Reliability of a Proposed Ultrasonographic Grading Scale for Severity of Primary Knee Osteoarthritis

Mohamed Mortada1, Ayman Zeid2, Mirvat Abd El-hamid Al-Toukhy1, Nillie Ezzeldin1 and M. Elgawish1

1Rheumatology and Rehabilitation Department, Faculty of Medicine, Zagazig University, Zagazig, Egypt. 2Radiodiagnosis Department, Faculty of Medicine, Zagazig University, Zagazig, Egypt.

ABSTRACT
OBJECTIVES: The objective of this study was to examine the concordance of a grading scale (0–4) of medial femoral osteophytes in knee joint detected by ultrasound (US) compared with the corresponding grades (0–4) of Kellgren–Lawrence (K&L) scale of conventional radiography and clinical joint examination.

PATIENTS AND METHODS: A cross-sectional observational study included 160 patients with knee pain who fulfilled the American College of Rheumatology (ACR) criteria for knee osteoarthritis (KOA) and 20 patients with knee pain who have not fulfilled the ACR criteria for KOA. All patients were subjected to clinical assessment (Western Ontario and McMaster Universities Index of Osteoarthritis and global visual analog scale) and radiological assessment in the form of X-ray grading according to K&L grading scale and ultrasonographic assessment of medial femoral osteophytes according to a scale that was proposed by the first author and consisted of five grades (0–4), where grade 0 denoted no osteoarthritis and grade 4 denoted the most advanced grade of KOA. Grade 2 was divided into two subgrades A and B with grade 2B considered as a more advanced stage than grade 2A.

RESULTS: The proposed US grading scale had high sensitivity and specificity in detecting the different grades of KOA compared with K&L grading scale (a total sensitivity of 94.6% and a total specificity of 93.3%). Intra- and interreader reliability of US was excellent (kappa >0.93 and >0.85, respectively).

CONCLUSIONS: US can reliably detect the severity of KOA. Good agreement was found between the proposed US grading scale and K&L grading scale. The proposed US grading scale is simple and reliable.

KEYWORDS: knee, osteoarthritis, ultrasound

Introduction
Knee osteoarthritis (KOA) in the elderly is one of the most common degenerative age-related joint diseases, leading to pain and limitation of joint motion.1 Conventional radiography (CR) is the standard imaging modality for the assessment of structural abnormalities in KOA.2 Determination of severity stage of knee OA has several implications in clinical practice and clinical trials.3,4

Kellgren–Lawrence (K&L) grading scheme is the most widely used and accepted standard for diagnosis of different grades of radiographic KOA.5 Recently, ultrasound (US) has enhanced OA diagnosis and management.6 Current US technology offers many advantages, including the ability to detect osteophytes, degeneration of articular cartilage, tears of menisci and ligaments, bursitis, the presence of effusion,7 and utility in interventional procedures.8 Furthermore, US is cost-effective.9 The published research priorities by the Outcome Measures in Rheumatology Ultrasound Task Force included a validation of ultrasonographic scoring systems for osteoarthritis.10 Till now, there is no consensus on a US scoring system for the assessment of knee joint in KOA.11–16

Osteophytes are very common as a radiographic feature of osteoarthritis. They have been used to define the presence of disease.17 Osteophytes are associated with KOA progression that is partly explained by the strong association of osteophytes with malalignment to the same side of the osteophyte.18 Furthermore, US has been shown to be more sensitive than CR in the detection of knee osteophytes.19

The main objective of the present study on KOA patients was to examine the concordance of a grading scale (0–4) of medial femoral osteophytes in knee joint detected by US compared with the corresponding grades (0–4) of K&L scale of CR and clinical joint examination.

