



Rasch analysis of the Forgotten Joint Score in patients undergoing knee arthroplasty

Didier D. Niama Natta¹ · Emmanuel Thienpont^{2,3} · Alexandre Bredin² · Gwenael Salaun² · Christine Detrembleur² 

Received: 25 May 2018 / Accepted: 10 August 2018 / Published online: 21 August 2018
© European Society of Sports Traumatology, Knee Surgery, Arthroscopy (ESSKA) 2018

Abstract

Purpose The purpose of this study was to test the hypothesis that the “Forgotten Joint Score” (FJS-12) is a unidimensional interval-level scale. Unidimensionality refers to measuring a single attribute, i.e., the single ability to forget the arthroplasty. If this property is not verified, the interpretation of the score can be confusing. Unidimensionality is an essential prerequisite of construct validity and required if FJS-12 response data are to be validly summated into a single score. Interval-level dimension is an essential prerequisite of the parametric statistics. Rasch analysis was used to test our study hypothesis.

Methods The FJS-12 questionnaire was validated in 248 unilateral knee arthroplasty patients. Successive analyses were used to select items with good psychometric qualities to constitute the new “FJS”. The external validity was assessed with the KUJALA questionnaire.

Results Quantity of relevant items was greater than 50%. Of the 12 original items, nine showed disturbed thresholds, indicating that patients were unable to discriminate among the five levels for these items. The data set was reanalyzed using a four-level scale. The new analysis indicated that the internal consistency was good ($r=0.84$). Three items did not fit with the model and they were removed. The nine items of the final scale defined a unidimensional and linear measure of the forgotten joint, and showed a continuous progression in their difficulty. The perception of difficulty was group-independent. The correlation coefficient was moderate between FJS and KUJALA score ($r=0.4$).

Conclusions This new and items reduced FJS can be used in clinical practice with good psychometric qualities. It provides a reliable tool to follow up patient’s evolution and document changes related to knee arthroplasty. This valid FJS is needed in evaluating patients’ assessment, one indicator of quality of care.

Level of evidence III—Therapeutic.

Keywords Rash model · FJS-12 score · Awareness joint · Total knee arthroplasty

Introduction

Total knee arthroplasty (TKA) is the most common treatment for end-stage osteoarthritis (OA) of the knee. The main goals of knee arthroplasty are to effectively relieve pain and

regain joint function, thus providing substantial improvement in quality of life. After surgery, approximately 30% of patients state that their expectations from joint replacement surgery were not fully achieved. One of the key expectations,

✉ Christine Detrembleur
christine.detrembleur@uclouvain.be

Didier D. Niama Natta
digitalesfr@yahoo.fr

Emmanuel Thienpont
emmanuel.thienpont@uclouvain.be

Alexandre Bredin
Alexandre.bredin@student.uclouvain.be

Gwenael Salaun
Gwenael.salaun@uclouvain.be

¹ Physical Medicine and Rehabilitation Department Cotonou, National University Hospital of Cotonou, Cotonou, Benin

² Neuro-musculoskeletal Lab (NMSK), Secteur des Sciences de la Santé, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Avenue Mounier 53, B1.53.07, 1200 Brussels, Belgium

³ Cliniques Universitaires Saint-Luc, Service d’Orthopédie et de Traumatologie de l’appareil Locomoteur, Avenue Hippocrate 10, 1200 Brussels, Belgium

and therefore goals of this surgery, is the ability of patients to forget their artificial joint during everyday life [2].

The “Forgotten Joint” Score (FJS-12) assesses awareness of the hip and knee joints during various activities of daily living (ADL) following joint replacement [6]. The FJS-12, an ordinal scale, uses a five-point Likert response format, consisting of 12 equally weighted questions with the raw score transformed to a scale of 0–100 points. High scores indicate a good outcome (i.e., being able to forget about the affected joint in daily life) [3].

The FJS-12 is assumed to be a unidimensional interval scale, but this assumption has not been tested. As highlighted by Van der Velde et al. [22], validating the hypothesis that the FJS-12 is a unidimensional scale is important, because unidimensionality is an essential prerequisite of construct validity. Interval-level dimension is also an essential prerequisite of the parametric statistics that are used to assess outcomes and compare treatment responses across groups [22]. Interval-level scaling allows the straightforward interpretation of changes in FJS-12 score, such that, for example, a change in score from 70 to 80 is equivalent to a change from 10 to 20. Thus, a measure of “joint awareness” was obtained that is free from bias and satisfies the guidelines for building interval-scale data [22].

