

# WHAT IS THE BEST QUESTIONNAIRE FOR MONITORING THE PHYSICAL CHARACTERISTICS OF PATIENTS WITH KNEE OSTEOARTHRITIS IN THE BRAZILIAN POPULATION?

Leonardo Metsavaht<sup>1</sup>, Gustavo Leporace<sup>2</sup>, Maria Matilde de Mello Sposito<sup>3</sup>, Marcelo Riberto<sup>4</sup>, Luiz Alberto Batista<sup>5</sup>

## ABSTRACT

**Objective:** To measure the validity and reliability of the WOMAC, IKDC and Lysholm questionnaires among patients with knee osteoarthritis and determine the influence of age on their scores. **Method:** Fifty-seven patients diagnosed with primary osteoarthritis of the knee answered the SF-36, WOMAC, IKDC and Lysholm questionnaires. Validity was tested by measuring the correlation (Pearson's correlation coefficient, "r") between the questionnaires. Internal consistency was measured using Cronbach's  $\alpha$  and concordance was measured through the Altman-Bland graphical representations and survival concordance. To determine the influence of age on the results, we correlated this variable with the scores from the three knee questionnaires using Pearson's coefficient of determination ("r<sup>2</sup>"). **Results:** IKDC (0.62) and WOMAC (0.642) showed moderate to strong correlations in relation to the summarized physical capacity score of the SF-36, while Lysholm

showed moderate correlations (0.555). The Cronbach  $\alpha$  values for IKDC, WOMAC and Lysholm were 0.811, 0.959 and 0.734, respectively. Despite the strong correlations between IKDC and WOMAC (0.843), Lysholm and WOMAC (0.759) and IKDC and Lysholm (0.858), the Bland-Altman graphical representations and survival concordance showed that the concordance between the three questionnaires was low. IKDC, Lysholm and WOMAC showed coefficients of determination (r<sup>2</sup>) with age of 0.004, 0.010 and 0.043, respectively. **Conclusion:** Age was not found to be a limiting factor for the use of any of the questionnaires applied in this study. The concordance tests and correlations with the physical components of SF-36 suggest that WOMAC is more appropriate for assessing physical capacities and limitations relating to physical traits, while IKDC seems more appropriate for assessing the functional limitations relating to pain.

**Keywords** – Osteoarthritis; Questionnaires; Psychometrics

## INTRODUCTION

Osteoarthritis (OA) is a disease of inflammatory and degenerative nature that causes destruction of the joint cartilage and can lead to acute and chronic painful conditions and joint deformities<sup>(1)</sup>. Its prevalence ranges from around 4% to 30%, and its incidence is especially high among the elderly population, although this disease is not an inevitable consequence of advancing age<sup>(2)</sup>. The knee is the

joint most affected by OA, and its functionality status is strongly associated with changes to individuals' activities of daily living and autonomy<sup>(3)</sup>. The World Health Organization (WHO) classifies OA in the knee as the fourth biggest cause of incapacity among women and the eighth among men<sup>(4)</sup>.

Different clinical examinations can be used to delineate the diagnosis of OA, such as radiography<sup>(2)</sup>, magnetic resonance<sup>(5)</sup> and, recently, bioimpedance<sup>(6)</sup>.

1 – MSc in Medicine from DOT/SOT/Federal University of Rio de Janeiro. CSO of the Brazilian Institute of Health Technology (IBTS).

2 – MSc student on the Biomedical Engineering Program, COPPE, Federal University of Rio de Janeiro (UFRJ); Researcher in the Biomechanics and Motor Behavior Laboratory, State University of Rio de Janeiro (UERJ).

3 – PhD in Medicine. Hospital Lucy Montoro, Institute of Physical Medicine and Rehabilitation, University of São Paulo (USP), São Paulo, Brazil.

4 – PhD in Medicine. Professor in the Ribeirão Preto Medical School, University of São Paulo (USP), São Paulo, Brazil.

5 – PhD in Sports Sciences. Coordinator of the Biomechanics Laboratory and Adjunct Professor of the Physical Education and Sports Course, State University of Rio de Janeiro (UERJ).

Work performed at the Brazilian Institute of Health Technology (IBTS), Rio de Janeiro, Brazil.

