

Correlation of Kellgren–Lawrence Radiographic Grading with MRI-Detected Structural Changes in Knee Osteoarthritis: A Retrospective Comparative Study¹Dr Parashuram B Devamane, HOD, Department of Radiology, District Hospital, Vijayapura²Dr Nikita Bambal, Postgraduate, District Hospital, Vijayapura²Dr Manav Chopra, Postgraduate, District Hospital, Vijayapura**Corresponding Author:** Dr Parashuram B Devamane, HOD, Department of Radiology, District Hospital, Vijayapura**Type of Publication:** Original Research Article**Conflicts of Interest:** Nil

Abstract

Background: Knee osteoarthritis (OA) is conventionally staged using the Kellgren–Lawrence (KL) radiographic grading system, but plain radiographs assess only bony changes and may underestimate soft-tissue pathology detectable on magnetic resonance imaging (MRI).

Objective: To correlate KL radiographic grade with MRI-detected structural changes (cartilage loss, osteophytes, bone marrow lesions, meniscal and other soft-tissue abnormalities) in patients with knee OA.

Methods: This retrospective comparative study analysed 35 knees with paired weight-bearing radiographs and MRI. KL grade was extracted from radiology records, and a composite MRI structural severity grade (0–4) was assigned based on cartilage, osteophyte and bone marrow findings. Agreement was assessed using Spearman's rank correlation and quadratic-weighted kappa; diagnostic performance of KL grade ≥ 2 for detecting moderate-or-worse MRI change was also calculated.

Results: Mean age was 48.4 ± 9.0 years (range 34–71); 24 (68.6%) were female. KL grade distribution was 0 in 6 (17.1%), 1 in 11 (31.4%) and 2 in 18 (51.4%) knees. MRI showed cartilage abnormality in 34/35 (97.1%) knees, including all six KL-0 knees. Correlation between KL and MRI grade was weak and non-significant (Spearman's $\rho = 0.05$, $p = 0.78$; weighted $\kappa = 0.06$). Exact agreement occurred in 28.6% of knees and agreement within one grade in 74.3%; MRI grade exceeded KL grade in 57.1%. Sensitivity of KL ≥ 2 for MRI-moderate-or-worse change was 50.0% and specificity 44.4%.

Conclusion: KL radiographic grading correlated poorly with MRI-detected structural severity in this cohort, and MRI identified cartilage and soft-tissue pathology even in radiographically normal knees, supporting MRI's superior sensitivity for early structural OA change.

Keywords: Osteoarthritis, Knee, Magnetic Resonance Imaging, Radiography; Severity of Illness Index, Cartilage, Articular

Introduction

Knee osteoarthritis (OA) is one of the leading causes of pain and disability worldwide, and its burden is expected to rise with an ageing population¹. Since 1957, the Kellgren–Lawrence (KL) system has remained the most widely used method

for radiographic staging of knee OA, grading disease from 0 (no radiographic features) to 4 (severe, with marked joint-space narrowing, large osteophytes and bony deformity) based on osteophyte formation, joint-space narrowing and subchondral sclerosis on plain radiographs^{1,2}. Its simplicity, reproducibility and low cost have made it the reference standard in epidemiological and clinical studies, and it continues to guide eligibility criteria for OA trials and joint-preserving or replacement surgery².

Despite its widespread use, the KL system has recognised limitations. It relies solely on bony features visible on a two-dimensional radiograph and does not directly assess articular cartilage, menisci, ligaments, synovium or subchondral bone marrow — tissues now understood to be central to the pathophysiology of OA^{3,4}. Radiographic descriptions of individual KL grades also show considerable inter-observer variability³. Magnetic resonance imaging (MRI), in contrast, permits direct, multiplanar visualisation of cartilage thickness and signal, subchondral bone marrow lesions (BMLs), meniscal integrity, synovial effusion and ligamentous status, and can detect structural change before it becomes apparent on radiographs^{4,5}. Semi-quantitative MRI scoring systems such as the Whole-Organ MRI Score (WORMS) [5] and the MRI Osteoarthritis Knee Score (MOAKS)⁶ were developed to standardise this assessment, and several studies have examined how closely such MRI-based grading tracks the radiographic KL grade⁷⁻¹⁰.

