

CONCORDANCE AND DISCORDANCE BETWEEN RADIOLOGY RESIDENTS AND CONSULTANT RADIOLOGISTS IN INTERPRETATION OF KNEE RADIOGRAPHS IN PATIENTS WITH KNEE PAIN

Samee Ur Rehman¹

How to cite this article

Rehman SU. Concordance And Discordance Between Radiology Residents and Consultant Radiologists in Interpretation of Knee Radiographs in Patients with Knee Pain. J Gandhara Med Dent Sci. 2026;13(1):101-105. <https://doi.org/10.37762/jgmds.13-1.741>

Date of Submission: 13-10-2025

Date Revised: 19-12-2025

Date Acceptance: 22-12-2025

Correspondence

Samee Ur Rehman, Trainee Medical Officer, Department of Radiology, Rehman Medical Institute, Peshawar

☎: +92-333-5829226

✉: samirehman6969@gmail.com

ABSTRACT OBJECTIVES

To evaluate the concordance and discordance between radiology residents and consultant radiologists in the interpretation of knee radiographs in patients presenting with knee pain, and to identify areas of clinically significant diagnostic discrepancy.

METHODOLOGY

This prospective observational study included 203 patients who underwent knee radiographs over two years. Four radiology residents independently interpreted all images, followed by a blinded review by four consultant radiologists with at least 5 years' post-qualification experience. Diagnostic categories assessed were normal findings, osteoarthritis, fractures, joint effusion, osteopenia, and bone/soft-tissue tumors. Concordance was defined as identical categorization between residents and consultants. Agreement was analyzed using Cohen's kappa coefficient, and discordant cases were further classified by type and frequency.

RESULTS

Out of the 203 patients, 113 (55.7%) were female, and 90 (44.3%) were male, with a mean age of 49.3 ± 17.1 years. Overall, 135 radiographs (66.5%) demonstrated abnormalities, with osteoarthritis being most frequent ($n=118$, 58.1%). Concordant interpretations occurred in 155 cases (76.4%), while 48 cases (23.6%) were discordant. Cohen's kappa for overall agreement was 0.629 ($p < 0.001$), indicating substantial concordance. Most discordance occurred in fracture detection and in residents' overdiagnosis of osteoarthritis. Less frequent findings such as joint effusion, osteopenia, and bone/soft-tissue tumors were also slightly under-reported by trainees.

CONCLUSION

Radiology residents show substantial agreement with consultants in interpreting knee radiographs; however, discrepancies remain in detecting fractures and rarer pathologies. Targeted training, structured supervision, and systematic review of interpretations are recommended to enhance diagnostic accuracy and patient safety.

KEYWORDS: Knee Radiographs; Diagnostic Concordance; Radiology Resident; Consultant Radiologist; Osteoarthritis; Fracture; Musculoskeletal Imaging

INTRODUCTION

Knee pain is one of the most common musculoskeletal complaints that occurs in clinical practice and is a primary source of disability at all ages. It comprises a significant percentage of outpatient visits, emergency department visits, and orthopedic referrals in the global context, with acute traumatic injuries and chronic degenerative conditions being the causes of the disease burden.¹ Clinical and socioeconomic implications of disorders in the knee have been further augmented by the increasing prevalence of obesity, extended lifespan, and the growing engagement in sports and occupational activities, among others.² Plain radiography is the most suitable imaging modality for assessing knee pain, as it

is widely available, inexpensive, and can detect a wide range of bone abnormalities. Knee radiographs are performed regularly to detect fractures, loss of joint space, osteophytes, malabsorption, and indirect evidence of internal rupture.³ Although cross-sectional imaging has improved, radiographs still have a central role in initial diagnosis and triage, and they determine subsequent imaging and management. Knee radiograph interpretation is critical because misinterpretation can lead to missed diagnoses, misdirected therapy, and poor patient outcomes. The fractures, minor cortical disturbances, and initial degenerative alterations can be difficult to detect, particularly when overlapping structures or poor image quality are present.⁴ Diagnostic mistakes in musculoskeletal imaging have