Although the measurement properties of the FJS-12 have been extensively examined using classic test theory methods, the assumption of interval-level measurements and unidimensionality of joint awareness has not been verified. The objective of this study was to use modern psychometric approaches, specifically the Rasch analysis, to test if the FJS-12 is a unidimensional interval scale.

Materials and methods

From the medical database, 248 patient files were selected. Patients had received unilateral knee arthroplasty at the Cliniques Universitaires Saint-Luc in Brussels between January 2012 and March 2015. All patients had been operated by a single, experienced surgeon. The following parameters were noted: age, sex, body mass index (BMI), laterality, and type of implant [total knee arthroplasty (TKA), unicompartmental arthroplasty (UKA) or patellofemoral arthroplasty (PFA)], FJS-12 scores, and Kujala Anterior Knee Pain Scale (KUJALA)]. Only patients with complete records were included. The scores were obtained pre-surgery and at 3, 6, 9, 12, and 18 month intervals post-surgery. The KUJALA questionnaire [11] consists of 13 questions: three of which refer to pain and physical alterations, eight to possible limitations in functional capacity, and two describe the ability to participate in sports.

The protocol of this retrospective study was approved by the local ethics committee of our Faculty of Medicine

of the Université Catholique de Louvain (Reference No. B403201523492) and registered on clinical trial.gov (NCT02355301).

Statistical analysis

Rasch analysis

The Rasch analysis is a process of testing whether data from a scale, such as the FJS-12, satisfy the rules for constructing interval-scale measurement [17]. The Rasch Unidimensional Measurement Models program 2030 (RUMM2030) was used to estimate: (1) patient performance level; (2) item difficulty; (3) item thresholds (i.e., the positions on the scale at which two adjacent categories are likely to be observed equally) [23]. Successive analyses were used to select items with good psychometric qualities to constitute the new FJS-12 scale. The methodology was described in detail in [17, 22, 23].

Relevant items

An item is considered relevant when it forms part of the daily reality of most patients. Items with a response rate < 50% are removed [17].

Ordered rating scale

Patients’ perceptions are reported on a five-level scale: 1—never, 2—nearly never, 3—rarely, 4—sometimes, and 5—frequently. Good discrimination of these different categories is highlighted by increasing levels of performance as represented by thresholds occurring in the expected order. When patients do not discriminate between categories, the thresholds are reversed. In this case, two categories can be merged [24].

Item and patient targeting

Item–patient targeting was also analyzed by comparing histogram frequencies (top and lower panels of Fig. 1) to detect potential gaps in the scale [16]. The average difficulty of the items is conventionally set at 0 logit. Items reported as ‘easy’ or ‘impossible’ by all the subjects are deleted.

To fit the Rasch model

When observed responses are equal to or do not differ greatly from the expected responses provided by the model, the data are said to fit the Rasch model. In general, Chi-square fit statistics need to be non-significant. Residual fit statistics are expected to be within a given range ± 2.5 for individual items, and with a mean fit residual value close to

