ORIGINAL ARTICLE **DOI:** https://doi.org/10.18502/fem.v10i1.22033

Diagnostic accuracy of inverted grayscale radiography for detection of tibial plateau fractures

Shakiba Mohammadi[1], Mohammadreza Mohammadi[1], Mohammad Afzalimoghaddam[1], Mohammad Mahdi Pirmoradi[1], Sara Parviz[2], Hadi Mirfazaelian[1,3]*

1. Emergency Medicine Department, Tehran University of Medical Sciences, Tehran, Iran.
2. Radiology Department, Tehran University of Medical Sciences, Tehran, Iran.
3. Prehospital and Hospital Emergency Research Center, Tehran University of Medical Sciences, Tehran, Iran.

***Corresponding author:** Hadi Mirfazaelian; **Email:** Mirfazaelian@sina.tums.ac.ir

Published online: 2026-02-22

Abstract: **Objective:** This study aimed to quantitatively evaluate the diagnostic performance of inverted grayscale radiography for detecting tibial plateau fractures.

Methods: In this retrospective diagnostic accuracy study, two emergency physicians independently assessed one-view (anteroposterior) knee radiographs in both conventional and inverted grayscale modes in separate, blinded sessions. Spiral knee computed tomography (CT) scans served as the reference standard for fracture confirmation. Sensitivity, specificity, and area under the receiver operating characteristic curve values were calculated. Inter-observer agreement was also assessed using Cohen's kappa.

Results: Of 112 included cases, 31 (27.67%) had tibial plateau fractures confirmed on CT. For the first physician, sensitivity and specificity were 54.84% (CI 95%: 36.00%,72.70%) and 85.19% (CI 95%: 75.60%,92.10%) with conventional images, compared with 35.48% (CI 95%: 19.20%,54.60%) and 83.95% (CI 95%: 74.10%,91.20%) with inverted grayscale images. For the second physician, sensitivity was identical for both modalities (38.71% (CI 95%: 21.90%,57.80%)) while specificity was higher with conventional radiography (100.00% (CI 95%: 95.60%,100.00%)) than inverted grayscale radiography (90.12% (CI 95%: 81.50%,95.60%)). The area under the operating characteristic curve values were consistently higher for conventional imaging: 0.70 (CI 95%: 0.60,0.78) versus 0.59 (CI 95%: 0.50,0.68) for the first physician (P-value=0.03) and 0.69 (0.57–0.81, CI 95%) versus 0.64 (CI 95%: 0.52,0.76) for the second physician (P-value=0.18). Inter-observer agreement was moderate for conventional radiographs ($\kappa = 0.51$) and lower for inverted grayscale images ($\kappa = 0.38$).

Conclusion: Inverted grayscale radiography did not improve the detection of tibial plateau fractures and was associated with reduced inter observer reliability.

Keywords: Diagnostic Accuracy; Emergency Physician; Grayscale Inversion Radiography; Image Interpretation; Tibial Plateau Fracture

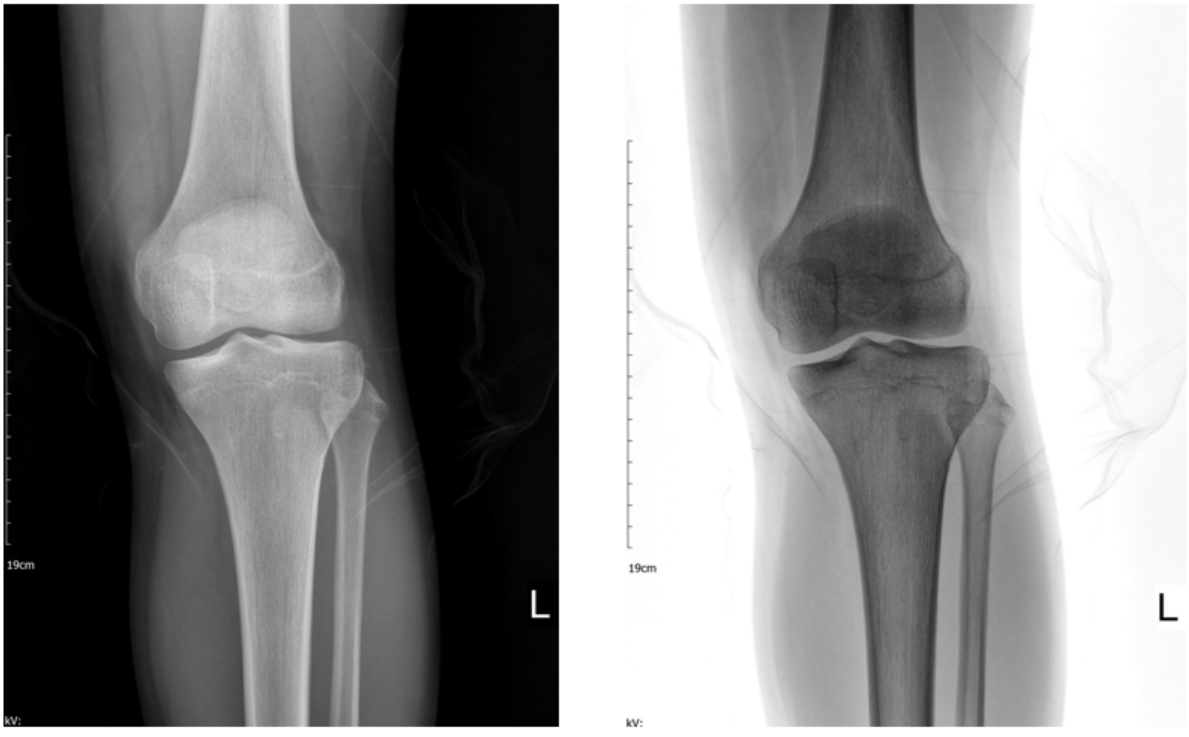
Cite this article as: Mohammadi S, Mohammadi M, Afzalimoghaddam M, Pirmoradi MM, Parviz S, Mirfazaelian H, et al. Diagnostic accuracy of inverted grayscale radiography for detection of tibial plateau fractures. Front Emerg Med. 2026;10(1):e5