been identified as a source of medicolegal claims and patient safety issues, and the need to focus on interpretive accuracy has been emphasized.⁵ Radiology resident participation in the initial interpretation of imaging studies is a common role in teaching hospitals, where they are often under time pressure and with varying degrees of supervision. Although residency training is designed to develop diagnostic competence over time, discrepancies in experience and exposure can affect interpretive performance. Previous research has established that differences between trainees and senior radiologists are not exceptional, especially in musculoskeletal imaging, where subtle findings can be missed or overreported.⁶ Interobserver variability in radiographic interpretation has been well documented, and models include factors such as level of training, case complexity, diagnostic confidence, and cognitive biases. Patients are likely to show a bias to overdiagnose degenerative alterations or underidentify minor traumatic results, whereas consultants depend on experience and pattern recognition to refine the diagnostic judgment.⁷ Knowledge of such trends in concordance and discordance is crucial for enhancing training interventions and patient care. Although some international researchers have studied consensus between radiology trainees and consultants across different imaging modalities, there is limited data on agreement between radiology trainees and consultants in the analysis of knee radiographs using plain radiography, especially in local and regional training environments.⁸ Given the high number of knee radiographs performed in everyday practice, even small diagnostic discrepancy rates can have clinical consequences. Assessment of diagnostic concordance between resident and consultant radiologists can help identify the root causes of errors, implement targeted educational interventions, and improve supervision in radiology training programs. These types of assessments also play a role in quality assurance activities and help develop systematic feedback systems intended to enhance interpretation accuracy and patient safety. This research aimed to determine the extent of agreement and disagreement between radiology residents and consultant radiologists in interpreting knee radiographs in patients with knee pain, and to identify clinically significant areas of diagnostic disagreement.

METHODOLOGY

This is an observational-analytical research study conducted at the Department of Radiology, Rehman Medical Institute, Peshawar, Pakistan, over 2 years from November 2022 to November 2024. The research aimed to test diagnostic consensus between radiology residents and consultant radiologists in interpreting

knee radiographs of patients presenting with knee pain. Prospective data collection was used, and all participants were measured at a single time point. Non-probability consecutive sampling was used to incorporate 203 patients who had knee radiography during the study period. All patients with referred plain knee radiographs and knee pain were eligible to participate. Radiographs of poor image quality, missing views, and those from patients who had undergone previous knee surgery were excluded to reduce interpretive bias. Each patient in the analysis provided one set of knee radiographs. Each knee radiograph consisted of anteroposterior and lateral images, obtained with the digital radiography systems available at the institution (manufacturer information was omitted because the imaging had to be performed with standard hospital equipment). Four radiology residents, each registered in an accredited residency training program and with daily exposure to musculoskeletal imaging in the clinical environment, were the first to interpret radiographs independently. Four consultant radiologists then reviewed the same radiographs with a minimum of 5 years of post-qualification experience in diagnostic radiology. Consultants did not see what the inhabitants interpreted and what they had preliminarily reported. Routine clinical referral information accompanying the imaging request was available to both residents and consultants, as is typical in clinical practice. Radiographic results were divided into the predetermined diagnostic categories: normal results, osteoarthritis, fracture, joint effusion, osteopenia, and bone or soft-tissue tumor. These were chosen categories that typify clinically significant knee pathologies in daily practice. Diagnostic concordance was stipulated as the same categorization of the results between residents and consultant radiologists. The reference standard of agreement analysis was consultant radiologist interpretation, as done in previous interobserver studies. Demographic information about the patient, such as age and gender, and radiographic diagnostic results were recorded in a structured Microsoft Excel data sheet. Cases with discordance were also reviewed to determine the nature and clinical significance of the discrepancies. The primary outcome measure was agreement between the residents' and the consultant's interpretations. The study was approved by the Research Ethics Committee of Rehman Medical Institute (IRBNo.RMI/RMI-REC/ArticleApproval/130). Patient confidentiality was maintained throughout the study, and all data were anonymized prior to analysis. Statistical analysis was performed using Statistical Package for the Social Sciences (SPSS) version 22 (IBM Corp., Armonk, NY, USA). Descriptive statistics were used to summarize demographic variables and diagnostic frequencies. Agreement between residents

and consultants was assessed using Cohen's kappa coefficient, with kappa values interpreted according to standard benchmarks. A p-value of less than 0.05 was considered statistically significant.