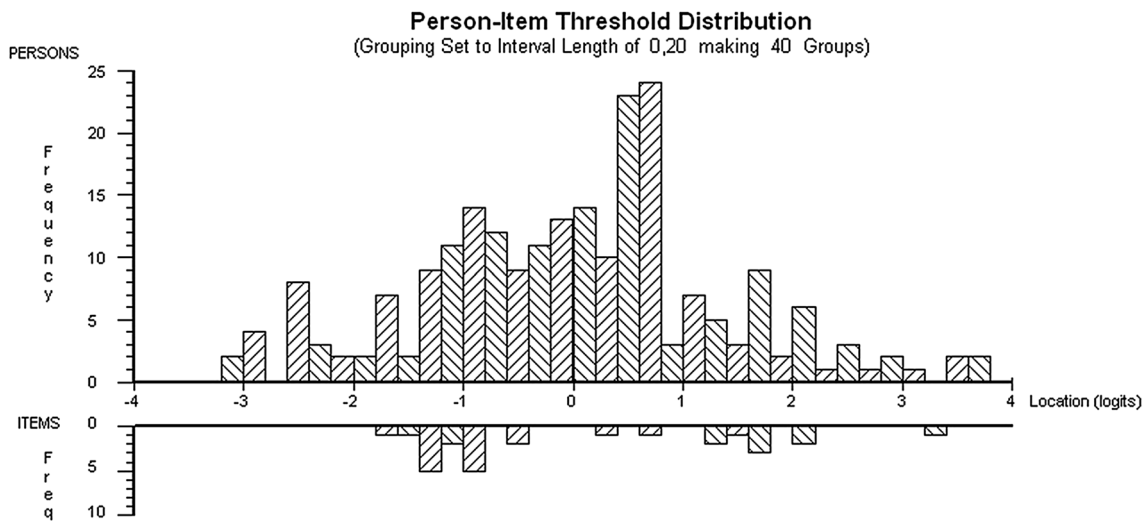


Fig. 1 Rasch-calibrated metric scale of FJS distribution of the patients ($n=248$) and items along the Rasch-calibrated metric scale. Upper panel shows the location of the patients

0.0 and an SD approaching 1.0 (usually < 1.4) for summary statistics [22].

Local independency

Local dependency is a phenomenon that occurs because of a high correlation between some items on the same scale. If there are residual correlations greater than 0.3 between two items, the item with the worst psychometric qualities is removed [13].

Unidimensional scale

Principal component analysis is used to establish two item subsets, estimates of which are compared using a paired t test. The unidimensional assumption is considered as met when the percentage of this test is outside the range -1.96 to 1.96 , but does not exceed 5% [16].

Differential item functioning (DIF)

Patients with the same Forgotten Joint Score, but different demographics or clinical characteristics, must have the same probability of choosing a given response for any particular item. If this is not the case, the item is biased or presents “differential functioning”. Invariance of the item difficulty hierarchy was tested for age (≤ 67.5 years old vs > 67.5 years old), sex, type of implant (TKA vs UKA/PFA), and delay post-surgery (≤ 3 months vs > 3 months) [9]. Items with DIF for age, sex, implant, or delay were removed from the scale.

Scale reliability

The degree of precision achieved in the measurement was reported using a Person Separation Index. A Person Separation Index > 0.8 is considered as good [5].

External validity

External validity was investigated by measuring the degree of association between FJS-12 and KUJALA scales. We used the Andresen’s guidelines to interpret the correlation coefficient [1].

Results

Population

Table 1 describes the main characteristics of study patients.

Table 1 Demographic data of patients with unilateral knee arthroplasty ($n=248$)

	Mean \pm SD	Range
Age (years)	67.5 \pm 10.9	28–92
Sex (male/female)	97/151	
Body mass index (%)	28.9 \pm 5.7	14.5–47.1
Operated knee (right/left)	137/111	
Type of implant (total/uni)	170/78	

Rasch analysis results

The rating scale analysis has been selected after the Chi-square likelihood ratio test has been completed (Chi-square = 19.23; $p = 0.57$).

Relevant items and ordered rating scale

For each item, the response rate was greater than or equal to 50%. No item was removed. Of the 12 original items, nine showed disordered thresholds, indicating that many patients were unable to discriminate among the five levels for these items. As the subjects were unable to discriminate between “nearly never” and “rarely”, these categories were merged into one category, “rarely”, and the entire data set was reanalyzed using a four-level scale: (0) never, (1) rarely, (2) sometimes, and (3) frequently.

Item-person targeting

Figure 1 shows good item-person targeting for FJS questionnaire. The FJS questionnaire showed no floor or ceiling effect.

Fit statistics and FJS properties

The first analysis showed that the item–trait interaction had good fit (Chi-square = 34.05; $p = 0.08$), one item did not fit the model, and three item pairs had correlations better than 0.3. Successive Rasch analyses were performed to construct the final FJS questionnaire. Three items (awareness when you are walking for more than 15 min, which is item 3; awareness when you are travelling in a car, which is item 5; awareness taking a walk/hiking, which is item 11) were deleted.

The final version of FJS thus included nine items on a four-level scale. The item–trait interaction indicated that all

nine items contributed to the definition of a unidimensional scale ($\chi^2 = 13.23$; $p = 0.77$). Table 2 reports items calibration and fit indices.