As a high reliability was recently shown for the scoring of hand osteoarthritis by the use of a US atlas as reference,20 the second objective of this study was to develop a US atlas, including a semiquantitative scoring of osteophytes, for the examination of reliability of the new sonographic scoring.
Patients and Methods

Patients. One hundred and sixty patients with knee pain who fulfilled the American College of Rheumatology (ACR) criteria for KOA and 20 patients with knee pain who have not fulfilled the ACR criteria for KOA were randomly recruited for the present cross-sectional observational study from Zagazig Rheumatology and Rehabilitation outpatient clinic during the period between June 2013 and May 2014. The inclusion criterion was fulfillment of the clinical or the radiological criteria for KOA of the ACR. Exclusion criteria were (1) secondary knee OA and (2) a history of knee surgery. An approval had been obtained from the institutional review board of our university, and all participants signed an informed consent. The research was conducted in accordance with the principles of the Declaration of Helsinki.

Clinical assessment. One experienced rheumatologist (MAA), blinded to the imaging results, performed clinical examinations of the selected knees, including detection of the degree of tenderness.

The patients also completed standard questionnaires including: demographic characteristics, body mass index (BMI), a 100-mm visual analog scale (VAS) for patient global assessment, and the Western Ontario and McMaster Universities (WOMAC) OA Index of pain, stiffness, and function subscales.

Conventional radiography. Two experienced rheumatologists (NE and MG), blinded to patient information during the evaluation, reached a consensus on the image assessments and used the K&L score to grade osteoarthritic cases on standing weight-bearing anteroposterior radiographs. In using the K&L score, an atlas was used to help grade the radiographs. The K&L score has five categories: grade 0 is given if no osteoarthritic features are present, grade 1 indicates minute osteophytes of doubtful importance, grade 2 indicates definite osteophytes without a reduction of the joint space, grade 3 is given when the joint space has diminished, and grade 4 indicates a greatly reduced joint space and sclerosis of the subchondral bone.

Ultrasoundography. Two sonographers (AZ, a radiologist, and MM, a rheumatologist; both are experienced in MSUS) performed ultrasonographic examinations of the selected knees. Grading of KOA was done according to a scale proposed by the first author, which depended on the shape of distal femoral osteophytes. The scale consisted of five grades (0–4), where grade 0 denoted no osteoarthritis and grade 4 denoted the most advanced grade of KOA, as presented in Table 1. Grade 2 was divided into two subgrades A and B with grade 2B considered as a more advanced stage than grade 2A. An atlas of different grades of the proposed scale was prepared by the first author and used by the sonographers as a guide for the grading of KOA. As illustrated in Figure 1, the atlas presented three examples of each score (18 examples).

The patients were positioned supine with their knees in full extension. The knees were scanned longitudinally in the coronal plane at the level of the knee joint, where the medial collateral ligament and the body of the medial meniscus could be detected. The probe was moved in the coronal plane from superior to inferior to detect osteophytes that were defined as cortical protrusions. The distal medial osteophytes were graded independent of the number, size, and location of other osteophytes in the same joint. If there were different shapes of the osteophyte during scanning, the most advanced one was considered according to the proposed scale.

A linear array transducer was used (5–12 MHz, Medison, R3) with fixed settings of the machine. To ensure standardization, the same US machine without software upgrading was used throughout the study. Each patient was examined by the two ultrasonographers on the same day. The images and its grades detected by each sonographer were recorded separately and delivered to another member of the research team. If a discrepancy was noticed, it was resolved by consensus of the two ultrasonographers. The sonographers were blinded to the results of other assessments.

Reliability of US. For the reliability (interreader and intrareader) of the grades of KOA detected by the proposed US grading scale, the patients were asked to participate in a follow-up US scanning (10–16 months after the first scanning). Eighty-six patients accepted and completed the follow-up US scanning, in which both ultrasonographers separately performed another two grading (one week apart) of the presented images according to the proposed US grading scale.

Statistical analysis. Analyses were performed using SPSS software (version 20, SPSS Inc.). The correlation between the proposed US scale and other measures of disease activity was estimated by Spearman’s rank correlation coefficient. The level of significance was chosen as $P = 0.05$. Intra- and interobserver’s reproducibility was assessed using the kappa statistic and 95% confidence interval.