RESULTS

Over the course of the trial, 203 knee radiographs were assessed. There were 90 males (44.3%) and 113 females (55.7%) in the study population. The majority of patients were aged 51-70, with a mean age of 49.3 ± 17.1 years. Based on the consultant radiologist's assessment, abnormal radiographic findings were identified in 135 cases (66.5%), while 68 radiographs (33.5%) were reported as usual. Among abnormal findings, degenerative changes consistent with osteoarthritis were the most frequently observed pathology. Other findings, including fractures, joint effusion, osteopenia, and bone or soft-tissue tumors, were less frequently identified. When resident and consultant interpretations were compared, agreement was the rule. A smaller subset of radiographs had discordant interpretations. Agreement analysis using Cohen's kappa coefficient demonstrated a value of 0.629 ($p < 0.001$), indicating substantial agreement between radiology residents and consultant radiologists.

Table 1: Age and Gender Distribution of Study Population (n = 203)

Variable	Category	n (%)
Gender	Male	90 (44.3)
	Female	113 (55.7)
Age (years)	≤30	30 (14.8)
	31-50	59 (29.1)
	51-70	96 (47.3)
	>70	18 (8.8)

Table 2: Overall Radiographic Outcome Based on Consultant Interpretation

Radiographic Outcome	Frequency	%age
Normal	68	33.5
Abnormal	135	66.5

Table 3: Spectrum of Abnormal Knee Radiographic Findings (Consultant-Based, n=135)

Radiographic Diagnosis	Frequency	Percentage of Abnormal Cases (%)
Osteoarthritis	118	87.4
Joint effusion	08	5.9
Fracture	03	2.2
Osteopenia	02	1.5
Bone/soft-tissue tumor	04	3.0

Table 4: Agreement Status Between Residents and Consultants (n=203)

Interpretation Outcome	Frequency
Concordant	155
Discordant	48

Table 5: Diagnostic Distribution Reported by Residents and Consultants (n = 203)

Diagnosis Category	Residents (n)	Consultants (n)
Normal	68	66
Osteoarthritis	129	118
Fracture	02	03
Joint effusion	07	08
Osteopenia	02	02
Bone/soft-tissue tumor	01	04

DISCUSSION

This study found that radiology residents and consultant radiologists showed a high correlation in interpreting knee radiographs, with a Cohen's kappa of 0.629 and an overall concordance rate of 76.4%.⁹ These findings suggest that radiology residents can recognize prevalent knee pathologies with reasonable reliability, especially when regular radiographic images are available. Osteoarthritis was the most commonly observed pathology, as previously reported in the literature, where the prevalence of degenerative changes on knee radiographs is high.¹⁰ Residents were more prone to overdiagnosing osteoarthritis than consultants, and this trend has been observed in other studies, with less experienced readers having a greater sensitivity and lower specificity to a common degenerative appearance.¹¹ This implies that there should be systematic training in distinguishing age-related changes from clinically significant osteoarthritis. Detection of fractures is one of the most critical sources of discord, as it directly affects patient management. This research demonstrated that residents miss a few fractures.¹² Other reports on similar patterns have indicated that subtle fractures and cortical irregularities are easily missed by trainees elsewhere.¹³ These findings highlight the necessity of special educational treatment and the systematic examination of trauma radiographs. The incidence of joint effusion and the less frequent occurrence of osteopenia and bone/soft-tissue tumors were rare in both groups.¹⁴ Residents slightly underreported these findings, consistent with previous studies reporting lower detection of rare pathologies.¹⁵ Subsequently, more knee conditions could be recognized better by exposing them to a broader range of conditions. The causes of diagnostic discordance include level of experience, case complexity, and image quality.¹⁶ Training biases, such as anchoring on degenerative changes and the satisfaction of search, could also affect trainee performance.¹⁷ Knowledge of these factors can inform the development of teaching interventions that could minimize interpretive error and enhance diagnostic confidence. Such research differs from previous studies in that it focused on routine clinical practice, in which residents interpret images