The items were classed in decreasing difficulty order (from 1.38 to -0.61 logits), with higher logit values representing more difficult items. The FJS is expressed in logits, a linear unit defined as the natural logarithm of the odds of successful achievement by a patient for any item. The item difficulty was estimated with a mean of 0 ± 0.59 logit and the mean residual standard of -0.33 ± 1.93 . The location of patients 175 ranged from -3.06 to 3.74 logits. The overall mean patient location on the scale was -0.001 logits with a standard deviation of 1.37 logits and residual standard of -0.47 ± 1.42 .

The nine items of the final scale defined a unidimensional and linear measure of forgotten joint and showed a continuous progression in their difficulty.

The new FJS scale is described in Table 3. Table 4 underlines the relationship between the total raw scores, and the scores expressed in logits and in percent.

Local independency and unidimensionality

Residual correlations between all items were weak ($r \leq 0.3$). None of the nine items was deleted.

All items defined a unidimensional scale of joint awareness, because p values did not show a significant difference between observed and expected scores. The percentage of paired t test results outside -1.96 to 1.96 was 4.05.

Differential item functioning (DIF)

Based on four criteria (age, sex, implant, and postoperative delay), patients were allocated into two groups. The perceptions of difficulty of each item by these two groups were plotted against one another (Fig. 2). As the 95% CI included the majority of items, the perception of difficulty

Table 2 Individual item fit for FJS-12 after knee arthroplasty

	Difficulty (logit)	SE (logit)	Residual (z)	Fit (χ^2)	p value
Awareness standing for long periods of time?	-0.61	0.10	-1.41	1.32	n.s
Awareness when standing up from a low-sitting position?	-0.46	0.09	-0.75	0.9	n.s
Awareness doing housework or gardening?	-0.28	0.10	-2.16	1.55	n.s
Awareness doing your favourite sport?	-0.21	0.14	-1.39	0.24	n.s
Awareness sitting on a chair for more than 1 h?	-0.15	0.09	0.81	1.56	n.s
Awareness walking on uneven ground?	-0.13	0.09	-2.10	1.84	n.s
Awareness climbing stairs?	0.22	0.09	-0.5	0.79	n.s
Awareness in bed at night?	0.23	0.09	4.12	2.79	n.s
Awareness taking a bath/shower?	1.38	0.09	0.48	2.23	n.s

The items are presented in order of increasing difficulty from the least to the most difficult
SE standard error, n.s non-significant

Table 3 Modified FJS with new scores

	Never (1)	Rarely (2)	Sometimes (3)	Frequently (4)
Awareness standing for long periods of time?				
Awareness when standing up from a low-sitting position?				
Awareness doing housework or gardening?				
Awareness doing your favourite sport?				
Awareness sitting on a chair for more than 1 h?				
Awareness walking on uneven ground?				
Awareness climbing stairs?				
Awareness in bed at night?				
Awareness taking a bath/shower?				

Table 4 Correspondence table between the total ordinal score and score in logit and in percent

Ordinal score	Score in logit	Score in %
0	-3.79	0.00
1	3.06	8.65
2	-2.57	14.45
3	-2.23	18.48
4	-1.97	21.56
5	-1.74	24.29
6	-1.54	26.66
7	-1.36	28.79
8	-1.18	30.92
9	-1.02	32.82
10	-0.85	34.83
11	-0.68	36.85
12	-0.51	38.86
13	-0.33	41.00
14	-0.15	43.13
15	0.05	45.50
16	0.26	47.99
17	0.49	50.71
18	0.72	53.44
19	0.98	56.52
20	1.25	59.72
21	1.54	63.15
22	1.85	66.82
23	2.19	70.85
24	2.59	75.59
25	3.07	81.28
26	3.74	89.22
27	4.65	100.00

was group-independent. For the age criteria, the intraclass correlation coefficient (ICC) was 0.97 with $p < 0.001$; for gender, ICC was 0.87, $p < 0.001$; for type of implant, ICC was 0.93, $p < 0.001$; for the delay, ICC was 0.93, $p < 0.001$.