Correspondence: Instituto Brasil de Tecnologias da Saúde - Rua Visconde de Pirajá, 407/905, Ipanema - 22410-003 - Rio de Janeiro, RJ, Brasil. E-mail: leonardo@brasilsaude.org

Work received for publication: March 22, 2010; accepted for publication: July 20, 2010.

The authors declare that there was no conflict of interest in conducting this work

One characteristic common to these examinations is the fact that they all require interpretation and evaluation by clinicians, both to establish the diagnosis and to estimate the prognosis for the impairment. This introduces an important measurement bias into the process<sup>(7)</sup>.

Over the last two decades, patients' perceptions of their state of health have been given value as a fundamental variable in making trustworthy clinical assessments and, consequently, in the therapeutic strategy<sup>(8)</sup>. Information on patients' perceptions of their state of health is usually gathered through applying questionnaires. With regard to OA, the ones most used are the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and the Lequesne Index<sup>(9,10)</sup>, since these were developed specifically for use in assessing this disease. However, other instruments such as the subjective assessment of the "International Knee Documentation Committee" and the Lysholm scale also provide satisfactory results when applied to this population<sup>(11)</sup>. The level of scientific authenticity of these instruments may vary depending on the sociocultural contexts and specific characteristics of populations, thus influencing the quality of the results from interpreting the questions<sup>(12)</sup>. It is therefore necessary to identify the degree of adequacy of these instruments for assessing issues relating to pain, physical limitations and functionality, among Brazilian patients affected by OA. Furthermore, because this is a disease that predominantly affects one age group, an estimate of the influence of age on the final score may signify a better fit with reality for the results.

The aim of this study was to measure the validity and reliability of the WOMAC, IKDC and Lysholm questionnaires among patients with OA in the knee, and to determine the influence of age on the scores from these questionnaires.

## METHODS

The patients agreed to participate in the study by signing a commitment statement, thereby authorizing their participation. This study was approved by the institution's Ethics Committee.

Fifty-seven patients with a diagnosis of primary OA of the knee answered the Brazilian versions of the SF-36, WOMAC, Lysholm and IKDC questionnaires. Forty-one patients were female (71.9%) and sixteen

were male (28.1%). The mean age of the sample was 61.7 years, with a range from 35 to 84 years.

According to de Vet et al<sup>(13)</sup>, validity refers to the capacity of an instrument to measure what it is designed to measure. To measure the construct validity of the specific questionnaires for the knee, we correlated the values from these questionnaires with the physical domains of the SF-36, which deal with the factors of functional capacity, presence of pain and limitations due to physical problems. The convergent and divergent validities were assessed based on comparison between the results obtained through the WOMAC, IKDC and Lysholm questionnaires and the results relating to the eight domains of the SF-36, taking the hypothesis that the knee questionnaires would correlate better with the physical domains of the SF-36 than with its mental domains. The content validity was assessed through the distribution and occurrence of floor and ceiling effects, which occur respectively when a response attains a score of zero (the minimum) and when a response attains a score of 100 (i.e. the maximum possible).

The internal consistency and concordance between the questionnaires for examining the homogeneity between the items of a scale were calculated<sup>(14)</sup>. Concordance refers to how close the values of two or more instruments are and, according to de Vet et al<sup>(13)</sup>, it denotes the absence of measurement errors.

## Statistical analysis

The convergent and divergent construct validities were tested by means of Pearson's correlation coefficient ( $r$ ). The internal consistency was calculated by means of Cronbach's  $\alpha$  coefficient<sup>(14)</sup>. We examined the concordance between the questionnaires using the graphical representations of Altman-Bland<sup>(15,16)</sup> and Concordance-Survival<sup>(17)</sup>. Altman and Bland<sup>(15,16)</sup> proposed that the concordance limits should be calculated from the differences observed, using the mean and standard deviation of the differences as the basis for calculating the limits. On the other hand, the concordance-survival curve<sup>(17)</sup> expresses the degree of discordance as a function of the diverse tolerance limits, similar to the Kaplan-Myer survival analysis. However, instead of absolute differences observed between the measurements, this technique locates the modulus of the differences on the X axis and the proportion of discordant cases on the Y axis.

To determine the influence of age on the results, we examined the association between age and the scores from the WOMAC, IKDC and Lysholm questionnaires by means of Pearson's coefficient of determination ( $r^2$ ). The statistical analyses were performed using the Statistical Package for the Social Sciences software for Windows (SPSS Science Inc, version 13.00, Chicago, IL, USA), taking a confidence level of 5%.