under supervision with standard referral information.¹⁸ This method is based on real-life conditions, unlike simulated test conditions or high-tech radiography research, thereby providing an idea of how diagnostic performance is in real-life situations for knee radiographs. Consultant interpretation is not a gold standard, and inter-consultant variability cannot be formally measured using a reference standard common to agreement studies.^{12,13,14,15,16,17,18,19}

LIMITATIONS

The research was conducted at one institution, and this does not mean it can be generalized to other training settings. The sample size of 203 patients is not sufficient to address rare pathologies. The experience of residents and the year of training were not considered, which could influence diagnostic performance. A reference standard was taken to be consultant interpretations, with no formal evaluation of inter-consultant variability. It was not correlated with high-resolution imaging modalities such as MRI or CT. Diagnostic confidence was not measured, so its effects on accuracy were not evaluated. Lastly, rare findings were limited to interpretation due to their low case count.

CONCLUSIONS

Radiology residents agreed significantly with consultant radiologists in interpreting knee radiographs, especially when considering the most common pathologies, such as osteoarthritis. Nonetheless, fractures and rarer identifications remained discordant, indicating that specific training, systematic supervision, and feedback should be provided. The accuracy of diagnosis, the reduction of clinically significant errors, and, finally, patient safety could be improved through enhanced education strategies and systematic review of residents' interpretations.