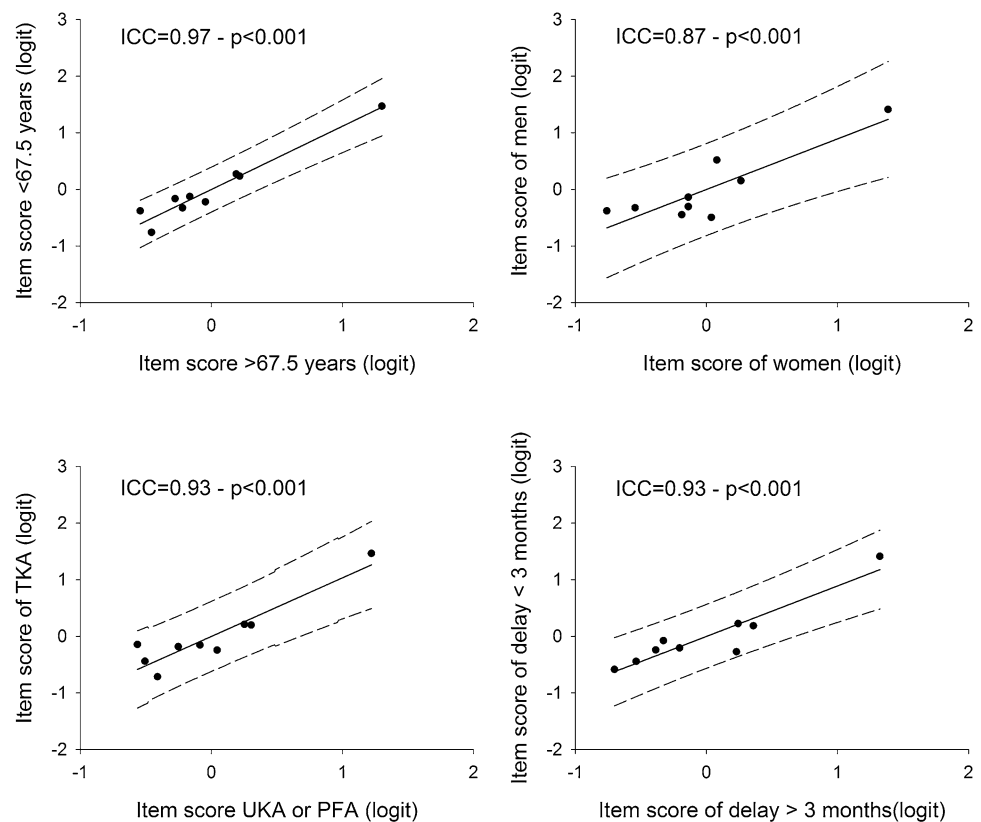
Scale reliability and external validity

The reliability index based on the person separation index was 0.84. The correlation coefficient was moderate between FJS-12 and KUJALA ($r = 0.4$; $p < 0.001$).

Discussion

The most important finding of this study was that the validated FJS scale allows the practice of calculating changes in scores and other parametric statistics. It responds to the key criteria for scale assessment defined in the guidelines for psychometric standards for rating scales [17]. This new FJS provides a reliable tool to follow patient's evolution and document changes related to knee arthroplasty. This valid FJS is required to evaluate patients' assessment, as one indicator of care quality. The "Forgotten Joint Score" was validated as a new questionnaire constructed with nine items with four response categories using a Rasch analysis (Tables 3, 4). The internal consistency of this version was good. This questionnaire enables three significantly different levels of joint awareness to be distinguished, which is greater than the two minimum levels required for a measurement tool [15]. The internal consistency of this study is less than that of the original version, which is 0.95 [3], the Turkish version, which is 0.92 [10], and the English version, which is 0.97 [7]. This difference can be explained by the fact that, in these studies, the Cronbach's alpha was calculated on raw scores [3, 7, 10] and on 14 items in the initial FJS scale [3]. In Turkish version [10], the 12th question (based on the sport practice) was not directly suitable to Turkish population for cultural reasons. The population life style was not adapted to sport activities before or after the arthroplasty surgery and all patients responded "never". Then, Cronbach's alpha was calculated on 11 questions. Nevertheless, it should be noted that a robust statistical method was used for calculations, whereas the other studies used a classical test theory method.

