correlated with the WOMAC. For scores within 5 and 10% of the maximum score, the FJS-12 had a lower ceiling effect than the WOMAC. Hence, this Italian version of the FJS-12 questionnaire seems to be an acceptable representation of the original questionnaire, and we would recommend its use in research concerning the outcomes of TKA as well as other joint arthroplasties in Italy. With increasing numbers of TKAs and increasing patient expectations, the capacity of the FJS-12 to distinguish between patients with good or excellent results after TKA could be advantageous.

Appendix

Forgotten Joint Score

Table 4 Questo questionario serve all'equipe dei suoi medici curanti a valutare quanto la sua articolazione o protesi si "*fa sentire*" durante le attività quotidiane. La preghiamo di compilarlo con la massima cura e sincerità possibile

	MAI	Quasi mai	Raramente	Qualche volta	Sempre
1) senti la protesi a letto durante la notte					
2) senti la protesi quando stai seduto su una sedia per più di un'ora?					
3) senti la protesi quando cammini per più di 15 minuti?					
4)senti la protesi quando fai la doccia/bagno?					
5) senti la protesi quando viaggi in macchina?					
6)senti la protesi quando sali le scale?					
7) senti la protesi quando cammini su un terreno sconnesso?					
8) senti la protesi quando ti alzi da seduto?					
9) senti la protesi quando stai in piedi per un lungo periodo di tempo?					
10) senti la protesi quando fai i lavori di casa?					
11) senti la protesi quando fai una camminata?					
12) senti la protesi quando fai sport o hobby?					

Author details
[1]Department of Orthopaedics, Università degli Studi di Milano, Via Festa del Perdono, 7, 20122 Milan, MI, Italy. [2]IRCCS Istituto Ortopedico Galeazzi, Via Riccardo Galeazzi, 4, 20161 Milan, MI, Italy. [3]AMR Advanced Medical Research GmbH, Männedorf, Switzerland. [4]Department of Statistics and Quantitative Methods, Unit of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, U7, stanza 2064, 20126, Milan, Italy.

References
1. Kurtz SM, Ong KL, Lau E, Widmer M, Maravic M, Gómez-Barrena E, et al. International survey of primary and revision total knee replacement. Int Orthop. 2011;35(12):1783–9. https://doi.org/10.1007/s00264-011-1235-5.
2. Romanini E, Decarolis F, Luzi I, Zanoli G, Venosa M, Laricchiuta P, et al. Total knee arthroplasty in Italy: reflections from the last fifteen years and projections for the next thirty. Int Orthop. 2019;43(1):133–8.
3. Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. Clin Orthop Relat Res. 2009;467:2606e12.
4. Sundberg M, Dahl AW, Lidgren L, Robertson O. The Swedish Knee Arthroplasty Register. Annual Report. 2017. https://www.myknee.se/pdf/SVK_2017_Eng_1.0.pdf. Accessed 30 July 2019.
5. Williams DP, Price AJ, Beard DJ, Hadfield SG, Arden NK, Murray DW, et al. The effects of age on patient-reported outcome measures in total knee replacements. Bone Joint J. 2013;95-B(1):38–44.
6. Witjes S, van Geenen RC, Koenraadt KL, van der Hart C, Blankevoort L, Kerkhoffs GM, et al. Expectations of younger patients concerning activities after knee arthroplasty: are we asking the right questions? Qual Life Res. 2017;26(2):403–17.
7. Marx RG, Jones EC, Atwan NC, et al. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. J Bone Joint Surg Am. 2005;87:1999.
8. Dawson J, Beard DJ, McKibbin H, et al. Development of a patient-reported outcome measure of activity and participation the OKS-APQ to supplement the Oxford knee score. Bone Joint J. 2014;96-B:332–8.
9. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988;15(12):1833–40.
10. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the knee society clinical rating system. Clin Orthop Relat Res. 1989;248:13–4.
11. Clement ND, Weir D, Holland JP, Gerrand CH, Deehan DJ. An overview and predictors of achieving the postoperative ceiling effect of the WOMAC ScoreFollowing Total knee Arthroplasty. J Arthroplast. 2019;34(2):273–80.
12. Dinjens RN, Senden R, Heyligers IC, Grimm B. Clinimetric quality of the new 2011 Knee Society score: high validity, low completion rate. Knee. 2014;21(3):647–54.
13. Behrend H, Giesinger K, Giesinger JM, et al. The forgotten joint as the ultimate goal n joint arthroplasty. Validation new patient-reported outcome measure J Arthroplasty. 2012;27:430.

Sansone *et al. BMC Musculoskeletal Disorders*   (2020) 21:23

Page 9 of 9

14. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993;46(12):1417–32.

15. 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(1):34–42.