1. Introduction

The tibiofemoral joint experiences various forces during daily activities, each influencing its structural integrity. Tibial plateau fractures, which represent approximately 1% of all fractures, typically result from high energy trauma and may involve adjacent structures, including blood vessels, nerves, ligaments, and menisci (1). Diagnosis is primarily based on clinical evaluation supported by standard radiographic imaging (2,3). However, because missed diagnoses are relatively common and may lead to adverse clinical outcomes, computed tomography (CT) imaging is recommended when uncertainty persists, as it provides superior visualization of bone morphology and fracture characteristics (4).

In healthcare settings where CT imaging is unavailable-such as rural or resource-limited hospitals-alternative diagnostic strategies may assist clinical decision-making (5,6). Inverted grayscale radiography is an image processing technique that reverses pixel intensity values across the entire grayscale spectrum. It has been proposed that grayscale inversion may enhance radiographic interpretation because human visual perception can sometimes detect density variations more effectively in inverted-contrast images, potentially improving the delineation of osseous abnormalities (7–9).

Although previous studies have evaluated inverted grayscale radiography either alone or alongside conventional radiography, their findings regarding its diagnostic value have been inconsistent (7). Furthermore, its diagnostic performance for knee fractures has not yet been investigated. This study aimed to quantitatively assess the diagnostic accuracy of inverted grayscale radiography for detecting tibial plateau fractures.



a. Conventional knee x-ray b. Inverted grayscale knee x-ray

Figure 1 Anteroposterior knee conventional and inverted grayscale X-ray images

Table 1 The diagnostic performance indices of emergency physicians using conventional and inverted images

Reader & modality	TP	FP	TN	FN	Sensitivity (CI 95%)	Specificity (CI 95%)	+LR	-LR	AUC (CI 95%)
Physician 1 Conventional	17	12	69	14	54.84% (36.00,72.70)	85.19% (75.60,92.10)	3.70 (2.01,6.83)	0.53 (0.36,0.79)	0.70 (0.60,0.78)
Physician 1 Inverted	11	13	68	20	35.48% (19.20,54.60)	83.95% (74.10,91.20)	2.21 (1.11,4.40)	0.77 (0.58,1.01)	0.59 (0.50,0.68)
Physician 2 Conventional	12	0	81	19	38.71% (21.90,57.80)	100.00% (95.60,100.00)	—	0.61 (0.46,0.81)	0.69 (0.57,0.81)
Physician 2 Inverted	12	8	73	19	38.71% (21.90,57.80)	90.12% (81.50,95.60)	3.92 (1.77,8.66)	0.68 (0.51,0.91)	0.64 (0.52,0.76)

AUC: Area Under The Receiver Operating Characteristic Curve; TN: True Negative; FN: False Negative; TP: True Positive; FP: False Positive; +LR: Positive Likelihood Ratio; -LR: Negative Likelihood Ratio. Estimates have been presented along with a CI 95%

2. Methods

This study was conducted in December 2023 at the Imam Khomeini Hospital Complex (IKHC), a tertiary care center affiliated with Tehran University of Medical Sciences. The aim was to evaluate the diagnostic accuracy of inverted grayscale radiography compared with standard radiography for detecting tibial plateau fractures. This study was reported according to the standards for reporting diagnostic accuracy studies (STARD) reporting guidelines (10). Ethical approval was obtained from the institutional review board (IR.TUMS.IKHC.REC.1400.226), and the study was conducted in accordance with the Declaration of Helsinki.