## RESULTS

The means, standard deviations and confidence intervals are presented in Table 1.

**Table 1** - Means, standard deviations (SD) and 95% confidence intervals for the IKDC, SF-36, Lysholm and WOMAC questionnaires in the study sample.

	Mean	SD	95% CI
IKDC	29.0	12.0	25.8-32.2
Lysholm Score	31.6	19.2	25.7-37.4
WOMAC (Total)	42.9	19.1	37.7-48.0
SF-36 SPC	34.1	19.6	28.9-32.2
SF-36 SMC	65.1	20.6	59.7-70.6
SF-36 FC	34.3	21.0	28.7-39.9
SF-36 LP	30.3	39.7	19.7-40.8
SF-36 Pain	35.1	21.9	29.3-40.9
SF-36 GH	58.9	25.6	52.1-65.7
SF-36 VT	57.9	21.4	52.2-63.6
SF-36 SF	68.9	28.0	61.4-76.3
SF-36 LE	43.3	43.6	31.7-54.9
SF-36 MH	66.5	21.8	60.7-72.3

### Validity

IKDC and WOMAC presented moderate to strong correlations in relation to the summary of physical capacity of SF-36, while the Lysholm questionnaire presented moderate correlations. All three questionnaires presented weak correlations with the mental components of SF-36, thus confirming the convergent and divergent validities (Table 2). The three questionnaires presented good content validity, given that no floor or ceiling effects were observed.

### Internal consistency

Cronbach's  $\alpha$  demonstrated values of 0.811 for IKDC, 0.959 for WOMAC and 0.734 for Lysholm. When one item at a time was excluded, Cronbach's  $\alpha$  was no larger than the original in any of the questionnaires, which negated any need to exclude any items from the scales when applied to patients with OA.

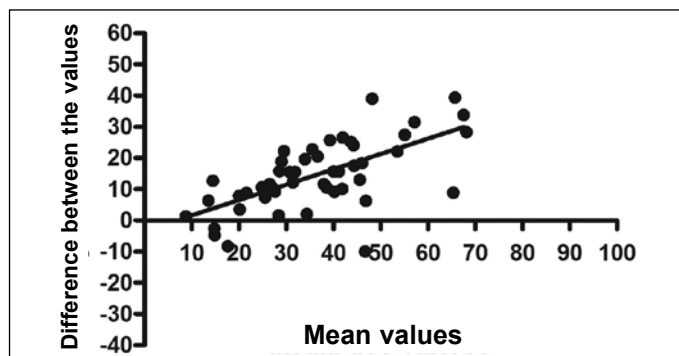
**Table 2** - Correlations between the questionnaires applied (IKDC, Lysholm and WOMAC) and the components of SF-36.

	Questionnaire	R	p-value
SF-36 SPC	IKDC	0.62	< 0.0001
	Lysholm	0.555	0.001
	WOMAC	0.642	< 0.0001
SF-36 SMC	IKDC	0.251	0.0592
	Lysholm	0.199	0.1945
	WOMAC	0.167	0.2217
SF-36 FC	IKDC	0.485	0.0001
	Lysholm	0.414	0.0052
	WOMAC	0.522	< 0.0001
SF-36 LP	IKDC	0.419	0.0012
	Lysholm	0.374	0.0125
	WOMAC	0.524	< 0.0001
SF-36 Pain	IKDC	0.651	< 0.0001
	Lysholm	0.652	< 0.0001
	WOMAC	0.580	< 0.0001
SF-36 GH	IKDC	0.337	0.0103
	Lysholm	0.205	0.182
	WOMAC	0.261	0.0546
SF-36 VT	IKDC	0.275	0.0384
	Lysholm	0.232	0.13
	WOMAC	0.244	0.073
SF-36 SF	IKDC	0.302	0.022
	Lysholm	0.256	0.093
	WOMAC	0.203	0.136
SF-36 LE	IKDC	0.286	0.0308
	Lysholm	0.282	0.0637
	WOMAC	0.350	0.0087
SF-36 MH	IKDC	0.148	0.2713
	Lysholm	0.099	0.5233
	WOMAC	0.061	0.6592