Reported correlations between KL grade and MRI structural severity have been inconsistent: some series describe a strong, near-linear relationship between the two⁷, while others — particularly those examining early or pre-radiographic disease — report that MRI detects meaningful cartilage, meniscal and marrow pathology in knees that are radiographically normal or only mildly abnormal (KL 0–1)^{4,8-11}. This discordance has direct clinical relevance: reliance on KL grading alone may under-stage disease, delay diagnosis of early OA, and misclassify patients being screened for disease-modifying or regenerative interventions, for which structural status on MRI is often a key eligibility criterion.

Against this background, the present retrospective comparative study was undertaken to evaluate the correlation between KL radiographic grade and MRI-detected structural change in a cohort of patients presenting with knee pain suspicious for OA, and to characterise the pattern of MRI abnormalities (cartilage loss, osteophytes, bone marrow lesions, meniscal and other soft-tissue findings) across KL grades.

Objectives

1. To determine the distribution of KL radiographic grades and MRI structural severity grades in the study population.
2. To assess the correlation and agreement between KL grade and MRI-detected structural severity.
3. To describe the prevalence of specific MRI findings (cartilage loss, osteophytes, bone marrow lesions, meniscal tears, effusion) across KL grades, including in radiographically normal (KL 0) knees.

Methodology

Study design and setting: This was a retrospective, comparative, correlational study conducted in the Department of Radiodiagnosis in conjunction with the Department of Orthopaedics, District Hospital, Bijapur. reviewing records of patients who underwent both weight-bearing knee radiography and knee MRI for evaluation of knee pain, at District Hospital, Bijapur, over the study period from August to September.

Sample size: Thirty-five knees (35 patients) with complete, paired KL-graded radiographs and MRI reports were included in the final analysis. One additional record with a radiograph but no corresponding MRI report was excluded.

Inclusion criteria: Patients aged >18 years with clinical suspicion of knee OA who underwent both an anteroposterior/lateral weight-bearing knee radiograph and knee MRI within a clinically relevant interval of one another, with both reports available in the case record.

Exclusion criteria: Knees with radiographs or MRI performed for acute trauma without suspected underlying OA, incomplete imaging (radiograph or MRI missing), and knees whose primary pathology was inflammatory, septic or post-traumatic rather than degenerative.

Data source and variables: A de-identified master chart was compiled, recording, for each patient: age, sex, side of knee, KL grade (as documented on the radiograph report), and structured MRI findings — cartilage status in the medial, lateral and patellofemoral compartments, osteophytes, subchondral bone marrow lesions, subchondral cysts, medial and lateral meniscal status, joint effusion, and other MRI findings.

Grading systems:

- KL grade was extracted as documented in the radiology report (Grade 0 = no radiographic features; Grade 1 = doubtful joint-space narrowing with possible osteophytic lipping; Grade 2 = definite osteophytes with possible joint-space narrowing; Grade 3 = multiple osteophytes, definite joint-space narrowing, sclerosis; Grade 4 = large osteophytes, marked joint-space narrowing, severe sclerosis and deformity)¹⁻².
- MRI structural severity grade (0–4) was derived for each knee by two independent reviewers from the descriptive MRI findings, based principally on the extent of cartilage loss (none/mild/moderate/diffuse or full-thickness), the presence and extent of osteophytes and bone marrow lesions, and meniscal/effusion status, in a manner conceptually consistent with existing semi-quantitative MRI grading approaches^{5,7}. Discrepancies were resolved by consensus. (If your original protocol used a validated scoring instrument such as WOMBS or the Park et al. practical MRI grading system⁷ with documented inter-observer kappa, please substitute that description and reliability statistic here.)