Results

Demographic variables and clinical data from the cohort are presented in Table 2.
The prevalence of different grades of KOA according to K&L scale of CR and the proposed ultrasonography in patients with KOA is shown in Figure 2.

As regard to concordance between the proposed US grading scale and the K&L grading scale, the proposed US grading scale had high sensitivity, specificity, positive, and negative prediction values in detecting the different grades of KOA compared with K&L grading scale as presented in Table 3.

Most of the subjects in the control group had grade 0 by both K&L and US grading scales (17 and 18), respectively. Three subjects had grade 1 by K&L scale (two of them had grade 0 and one had grade 1 by US scale).

The interreader reliability of US detected grades was very good (agreement: 82.3%, difference: 13.5%, kappa: 0.81, \( P \leq 0.001 \)). The intrareader reliability for the two ultrasonographers was very good; for the first reader (MM), agreement: 89.4%, difference: 12.4%, kappa: 0.87, \( P \leq 0.001 \), and for the second reader (AZ), agreement: 86.5%, difference: 13.5%, kappa: 0.82, \( P \leq 0.001 \).

The proposed US grading scale was highly significantly correlated with age (\( r = 0.26, \ P = 0.02 \)), disease duration (time since an X-ray established the diagnosis; \( r = 0.23, \ P \leq 0.001 \)), and BMI (\( r = 0.39, \ P \leq 0.001 \)), whereas there was a nonsignificant correlation between the proposed US grading scale and

---

**Figure 1.** Different grades of KOA according to the proposed ultrasonographic scale.

**Abbreviations:** f, femur; t, tibia.
VAS, WOMAC pain subscale, stiffness subscale, function subscale, and total WOMAC (Table 4).

**Discussion**

To the best of our knowledge, the presented US grading scale of KOA is the first US grading that consisted of four grades according to medial femoral osteophytes with each grade directly correlated with its corresponding grade of the gold standard K&L grading of CR.

There were previous studies that graded the ultrasonographic features of KOA according to the thickness of femoral condylar cartilage. As the degeneration of the femoral cartilage usually results in an irregular margin and different thicknesses within the same degenerated joint, in our experience, the application of grading systems that depend on the shape or thickness of femoral cartilage in daily practice is liable to inter- and intrareader inconsistency.

Wu et al. and Saarakkala et al. graded femoral osteophytes by US into only three semiquantitative grades (mild, moderate, and large osteophytes), while K&L grading scale of CR consisted of four grades (suspicious, mild, moderate, and severe) in addition to grade zero, in which there was no evidence of KOA. In addition, both the studies did not aim or result in a global scale for KOA with the corresponding grades of the gold standard K&L grading of CR. Also, Hernandez Díaz et al. proposed a scoring system of both femoral and tibial (medial and lateral) osteophytes and found that the total score of osteophytes was correlated with K&L grading scale of CR, but the score depends on multiple factors and four different areas, which may also be liable to inter- and intrareader inconsistency.

The choice of the US shape of medial femoral osteophytes to be the base of the proposed scale may be supported by the results of Saarakkala et al, who found that the grades of US-detected osteophytes at medial compartment had a better correlation with arthroscopic grading of KOA than the grades of US-detected osteophytes at lateral compartment.

The good concordance between the proposed US grading scale and the K&L grading scale and the ability of the

**Table 2.** Demographic and clinical characteristics of the 180 patients with primary knee OA.

| VARIABLES | NO | (%) |
|-----------|----|-----|
| Women     | 115 | 63.9 |
| Males     | 65  | 36.1 |
| Age, years (mean ± SD) | 50.1 (±9.2) |
| Disease duration, years (mean ± SD) | 14.3 (±5.2) |
| Body mass index, kg/m² (mean ± SD) | 27.4 (±3.1) |
| WOMAC pain subscale (mean ± SD) | 12.3 (±3.6) |
| WOMAC stiffness subscale (mean ± SD) | 4.5 (±1.5) |
| WOMAC function subscale (mean ± SD) | 42.4 (±9.2) |
| WOMAC total scale (mean ± SD) | 59.2 (±11.2) |
| VAS (mean ± SD) | 63.0 (±16.1) |

**Abbreviations:** No, number; SD, standard deviation; OA, osteoarthritis; WOMAC, Western Ontario and McMaster Universities index; VAS, visual analog scale.