Fig. 2 Item difficulty hierarchy plots of four different groups. Item difficulty hierarchy estimated by age, gender, type of implant, and delay after surgery. Most difficult items are plotted in the upper right part of the figure. Control lines (solid lines) indicate the 95% confidence intervals of the ideal invariance. All items (dots) lie within the control line, indicating that both groups estimated the item difficulty consistently



The item-individual targeting for the questionnaire was excellent, meaning that it was not too easy or too difficult for the population. For good targeting, the average difficulty for subjects must be close to 0 [17]. Thus, this version of the forgotten joint questionnaire has no ceiling or floor effect compared to the original version, which has a ceiling effect of 9.2% [3], and to the Danish version, which has a ceiling effect of 16% [21].

The validated questionnaire displayed no evidence of DIF within the following groups—sex, age, type of implant, and postoperative delay—indicating that the questionnaire works similar for these groups [15]. Indeed, Behrend et al. [3] and Thienpont et al. [18] reported that age did not significantly affect the FJS. However, Hiyama et al. showed that knee pain and quadriceps strength did affect joint awareness [8]. Thienpont et al. [18] also reported that the type of implant (UKA vs TKA) did not affect the outcome of the FJS score, but did observe a higher score for fixed-bearing TKA than for mobile-bearing TKA [19]. In contrast, Zuiderbaan et al. [25] showed a significant difference between total and unicompartamental FJS-12 outcome scores not observed by the previous author [18]. This difference may be related to the way that the statistical analysis was performed. The latter authors first summed the ordinal scores. Then, the total scores were analyzed as a continuous variable with a parametric statistic, which may explain the observed differences.

The Rasch analysis confirmed that the questionnaire does not measure any latent variable other than joint awareness. This observation was also made in the validation of the English version using classical testing theory [7]. The questionnaire cannot be used to estimate pain, functional limitations, or other symptoms. Analysis of the external validity of the FJS showed that it has moderate correlation with KUJALA. Kujala documents response to six activities thought to be associated specifically with anterior knee pain syndrome. The FJS is further questionnaire assessing contextual factors.

This new calibrated scale would be interesting to use, for example, in the study of Li et al. [12]. In fact, this revised FJS has good psychometric properties which are essential to orientate treatments and assess their efficacy. Li et al. showed that female, younger patients and patients with higher BMIs had significantly worse pre-operative joint awareness. However, the authors used parametric tests, a multiple regression analysis to identify possible prognostic factors. These results could be revised with our revised scale adapted to this parametric statistic test.

The main limitation of this study was that the FJS should be tested in larger and more diverse populations. For example, as highlighted by Thienpont et al. [20], total hip arthroplasty is cited more often as a forgotten joint than TKA. Other disorders should be evaluated with this new score, such as joint awareness in patients who have undergone

anterior cruciate ligament reconstruction [4]. Another limitation is that the results were only validated in unilateral knee arthroplasty. Nielsen et al. showed in 99 patients who had undergone bilateral simultaneous TKA that knee awareness did not differ significantly between the “best” and the “worst” knee. The FJS was not able to detect a significant difference between the knees [14].

This tool is clinically relevant to assess contextual factors according to the International Classification of Functioning, Disability, and Health. Contextual factors are very rarely assessed in day-by-day clinical work. This tool is now a unidimensional interval scale allowing to follow patient with a single score expressed in percentages.

Conclusion

Rasch analysis of the FJS demonstrated the unidimensionality of its items, whose psychometric properties were suitable in terms of validity and internal consistency. FJS may be used in research and clinical practice as a tool for measuring the effectiveness of TKA surgery using the conversion table of raw scores into interval scores.

Author contributions All authors worked collectively to develop the protocols and methods described in this paper. DN, ET, and CD were principal investigators responsible for the fieldwork, the statistical analysis, and the writing of the manuscript. All authors contributed to the interpretation of results. All authors read and approved the final manuscript.

Funding This project is not supported by a research grant or funding source.

Compliance with ethical standards

Conflict of interest The authors declare that they have no proprietary, financial, professional, or other personal competing interests of any nature or kind.

Ethics approval This study was conducted according to the declaration of Helsinki and had ethical approval from the Comité d’Ethique Hospitalo-Facultaire de l’Université Catholique de Louvain (B403201523492).