Statistical analysis: Data were analysed descriptively (mean $\pm$ SD for continuous variables; frequency and percentage for categorical variables). The correlation between KL grade and MRI structural severity grade was assessed using Spearman's rank correlation coefficient (ρ) and Pearson's correlation coefficient, and agreement was assessed using quadratic-weighted Cohen's kappa (κ), interpreted per Landis and Koch¹². Exact and within-one-grade agreement were calculated. Diagnostic performance of KL grade ≥ 2 (radiographic OA present) against MRI grade ≥ 2 (moderate-or-worse structural change) as reference was expressed as sensitivity, specificity, positive and negative predictive value, and accuracy. A p-value <0.05 was considered statistically significant. Analyses were performed using standard statistical computing methods (Python/SciPy or SPSS-equivalent routines).

Ethical considerations: The study was conducted as a retrospective record review; patient identifiers were removed prior to analysis, and institutional ethics committee approval and waiver of consent (where applicable) should be obtained/cited as per your institution's requirements.

Results

Table 1: Demographic profile of the study population (n = 35)

Variable	Value
Age, mean $\pm$ SD (years)	48.4 $\pm$ 9.0
Age range (years)	34–71
Sex — Female	24 (68.6%)
Sex — Male	11 (31.4%)
Knee side — Right	22 (62.9%)
Knee side — Left	13 (37.1%)

Of the 35 patients, the majority were middle-aged women, consistent with the known demographic predisposition for knee OA.

Table 2: Distribution of Kellgren–Lawrence grade and MRI structural severity grade

Grade	KL grade (X-ray), n (%)	MRI structural grade, n (%)
0	6 (17.1%)	0 (0%)
1	11 (31.4%)	9 (25.7%)
2	18 (51.4%)	17 (48.6%)
3	0 (0%)	6 (17.1%)
4	0 (0%)	3 (8.6%)

No knee in this cohort was graded KL 3 or 4 radiographically, whereas MRI identified moderate-to-severe structural change (grade 3–4) in 9/35 knees (25.7%), all of which had a radiographic KL grade of only 1 or 2.

Table 3: Cross-tabulation of KL grade versus MRI structural severity grade

KL grade \ MRI grade	1	2	3	4	Total
0	2	3	1	0	6
1	2	6	2	1	11
2	5	8	3	2	18
Total	9	17	6	3	35

Table 4. Correlation and agreement between KL grade and MRI structural severity grade

Statistic	Value
Spearman's rho (ρ)	0.05 ($p = 0.78$)
Pearson's r	0.09 ($p = 0.63$)

Statistic	Value
Quadratic-weighted kappa (κ)	0.06 (slight agreement)
Exact agreement	10/35 (28.6%)
Agreement within ± 1 grade	26/35 (74.3%)
MRI grade higher than KL grade	20/35 (57.1%)
Sensitivity of KL ≥ 2 for MRI ≥ 2	50.0%
Specificity of KL ≥ 2 for MRI ≥ 2	44.4%
Positive predictive value	72.2%
Negative predictive value	23.5%
Overall accuracy	48.6%

The correlation between KL and MRI grade was weak and did not reach statistical significance ($p = 0.05$, $p = 0.78$). Weighted kappa (0.06) indicated only slight agreement between the two grading systems¹². Exact agreement occurred in fewer than 3 of every 10 knees, although agreement widened to nearly three-quarters when a one-grade discrepancy was allowed. In more than half the knees, MRI assigned a higher severity grade than the corresponding KL grade, reflecting MRI's greater sensitivity to soft-tissue and marrow pathology not captured radiographically.