**Figure 2.** The frequency distribution of different grades of KOA according to K&L scale of CR and the proposed ultrasonographic scale in patients with KOA.

**Abbreviations:** CR, conventional radiography; US, ultrasound; PPV, positive predictive values; NPV, negative predictive values; NLR, Negative likelihood ratio.

**Table 3.** Agreement between the grades detected according to K&L scale of CR and the proposed ultrasonography in patients with KOA.

| OSTEOPTHES BY US | OSTEOPTHES GRADES BY CR | SENSITIVITY OF US* | SPECIFICITY OF US* | PPV  | NPV  | NLR |
|------------------|------------------------|--------------------|--------------------|------|------|-----|
|                  | 0 1 2 3 4              | TOTAL SENSITIVITY: | TOTAL SPECIFICITY: |      |      |     |
|                  |                         | 95%                | 95%                |      |      |     |
| 0                | 16 2 0 0 0             | 94.1%              | 98.8%              | 88.9%| 99.4%| 0.06|
| 1                | 1 26 0 0 0             | 92.9%              | 99.3%              | 96.3%| 98.7%| 0.07|
| 2                | 2A 0 0 29 0            | 93.6%              | 100%               | 100% | 98.4%| 0.06|
|                  | 2B 0 0 22 0            | 91.6%              | 100%               | 100% | 98.5%| 0.08|
| 3                | 0 0 2 44 3             | 88.2%              | 96.3%              | 89.8%| 99.2%| 0.02|
| 4                | 0 0 0 1 34             | 91.9%              | 99.3%              | 97.1%| 97.9%| 0.08|

**Note:** *Sensitivity of US with the reference being the K&L grading of CR.

**Abbreviations:** CR, conventional radiography; US, ultrasound; PPV, positive predictive values; NPV, negative predictive values; NLR, Negative likelihood ratio.
Table 4. Correlation between the CR/ultrasonography-detected grades of KOA and clinical parameters.

|        | Grades by CR | Grades by US |
|--------|--------------|--------------|
| Age    | 0.27 (0.01*) | 0.26 (0.02*) |
| BMS    | 0.42 (0.03*) | 0.39 (0.00*) |
| Disease duration | 0.49 (0.00*) | 0.23 (0.00*) |
| WOMAC pain | 0.14 (0.17) | 0.35 (0.09) |
| WOMAC stiffness | 0.11 (0.23) | 0.12 (0.19) |
| WOMAC function | 0.13 (0.18) | 0.11 (0.23) |
| WOMAC total | 0.14 (0.17) | 0.15 (0.16) |
| VAS    | 0.15 (0.16) | 0.14 (0.17) |

Abbreviations: CR, conventional radiography; US, ultrasound.

proposed scale to discriminate adequately between the different degrees of K&L scale may denote that the proposed US scale could be used as the comparable US scale of the K&L scale of CR.

As grade 0 in the proposed US grading scale denotes the absence of degenerative changes, using the proposed scale may allow the discrimination between knee joints with and without osteoarthritis.

There were high intra-/interreader reliabilities of the detected grades of KOA by the US scale, which may relate to the simplicity of the proposed US scale being dependent only on the ultrasonographic shape of medial femoral osteophyte.

In the present study, there was a discrepancy between the K&L grading scale and the disease activity parameters including VAS, tenderness, WOMAC pain, stiffness subscales, and WOMAC total score. This was in agreement with previous studies and may be attributed to the difference between knee OA activity and severity parameters. Also, such discrepancy was found between the grades of the proposed US scale for knee OA severity and disease activity parameters. This may be explained by the fact that the proposed US scale was aimed mainly to be correlated with K&L grading scale.