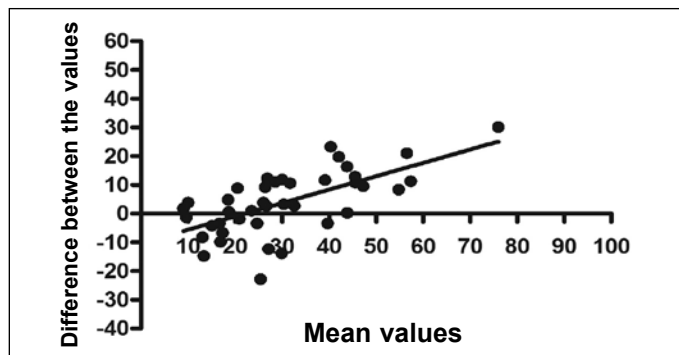
SPC: Summary of physical capacities; SMC: Summary of mental capacities; FC: Functional capacity; LP: Limitation due to physical factors; GH: General state of health; VT: Vitality; SF: Social factors; LE: Limitation due to emotional factors; MH: Mental health.

### Concordance

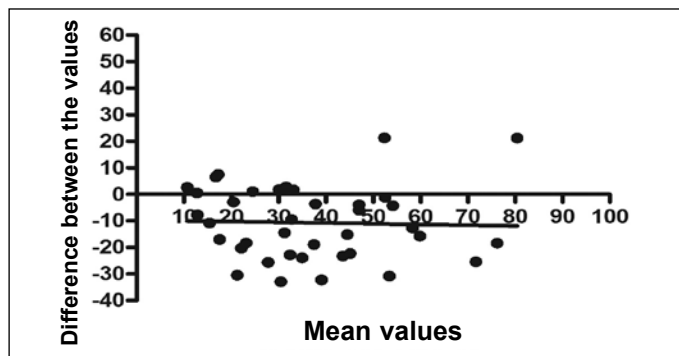
Despite the strong correlations between WOMAC and IKDC (0.843), WOMAC and Lysholm (0.759) and IKDC and Lysholm (0.858), the graphical representations of Altman-Bland and Concordance-Survival showed that the concordance between the three questionnaires was low. The linear regression curve represented in the Altman-Bland graph demonstrated the presence of proportional bias between IKDC and WOMAC (Figure 1) and IKDC and Lysholm (Figure 2), while between WOMAC and Lysholm it seemed that there was a fixed bias, such that Lysholm was on average ten points lower than WOMAC (Figure 3). The representation of concordance-survival (Figure 4) confirmed the findings from the Altman-Bland graphs, regarding the lack of concordance between the three questionnaires.



**Figure 1** - Altman-Bland representation of the comparison between IKDC and WOMAC.



**Figure 2** - Altman-Bland representation of the comparison between IKDC and Lysholm.



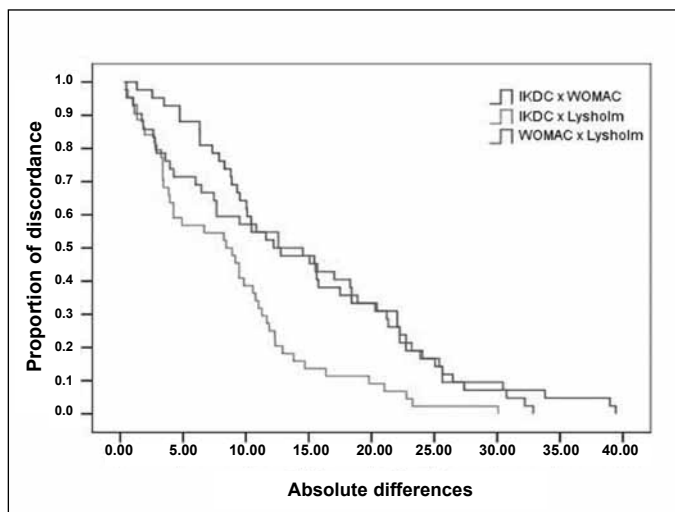
**Figure 3** - Altman-Bland representation of the comparison between WOMAC and Lysholm.

### Influence of age

The IKDC, Lysholm and WOMAC questionnaires presented Pearson's coefficients of determination ( $r^2$ ) of 0.004, 0.010 and 0.043 with age, respectively.