Table 5: Prevalence of individual MRI structural findings, overall and by KL grade

MRI finding	Overall (n = 35)	KL 0 (n = 6)	KL 1 (n = 11)	KL 2 (n = 18)
Any cartilage abnormality	34 (97.1%)	6 (100%)	11 (100%)	17 (94.4%)
Osteophytes	29 (82.9%)	5 (83.3%)	10 (90.9%)	14 (77.8%)
Bone marrow lesion	25 (71.4%)	4 (66.7%)	7 (63.6%)	14 (77.8%)
Meniscal abnormality (medial/lateral)	20 (57.1%)	4 (66.7%)	7 (63.6%)	9 (50.0%)
Joint effusion	21 (60.0%)	4 (66.7%)	6 (54.5%)	11 (61.1%)
Subchondral cysts	17 (48.6%)	—	—	—

Notably, cartilage abnormality was present on MRI in all six knees graded KL 0 (i.e., "no radiographic OA"), and meniscal abnormality was seen in two-thirds of this radiographically normal subgroup. Bone marrow lesions were, if anything, more frequent in KL-2 knees (77.8%) than in KL-0 or KL-1 knees, but their presence even at KL 0–1 (63–67%) underscores that marrow and cartilage pathology commonly precede definite radiographic change.

Discussion

In this retrospective series of 35 knees, KL radiographic grade correlated only weakly and non-significantly with a composite MRI structural severity grade (Spearman's $\rho = 0.05$; weighted $\kappa = 0.06$), and MRI detected structural pathology — most consistently cartilage thinning or loss — in virtually every knee examined, including all knees that were radiographically normal (KL 0). This pattern is broadly consistent with the physiological sequence of joint degeneration

described in the literature: cartilage and meniscal changes are now recognised to precede radiographically visible osteophyte formation and joint-space narrowing, sometimes by a considerable margin^{4,8,11}. A recent longitudinal analysis of the Multicenter Osteoarthritis Study cohort similarly found that cartilage loss in the central femoral compartment develops even in KL grade 0 knees, before any radiographic change is apparent, while meniscal extrusion and tears were the strongest predictors of subsequent radiographic progression¹¹. Our cross-sectional findings — cartilage abnormality in 100% of KL-0 knees and meniscal abnormality in two-thirds of them — are directly in keeping with this observation, and support the view that plain radiography systematically under-detects early structural OA^{4,11}.

Our findings of weak KL–MRI correlation contrast with some earlier reports describing a strong or "very high" correlation between a dedicated four-point MRI cartilage-based grading system and KL grade⁷, and with studies showing significant, if moderate, associations between MRI-measured cartilage volume and KL grade [9,10]. Several factors likely explain this discordance. First, those studies typically used formally validated, single-domain MRI scoring instruments applied prospectively by trained readers with documented inter- and intra-observer reliability⁵⁻⁷, whereas the present analysis derived a composite severity grade retrospectively from free-text radiology reports — a source of measurement variability. Second, our sample was younger (mean 48.4 years) than most published cohorts correlating KL and MRI grade (often >55–60 years)^{7,8}, and no knee in our series reached KL grade 3 or 4; restriction of range at the severe end of the KL scale is a well-recognised cause of attenuated correlation coefficients. Third, bone marrow lesions and meniscal pathology — both heavily weighted in our composite MRI grade — are known to correlate poorly with the purely bone-based KL system, since KL grading does not directly assess either feature³⁻⁶.

The practical implication of these findings is that, in a population presenting relatively early with knee symptoms, a normal or low KL grade should not be taken as evidence against clinically significant intra-articular structural pathology. Clinicians relying on radiographs alone risk under-staging disease, particularly with respect to cartilage and meniscal status, both of which carry prognostic and therapeutic relevance — for example, in counselling patients about the likely benefit of intra-articular or regenerative therapies, or in selecting candidates for cartilage-preservation procedures, where structural eligibility is increasingly defined on MRI rather than radiographic grounds^{4,13}. Conversely, MRI's greater sensitivity comes with recognised trade-offs in cost, availability and access^{4,13}, meaning that radiography will remain the first-line, pragmatic screening tool in most settings; the challenge, echoed in existing literature, is to identify which radiographic features (e.g., early tibial osteophyte formation) may act as accessible surrogate markers that trigger further MRI evaluation when clinically indicated¹¹.