16. Thienpont E, Opsomer G, Koninckx A, et al. Joint awareness in different types of knee Arthroplasty evaluated with the forgotten joint score. J Arthroplast. 2013;29:48.

17. Baumann F, Ernstberger T, Loibl M, Zeman F, Nerlich M, Tibesku C. Validation of the German Forgotten Joint Score (G-FJS-12) according to the COSMIN checklist: does a reduction in joint awareness indicate clinical improvement after arthroplasty of the knee? Arch Orthop Trauma Surg. 2016;136(2):257–64.

18. Shadid MB, Vinken NS, Marting LN, Wolterbeek N. The Dutch version of the Forgotten Joint Score: test-retesting reliability and validation. Acta Orthop Belg. 2016;82(1):112–8.

19. Cao S, Liu N, Han W, Zi Y, Peng F, Li L, et al. Simplified Chinese version of the forgotten joint score FJS-12 for patients who underwent joint arthroplasty: cross-cultural adaptation and validation. J Orthop Surg Res. 2017;12(1):6.

20. Gandek B. Measurement properties of the Western Ontario and McMaster Universities Osteoarthritis Index: a systematic review. Arthritis Care Res Hoboken. 2015;67(2):216–29.

21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. COSMIN checklist manual 2012. http://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf. Accessed 1 June 2019

22. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003;41(5):582–92.

23. Jette AM, Haley SM, Tao W, Ni P, Moed R, Meyers D, Zurek M. Prospective evaluation of the AM-PAC-CAT in outpatient rehabilitation settings. Phys Ther. 2007;87(4):385–98.

24. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4(4):293–307.

25. di Laura FG, Zaffagnini S, Filardo G, Romandini I, Fusco A, Candrian C. Total Knee Arthroplasty in Patients With Knee Osteoarthritis: Effects on Proprioception. A Systematic Review and Best Evidence Synthesis. J Arthroplasty. 2019;S0883–5403(19):30556.

26. Thomsen MG, Latifi R, Kallemose T, Barfod KW, Husted H, Troelsen A. Good validity and reliability of the forgotten joint score in evaluating the outcome of total knee arthroplasty. Acta Orthop. 2016;87(3):280–5.

27. Thompson SM, Salmon LJ, Webb JM, Pinczewski LA, Roe JP. Construct validity and test re-test reliability of the forgotten joint score. J Arthroplast. 2015;30(11):1902–5.

28. Ingelsrud LH, Roos EM, Terluin B, Gromov K, Husted H, Troelsen A. Minimal important change values for the Oxford knee score and the forgotten joint score at 1 year after total knee replacement. Acta Orthop. 2018 Oct;89(5):541–7.

29. Clement ND, Weir D, Holland JP, Gerrand CH, Deehan DJ. An Overview and Predictors of Achieving the Postoperative Ceiling Effect of the WOMAC Score Following Total Knee Arthroplasty. J Arthroplasty. 2019;34(2):273–80.

30. Marx RG, Jones EC, Atwan NC, Closkey RF, Salvati EA, Sculco TP. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. J Bone Joint Surg Am. 2005;87(9):1999–2005.

31. Culliton SE, Bryant DM, MacDonald SJ, Hibbert KM, Chesworth BM. Validity and internal consistency of the new knee society knee scoring system. Clin Orthop Relat Res. 2018;476(1):77–84.

32. 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.

33. Zambianchi F, Franceschi G, Rivi E, Banchelli F, Marcovigi A, Khabbazè C, Catani F. Clinical results and short-term survivorship of robotic-arm-assisted medial and lateral unicompartmental knee arthroplasty. Knee Surg Sports Traumatol Arthrosc. 2019. https://doi.org/10.1007/s00167-019-05566-4. [Epub ahead of print]

34. Lyons MC, MacDonald SJ, Somerville LE, Naudie DD, McCalden RW. Unicompartmental versus total knee arthroplasty database analysis: is there a winner? Clin Orthop Relat Res. 2012;470(1):84–90. https://doi.org/10.1007/s11999-011-2144-z.

35. Hamilton DF, Loth FL, Giesinger JM, Giesinger K, MacDonald DJ, Patton JT, Simpson AH, Howie CR. Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population. Bone Joint J. 2017;99-B(2):218–24.

36. Thienpont E, Vanden Berghe A, Schwab PE, Forthomme JP, Cornu O. Joint awareness in osteoarthritis of the hip and knee evaluated with the 'forgotten joint' score before and after joint replacement. Knee Surg Sports Traumatol Arthrosc. 2016;24:3346–51.

37. Giesinger JM, Giesinger K, Federico B, Howie CD, Hamilton DF. Differences in case mix and outcomes between Swiss and Scottish total knee arthroplasty patients. Knee Surg Sports Traumatol Arthrosc. 2019. https://doi.org/10.1007/s00167-019-05597-x.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.