Because grade 2 of K&L scale of CR correlated with the different shapes of femoral osteophytes by US, the proposed US scale grade 2 was divided into two subgrades (2A and 2B) with subgrade 2B considered as a progression of subgrade 2A. The clinical value of this subdivision is still needed to be detected in further studies.

The limitations of this study included the absence of correlation of the proposed grades with MRI findings, which should be looked for in future studies. Also, the proposed scale depended only on ultrasonographic features, which did not provide information about the changes that occur in the subchondral bone marrow and that are related to structural and symptomatic osteoarthritis progression.

In conclusion, the proposed ultrasonographic grading scale of knee OA was proved to be a reliable and a comparable tool to K&L scale of CR. Good agreement was found between the proposed US grading scale and K&L grading scale. One of the advantages of the proposed scale is its simplicity, being dependent only on one item (medial femoral osteophyte), which proved to be in a good concordance with K&L grading of CR in this study.

Acknowledgments
A preliminary version of this work was presented at EULAR 2014 in Paris, France, and the abstract was subsequently published in Ann Rheum Dis as part of the conference proceedings. Mortada M., Al-Toukhy M., Ezz Eldein N., Zeid A. SAT0434 Validation of A Proposed Ultrasonographic Grading Scale for Severity of Primary Knee Osteoarthritis. Ann Rheum Dis 2014;73:751 doi:10.1136/annrheumdis-2014-eular.1703.

Author Contributions
Conceived and designed the experiments: MM, AZ, MAI, NE, ME. Analyzed the data: MM, AZ, MAI, NE, ME. Wrote the first draft of the manuscript: MM, AZ, MAI, NE, ME. Contributed to the writing of the manuscript: MM, AZ, MAI, NE, ME. Agree with manuscript results and conclusions: MM, AZ, MAI, NE, ME. Jointly developed the structure and arguments for the paper: MM, AZ, MAI, NE, ME. Made critical revisions and approved final version: MM, AZ, MAI, NE, ME. All authors reviewed and approved of the final manuscript.

REFERENCES
1. Lawrence RC, Felton DT, Helmick CG, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum. 2008;59(1):26–35.
2. Roemer FW, Eckstein F, Hayashi D, Guermazi A. The role of imaging in osteoarthritis. Best Pract Res Clin Rheumatol. 2014;28(1):31–60.
3. Keurentjes JC, Fiocco M, So-Osman C, et al. Patients with severe radiographic osteoarthritis have a better prognosis in physical functioning after hip and knee replacement: a cohort-study. PLoS One. 2013;8:e59506.
4. Buckwalter JA, Lohmander S. Operative treatment of osteoarthritis. Current practice and future development. J Bone Joint Surg Am. 1994;76:1405–18.
5. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthritis. Ann Rheum Dis. 1957;16(4):494–502.
6. Braun HJ, Gold GE. Diagnosis of osteoarthritis: imaging. Bone. 2012;51(2):278–88.
7. Wakefield RJ, Gibbon WW, Emery P. The current status of ultrasonography in rheumatology. Rheumatology. 1999;38:195–8.
8. Qvistgaard A, Kristofferson H, Terslev L, Danneskiold-Samsoe B, Torp-Pedersen S, Bliddal H. Guidance by ultrasound of intra-articular injections in the knee and hip joints. Osteoarthrits Cartilage. 2001;9:512–7.
9. Guermazi A, Eckstein F, Helio Le Graverand-Gastineau MP, et al. Osteoarthritis: current role of imaging. Med Clin North Am. 2009;93:101–26, xi.
10. Naredo E, Wakefield RJ, Iagnocco A, et al. The OMERACT ultrasound task force status and perspectives. J Rheumatol. 2011;38:2063–7.
11. Podlipská J, Koski JM, Pulkkinen P, Saarakkala S. In vivo quantitative ultrasound image analysis of femoral subchondral bone in knee osteoarthritis. Scientific World Journal. 2013;2013:182562.
12. Saarakkala S, Waris F, Waris V, et al. Diagnostic performance of knee ultrasound for detecting degenerative changes of articular cartilage. Osteoarthrits Cartilage. 2012;20(5):376–81.
13. Tsai CY, Lee CL, Chai CY, et al. The validity of in vitro ultrasonographic grading of osteoarthritic femoral condylar cartilage—a comparison with histologic grading. Osteoarthrits Cartilage. 2007;15(3):245–50.
14. Lee CL, Huang MH, Chai CY, Chen CH, Su JY, Tien YC. The validity of in vivo ultrasonographic grading of osteoarthritic femoral condylar cartilage: a comparison with in vitro ultrasonographic and histologic grading. Osteoarthrits Cartilage. 2008;16(3):352–8.
15. Wu P-T, Shao C-J, Wu K-C, et al. Pain in patients with equal radiographic grades of osteoarthritis in both knees: the value of gray scale ultrasound. Osteoarthrits Cartilage. 2012;20:1507–13.
16. Saarakkala S, Kamel A, Szkudlarek M, et al. Detection of knee osteophytes with ultrasonography and conventional radiography: intra- and inter-reader reliability and comparison to arthroscopic degeneration of articular cartilage [abstract]. In: Orthopedic research society 2013 Annual Meeting. Available at: http://www.ors.org/Transactions/59/001/0006.html. Accessed January 5, 2015.