## DISCUSSION

Many studies have indicated the importance of using questionnaires for functional evaluations on patients affected by musculoskeletal diseases. Some of them have already been developed, translated and validated for the Brazilian population in relation to



**Figure 4** - Concordance-survival test on the three comparisons (IKDC versus WOMAC, IKDC versus Lysholm and WOMAC versus Lysholm).

evaluations on patients with OA in the knee, including WOMAC<sup>(18)</sup>, IKDC<sup>(12)</sup> and Lysholm<sup>(19)</sup>. However, selection of the ideal questionnaire for assessing the limitations caused by specific diseases like OA is a matter that has been little discussed.

Although all of these questionnaires present excellent psychometric properties for populations with joint diseases, it has not been determined which one would have the greatest capacity for ascertaining the physical limitations caused by OA. Since the incidence of OA increases with age, there is a need to estimate whether the responses obtained from the questionnaires would be influenced by this variable, independent of the severity of the clinical condition. In the present study, we chose to compare the criteria of scientific authenticity, validity, internal consistency and concordance between the WOMAC, IKDC and Lysholm questionnaires, in relation to patients with OA in the knee, and to identify the influence of the variable of age in determining the values of these parameters. No test-retest reproducibility verification was performed, given that the original studies of the Brazilian versions of the three questionnaires already presented excellent results. Consequently, it could be presumed that these properties would remain unchanged<sup>(18)</sup>.

Regarding the validity of the questionnaires, it was found that WOMAC and IKDC presented stronger correlations with the physical components of SF-36 than did the Lysholm questionnaire. Nevertheless, none of the questionnaires presented very high

correlations with these components (Table 2). This suggests that WOMAC is more suitable for assessing the limitations relating to physical factors, while IKDC and Lysholm are more suited for assessing pain-related factors among these patients. Both WOMAC and IKDC were shown to be suitable for analyzing functional capacity. Despite these differences, all three questionnaires presented results that were more strongly associated with the physical capacities of SF-36 than with its mental capacities, thus ratifying the convergent and divergent validities of these questionnaires.

The internal consistency of the three questionnaires was shown to be adequate, although WOMAC presented the highest values, followed by IKDC. When one item was excluded from the analysis of the Cronbach index, the values remained smaller than with the complete scale, which does away with the need to exclude any item, for these questionnaires to be used with these patients. The higher values presented by WOMAC were expected, given that the original selection of the items<sup>(9)</sup> was developed specifically for patients with osteoarthritis, while IKDC and Lysholm are questionnaires for overall assessments on knees.

Despite the better correlation between IKDC and WOMAC in relation to Lysholm and WOMAC, the Altman-Bland graph (Figures 1, 2 and 3) showed low concordance between the questionnaires, which was confirmed in the concordance-survival representation (Figure 4). According to de Vet et al<sup>(13)</sup>, concordance (agreement) and reliability are two psychometric properties that are included in the concept of reproducibility. They signify how close the values of two or more scales are to each other. De Vet et al<sup>(13)</sup> defined concordance as the absence of measurement errors. In our study, the concordance-survival graph demonstrated that the best agreement was between the IKDC and Lysholm questionnaires (Figure 4).

To achieve 70% concordance between individuals, the mean difference needed between these two questionnaires was approximately 10 points, while between IKDC and WOMAC and between WOMAC and Lysholm, it was 20 points. The presence of proportional bias between IKDC and WOMAC (Figure 1) and between IKDC and Lysholm (Figure 2) suggests that the greater the mean from the responses to the two questionnaires was, the greater the difference between

them would be. On the other hand, between WOMAC and Lysholm (Figure 3), we found that there was a relatively fixed bias, with a mean difference of approximately 10 points between the two questionnaires, such that WOMAC tended to have higher values than shown by Lysholm.

None of the questionnaires was influenced by the variable of age, since the low values of the coefficient of determination demonstrated that it would, at most, interfere by 5% in the final value from the questionnaires. If there had been a large negative correlation, we might have affirmed that age was influencing the values of the questionnaires, independent of the severity of the OA, but this did not happen. Hence, age was not shown to be a limiting factor for the use of any of the questionnaires applied in this study.