References

- Andresen EM (2000) Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 81:S15–S20
- Baker PN, van der Meulen JH, Lewsey J et al (2007) The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. *J Bone Joint Surg Br* 89:893
- Behrend H, Giesinger K, Giesinger JM, Kuster MS (2012) The “forgotten joint” as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty* 27(3):430–436
- Behrend H, Zdravkovic V, Giesinger JM, 292 Giesinger K (2017) Joint awareness after ACL reconstruction: patient-reported outcomes measured with the Forgotten Joint Score-12. *Knee Surg Sports Traumatol Arthrosc* 25:1454–1460
- Bond TG, Fox CM (2007) Applying the Rasch model: fundamental measurement in the human sciences, 2nd edn. Lawrence Erlbaum Associates, Mahwah
- Giesinger JM, Kesterke N, Hamilton DF, Holzner B, Jost B, Giesinger K (2015) Development of item list to assess the forgotten joint concept in shoulder patients. *BMC Musculoskelet Disord* 16:67
- Hamilton DF, Loth FL, Giesinger JM, MacDonald DJ, Patton JT, Simpson AH, Howie CR (2017) Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population. *J Bone Joint Br* 99:218–224
- Hiyama Y, Wada O, Nakakita S, Mizuno K (2016) Joint awareness after total knee arthroplasty is affected by pain and quadriceps strength. *Orthop Traumatol Surg Res* 102(4):435–439
- Holland PH, Wainer H (1993) Differential item functioning. Lawrence Erlbaum, Hillsdale.
- Kınıklı G, Güney Deniz H, Karahan S, Yüksel E, Kalkan S, Dönder Kara D, Önal S, Sevinç C, Çağlar Ö, Atilla B, Yüksel İ (2017) Validity and reliability of Turkish version of the Forgotten Joint Score-12. *J Exerc Ther Rehabil* 4:18–25
- Kujala UM, Jaakkola LH, Koskinen SK, Taimela S, Hurme M, Nelimarkka O (1993) Scoring of patellofemoral disorders. *Arthroscopy* 9(2):159–163
- Li D, Troelsen A, Ingelsrud L, Husted H, Gromov K (2017) Females, younger patients and patients with high BMI have the highest pre-operative knee awareness measured using the Forgotten Joint Score. *Knee Surg Sports Traumatol Arthrosc* 16:4446–4445
- Linacre JM (1998) Structure in Rasch r 317 residual: why principal component analysis? *Rasch Meas Trans* 12(2):636
- Nielsen KA, Thomsen MG, Latifi R, Kallelose T, Husted H, Troelsen A (2016) Does post-operative knee awareness differ between knees in bilateral simultaneous total knee arthroplasty? Predictors of high or low knee awareness. *Knee Surg Sports Traumatol Arthrosc* 24:3352–3358
- Pallant JF, Tennant A (2007) An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol* 46:1–18
- Smith EV Jr (2002) Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 3(2):205–231
- Tennant A, Conaghan PG (2007) The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 57(8):1358–1362
- Thienpont E, Opsomer G, Koninckx A, 341 Houssiau F (2014) Joint awareness in different types of knee arthroplasty evaluated with the Forgotten Joint score. *J Arthroplasty* 29(1):48–51
- Thienpont E, Zorman D (2016) Higher forgotten joint score for fixed-bearing than for mobile-bearing total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 24:2641–2645
- Thienpont E, Vanden Berghe A, Schwab PE, Forthomme JP, Cornu O (2016) 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* 24:3346–3351
- Thomsen MG, Latifi R, Kallelose T, Barfod KW, Husted H, Troelsen A (2016) Good validity and reliability of forgotten

- joint score in evaluating the outcome of total knee arthroplasty. *Acta Orthop* 87(3):280–285
22. Van der velde G, beaton D, Hogg-Johnston S, Hurwitz E, Tennant A (2009) Rasch analysis provides new insights into the measurement properties of the neck disability index. *Arthritis Rheum* 61(4):544–551
 23. Vandervelde L, Van den Bergh PYK, Goemans N, Thonnard JL (2007) ACTIVLIM: a Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscul Disord* 17:459–469
 24. Zhu W, Updyke WF, Lewandowski C (1997) Post-hoc Rasch analysis of optimal categorization of an ordered-response scale. *J Outcome Meas* 1(4):286–304
 25. Zuiderbaan HA, van der List JP, Khamaisy S, Nawabi DH, Thein R, Ishmael C, Paul S, Pearle AD (2017) Unicompartmental knee arthroplasty versus total knee arthroplasty: which type of artificial joint do patients forget? *Knee Surg Sports Traumatol Arthrosc* 25:681–686