This study has several limitations. Its retrospective design and modest sample size ($n = 35$) limit statistical power and the precision of the correlation estimates; the wide confidence intervals implied by the small sample mean these findings should be regarded as hypothesis-generating rather than definitive. The absence of any KL grade 3–4 knee in this cohort restricts the range over which correlation could be assessed and likely contributed to the weak observed association. The MRI structural severity grade was derived retrospectively from descriptive reports rather than from a prospectively applied, validated semi-quantitative instrument (e.g., WOMBS, MOAKS, or the Park et al. system)⁵⁻⁷ with documented inter-observer reliability, introducing potential misclassification; prospective studies applying a validated MRI scoring

tool with masked, independent readers are needed to confirm these findings. Finally, as a single-centre study, generalisability to other populations and imaging protocols may be limited.

Conclusion

In this retrospective comparative study of 35 knees, Kellgren–Lawrence radiographic grading correlated poorly with MRI-detected structural severity, and MRI identified cartilage, meniscal and marrow abnormalities even in knees graded as radiographically normal. These findings reinforce that radiography alone, while remaining a practical first-line tool, underestimates the true burden of structural change in knee osteoarthritis, and support selective use of MRI — particularly in patients with symptoms disproportionate to their radiographic grade — to avoid under-staging of disease and to better inform clinical and therapeutic decision-making. Larger, prospective studies using validated semi-quantitative MRI scoring systems are warranted to confirm and extend these observations.

References

1. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *Ann Rheum Dis.* 1957;16(4):494-502.
2. Kohn MD, Sassoon AA, Fernando ND. Classifications in brief: Kellgren-Lawrence classification of osteoarthritis. *Clin Orthop Relat Res.* 2016;474(8):1886-93.
3. Schiphof D, Boers M, Bierma-Zeinstra SM. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. *Ann Rheum Dis.* 2008;67(7):1034-6.
4. Culvenor AG, Engen CN, Øiestad BE, Engebretsen L, Risberg MA. Defining the presence of radiographic knee osteoarthritis: a comparison between the Kellgren and Lawrence system and OARSI atlas criteria. *Knee Surg Sports Traumatol Arthrosc.* 2015;23(12):3532-9.
5. Peterfy CG, Guermazi A, Zaim S, Tirman PFJ, Miaux Y, White D, et al. Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis. *Osteoarthritis Cartilage.* 2004;12(3):177-90.
6. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). *Osteoarthritis Cartilage.* 2011;19(8):990-1002.
7. Park HJ, Kim SS, Lee SY, et al. A practical MRI grading system for osteoarthritis of the knee: association with Kellgren-Lawrence radiographic scores. *Eur J Radiol.* 2013;82(1):112-7.
8. Hayes CW, Jamadar DA, Welch GW, Jannausch ML, Lachance LL, Capul DC, et al. Osteoarthritis of the knee: comparison of MR imaging findings with radiographic severity measurements and pain in middle-aged women. *Radiology.* 2005;237(3):998-1007.
9. Cicuttini FM, Wluka AE, Forbes A, Wolfe R. Comparison of tibial cartilage volume and radiologic grade of the tibiofemoral joint. *Arthritis Rheum.* 2003;48(3):682-8.
10. Harada Y, Tokuda O, Fukuda K, et al. Relationship between cartilage volume using MRI and Kellgren-Lawrence radiographic score in knee osteoarthritis with and without meniscal tears. *AJR Am J Roentgenol.* 2011;196(3):298-304.

11. Lee DW, Kim JS, Han HS, Ro DH. Longitudinal correlation between X-ray and MRI findings in medial compartment knee osteoarthritis: insights into early cartilage loss and structural changes. *Knee Surg Sports Traumatol Arthrosc.* 2026;34(6):2047-56.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-74.
13. Guermazi A, Hunter DJ, Roemer FW. Plain radiography and magnetic resonance imaging diagnostics in osteoarthritis: validated staging and scoring. *J Bone Joint Surg Am.* 2009;91 Suppl 1:54-62.
14. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage.* 2007;15 Suppl A: A1-56.
15. Favero M, Ramonda R, Goldring MB, Goldring SR, Punzi L. Early knee osteoarthritis. *RMD Open.* 2015;1(1):e000062.