We take the view that the absence of the Lequesne Index in this assessment, an index that is specific for patients with OA, as is WOMAC, may be considered to be a limiting factor in this study. However, it is well known that extensive tests discourage correct responses from patients. Moreover, it is believed that between WOMAC and the Lequesne Index, the response would be specific, given that the construct validity is similar between these two questionnaires<sup>(20)</sup>. The lack of tests commonly used to grade the severity of OA, such as the classifications of Kellgren or Ahlback, for comparisons with the values obtained from the questionnaires may also be considered to be a limiting factor, but this was not an objective of the present study.

## CONCLUSION

The concordance tests and the correlations with the physical components of the SF-36 suggest that none of the three questionnaires alone is capable of assessing all the matters relating to the physical limitations of patients with osteoarthritis. WOMAC in association with IKDC presented the best correlation values with the summary of the physical components of SF-36. While WOMAC assesses functional capacities and limitations relating to physical factors, IKDC seems to be better suited for assessing the functional limitations relating to pain. Age was not shown to be a factor limiting the use of any of the questionnaires applied in this study.

## REFERENCES

1. Martin DF. Pathomechanics of knee osteoarthritis. *Med Sci Sports Exerc.* 1994;26(12):1429-34.
2. Lanyon P, O'Reilly S, Jones A, Doherty M. Radiographic assessment of symptomatic knee osteoarthritis in the community: definitions and normal joint space. *Ann Rheum Dis.* 1998;57(10):595-601.
3. Wilson MG, Michet CJ Jr, Ilstrup DM, Melton LJ 3rd. Idiopathic symptomatic osteoarthritis of the hip and knee: a population-based incidence study. *Mayo Clin Proc.* 1990; 65(9):1214-21.
4. Murray CJL, Lopez AD. The global burden of disease. Geneva: World Health Organization; 1997.
5. Pepin SR, Griffith CJ, Wijdicks CA, Goerke U, McNulty MA, Parker JB et al. A comparative analysis of 7.0-Tesla magnetic resonance imaging and histology measurements of knee articular cartilage in a canine posterolateral knee injury model: a preliminary analysis. *Am J Sports Med.* 2009; 37(Suppl 1):119S-24S.
6. Neves EB. Desenvolvimento de instrumentos de diagnóstico e acompanhamento da osteoartrite em pára-quedistas militares [tese]. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2009.
7. Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum.* 1995;38(8):1134-41.
8. Marx RG. Knee rating scales. *Arthroscopy.* 2003;19(10):1103-8.
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. Lequesne MG. The algofunctional indices for hip and knee osteoarthritis. *J Reumatol.* 1997;24(4):779-781.
11. Hamby K, Griva K. IKDC or KOOS? Which measures symptoms and disabilities most important to postoperative articular cartilage repair patients? *Am J Sports Med.* 2008;36(9):1695-704.
12. Metsavaht L, Leporace G, Riberto M, de Mello Sposito MM, Batista LA. Translation and cross-cultural adaptation of the Brazilian version of the International Knee Documentation Committee Subjective Knee Form: validity and reproducibility. *Am J Sports Med.* 2010;38(9):1894-9.
13. de Vet HCW, Terwee CB, Bouter LM. Current challenges in clinimetrics. *J Clin Epidemiol.* 2003;56(12):1137-41.
14. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull.* 1955;52(4):281-302.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-10.
16. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *Statistician.* 1983;32:307-17.
17. Luiz, RR, Costa AJL, Kale PL, Werneck GL. Assessment of agreement of a quantitative variable: a new graphical approach. *J Clin Epidemiol.* 2003;56(10):593-7.
18. Fernandes MI. Tradução e validação do questionário de qualidade de vida específico para osteoartrose WOMAC (Western Ontario McMaster Universities) para a língua portuguesa. [dissertação]. São Paulo: Universidade Federal de São Paulo, Escola Paulista de Medicina; 2003.
19. Peccin MS, Ciconelli R, Cohen M. Questionário específico para sintomas do joelho "Lysholm Knee Scoring Scale": tradução e validação para a língua portuguesa. *Acta Ortop. Bras.* 2006;14(5):268-72.
20. Marx FC, Oliveira LM, Belline CD, Ribeiro MC. Translation and cultural validation of the Lequesne's Algofunctional questionnaire for osteoarthritis of knee and hip for Portuguese language. *Braz J Reumatol.* 2006;46(4):253-60.