CONFLICT OF INTEREST: None

FUNDING SOURCES: None

REFERENCES

- Brejnebo MW, Lenskjold A, Ziegeler K, Ruitenbeek H, Müller FC, Nybing JU, et al. Interobserver agreement and performance of concurrent AI assistance for radiographic evaluation of knee osteoarthritis. *Radiology*. 2024;312(1):e233341. doi:10.1148/radiol.233341.
- Huhtanen JT, Nyman M, Sequeiros RB, Koskinen SK, Pudas TK, Kajander S, et al. Discrepancies between radiology specialists and residents in fracture detection from musculoskeletal radiographs. *Diagnostics* (Basel). 2023;13(20):3207. doi:10.3390/diagnostics13203207.
- Cain G, Pittock LJ, Piper K, Venumbaka MR, Bodoceanu M. Agreement in the reporting of general practitioner requested musculoskeletal radiographs: reporting radiographers and consultant radiologists compared with an index radiologist. *Radiography* (Lond). 2022;28(2):288–95. doi:10.1016/j.radi.2021.12.004.
- Seow D, Yasui Y, Chan L, et al. Inconsistent radiographic diagnostic criteria for Lisfranc injuries: a systematic review. *BMC Musculoskelet Disord*. 2023;24:915. doi:10.1186/s12891-023-07043-z.
- Shahzad M, Zulfiqar T, Ali A. Radiographic evaluation of knee joint in patients with knee pain and its correlation with osteoarthritis and gender. *Adv Life Sci*. 2022;9(3):309–12.
- Brejnebo MW, Hansen P, Nybing JU, Bachmann R, Ratjen U, Hansen IV, et al. External validation of an artificial intelligence tool for radiographic knee osteoarthritis severity classification. *Eur J Radiol*. 2022;150:110249. doi:10.1016/j.ejrad.2022.110249.
- Pham HH, Nguyen HQ, Nguyen HT, Le LT, Lam K. Evaluating the impact of an explainable machine learning system on the interobserver agreement in chest radiograph interpretation. *arXiv [Preprint]*. 2023 Apr 1. arXiv:2304.01220.
- Zhang J, Santos C, Park C, Mazurowski MA, Colglazier R. Improving image classification of knee radiographs: an automated image labeling approach. *J Digit Imaging*. 2023;36(6):2402–10. doi:10.1007/s10278-023-00897-1.
- Guria A, Chandra V, Bej B, Vimal K, Barnwal RK, Singh US, et al. Clinical and radiological evaluation of osteoarthritis in knee pain patients and its association with inflammatory markers at MGM Medical College Hospital, Jamshedpur. *J Family Med Prim Care*. 2025;14(6):2242–50. doi:10.4103/jfmpe.jfmpe_365_25.
- Parsons C, Fuggle N, Edwards MH, et al. Concordance between clinical and radiographic evaluations of knee osteoarthritis. *Aging Clin Exp Res*. 2017;30(1):17–25. doi:10.1007/s40520-017-0845-0.
- Haas R, Gorelik A, O'Connor DA, Pearce C, Mazza D, Buchbinder R. Patterns of imaging requests by general practitioners for people with musculoskeletal complaints: an analysis from a primary care database. *Arthritis Care Res* (Hoboken). 2025;77(3):402–11. doi:10.1002/acr.25302.
- Eckersley T, Faulkner J, Al-Dadah O. Inter- and intra-observer reliability of radiological grading systems for knee osteoarthritis. *Skeletal Radiol*. 2021;50(10):2069–78. doi:10.1007/s00256-021-03749-3.
- Smolle MA, Goetz C, Maurer D, Vielgut I, Novak M, Zier G, et al. Artificial intelligence-based computer-aided system for knee osteoarthritis assessment increases experienced orthopaedic surgeons' agreement rate and accuracy. *Knee Surg Sports Traumatol Arthrosc*. 2023;31(3):1053–62. doi:10.1007/s00167-022-07313-1.
- Yasa Y, Çoban D, Özbey F, Tuna T, Yılmaz BN, Erzurumlu ZÜ, Sadık E. Evaluation of common diagnostic errors in panoramic radiographs and interobserver agreement in error identification. *Med Records*. 2025;7(1):114–9. doi:10.5455/medrec.2025.114.
- Borotikar B, Lempereur M, Lelievre M, Burdin V, Ben Salem D, Brochard S. Dynamic MRI to quantify musculoskeletal motion: a systematic review of concurrent validity and reliability, and perspectives for evaluation of musculoskeletal disorders. *PLoS One*. 2017;12(12):e0189587. doi:10.1371/journal.pone.0189587.
- Harkey P, Duszak R Jr, Gyftopoulos S, Rosenkrantz AB. Who refers musculoskeletal extremity imaging examinations to radiologists? *Am J Roentgenol*. 2018;210(4):834–41. doi:10.2214/AJR.17.19079.
- Meredith S, McBurnie A, Tavete I, Edwards C. Radiographer-clinician agreement on imaging selection for extremity injuries in the emergency department: a pilot study. *Radiography* (Lond). 2025;31(6):103166. doi:10.1016/j.radi.2025.103166.

18. Parsons C, Clynes M, Syddall H, Jagannath D, Litwic A, van der Pas S, et al. How well do radiographic, clinical and self-reported diagnoses of knee osteoarthritis agree? Findings from the Hertfordshire cohort study. SpringerPlus. 2015;4:177. doi:10.1186/s40064-015-0962-0.
19. Benchoufi M, Matzner-Lober E, Molinari N, Jannot AS, Soyer P. Interobserver agreement issues in radiology. Diagn Interv Imaging. 2020;101(10):639–41. doi:10.1016/j.diii.2020.07.005.

AUTHORS CONTRIBUTION

Samee Ur Rehman - Concept & Design; Data Acquisition; Data Analysis/Interpretation; Drafting Manuscript; Critical Revision; Supervision; Final Approval

The authors accept responsibility for all aspects of the work and will ensure that any concerns regarding the accuracy or integrity of any part are properly investigated and resolved.



LICENSE: JGMDS publishes its articles under a Creative Commons Attribution Non-Commercial Share-Alike license (CC-BY-NC-SA 4.0).

COPYRIGHTS: Authors retain the rights without any restrictions to freely download, print, share and disseminate the article for any lawful purpose.

It includes scholarly networks such as Research Gate, Google Scholar, LinkedIn, Academia.edu, Twitter, and other academic or professional networking sites.