17. Altman R, Asch E, Bloch D, et al. Development of criteria for the classification and reporting of osteoarthritis: Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum.* 1986;29:1039–49.

18. Felson DT, Gale DR, Eton DA, et al. Osteophytes and progression of knee osteoarthritis. *Rheumatology.* 2005;44(1):100–4.

19. Slimani S, Haddouche A, Bencharif I, Ladjouze-Rezig A. Superiority of knee ultrasound over radiographs in osteophyte detection in knee osteoarthritis. [Abstract]. *Osteoarthritis Cartilage.* 2013;21:S31–63.

20. Mathiessen A, Haugen IK, Slatkowsky-Christensen B, Bøyesen P, Kvien TK, Hammer HB. Ultrasonographic assessment of osteophytes in 127 patients with hand osteoarthritis: exploring reliability and associations with MRI, radiographs and clinical joint findings. *Ann Rheum Dis.* 2013;72:51–6.

21. Ahlman R, Asch E, Bloch D. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the knee. *Arthritis Rheum.* 1986;29:1039–49.

22. Bellamy N, Buchanan WW, Goldsmith CH. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to anti-rheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol.* 1988;15:1833–40.

23. Wakefield RJ, Balint PV, Szkudlarek M, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol.* 2005;32:2485–7.

24. Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions.* 3rd ed. New York, NY: John Wiley; 2003:598–605.

25. Hernández Díaz C, León-Hernández SR, Bernal A, Espinosa R, Hofmann F, Peña A. A simplified ultrasound score to assess knee osteoarthritis [abstract]. *Osteoarthritis Cartilage.* 2011;19(Suppl):S173.

26. Cubukcu D, Saracan A, Alkan H. Relationships between pain, function and radiographic findings in osteoarthritis of the knee: a cross-sectional study. *Arthritis.* 2012;2012:984060.

27. Finan PH, Buenaver LF, Bounds SC, et al. Discordance between pain and radiographic severity in knee osteoarthritis: findings from quantitative sensory testing of central sensitization. *Arthritis Rheum.* 2013;65:363–72.

28. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskelet Disord.* 2008;9(1):116.

29. Lequesne M. Indices of severity and disease activity for osteoarthritis. *Semin Arthritis Rheum.* 1991;20(6):48–54.

30. Zhang M, Driban JB, Price LL, Lo GH, Miller E, McAlindon TE. Development of a rapid cartilage damage quantification method for the lateral tibiofemoral compartment using magnetic resonance images: data from the osteoarthritis initiative. *Biomed Res Int.* 2015;2015:634275.