2.1. Design and setting

In this retrospective diagnostic accuracy study, we analyzed imaging data from a randomly selected sample of emergency department (ED) patients who underwent both knee radiography and CT for acute knee trauma, either as an isolated injury or as part of multiple traumas. Imaging decisions were made at the discretion of the attending emergency physician or the trauma team leader. Exclusion criteria included: (1) age under 18 years, (2) open fractures, and (3) incomplete radiologic records. A total of 112 adult cases were reviewed between December 2022 and December 2023. Of these, 31 patients (27.7%) had tibial plateau fractures, while 81 (72.3%) did not. Grayscale inversion was performed using the built in tools of the picture archiving and commu-

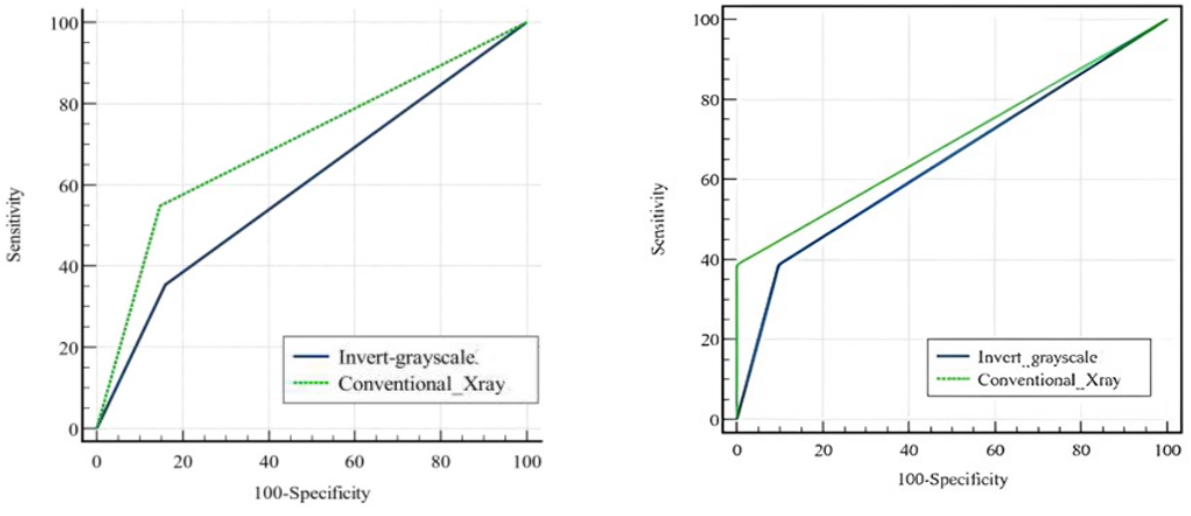


Figure 2 AUROC curves of the tibial plateau fracture detection by conventional and inverted images

Table 2 Comparison of the area under the receiver operating characteristic (AUROC) curve for each emergency physician using conventional and inverted image

Reader	Difference in AUC	P value*
Physician 1 (conventional – inverted)	0.10	0.025
Physician 2 (conventional – inverted)	0.05	0.180

*: Calculated using the DeLong method

nication system (PACS) workstation. Both the conventional and inverted grayscale images were obtained in the anteroposterior (AP) view only due to the availability of data. Images were exported from the PACS in JPG format at 96 dpi, which is the standard anonymized export format supported by our institution. Direct DICOM export was not feasible due to system restrictions on de identification. Although JPG is a compressed format, the PACS export function does not apply additional compression or post processing, and no external editing was performed. The same exported images were used consistently across both modalities and for both readers, ensuring uniform image quality. All images were irreversibly anonymized and displayed on a 21 inch LCD monitor with a resolution of 2048 × 1536 pixels. An example of the conventional and inverted images is provided in figure 1.

2.2. Study protocol and measurements

Two emergency physicians, each with more than 10 years of experience, independently evaluated the conventional and inverted images. Both readers were board certified emergency physicians and were familiar with the grayscale inversion function available in the PACS system. No formal calibration or training was conducted before image interpretation, as the study aimed to reflect real world practice. Intraobserver reliability was not assessed because each modality was interpreted only once per reader. To ensure unbiased assessment, both physicians were blinded to the CT findings. Image interpretation was performed in two separate reading sessions. In the first session, the physicians reviewed randomly ordered, anonymized conventional images and recorded the presence or absence of fractures. To minimize recall bias, the sessions were separated by a one-week washout interval. In the second session, inverted grayscale versions of the same images were presented in a different random sequence for independent reassessment. During both sessions, physicians were permitted to adjust magnification, but were not allowed to make any other image modifications. Breaks were allowed, and no time limits were imposed on the evaluations.

2.3. Outcome

The primary outcome was the diagnostic accuracy of inverted grayscale radiography compared with conventional radiography for detecting tibial plateau fractures. The reference diagnosis for each case was based on the radiologist's interpretation of the spiral knee CT scan. The board-certified radiologist had more than 10 years of experience.

2.4. Statistical analysis

Normally distributed continuous variables were reported as means with standard deviations, and categorical variables as counts and percentages. The area under the receiver operating characteristic (AUROC) curve was used to compare the proportion of discordant patient pairs in which the patient with the fracture had the higher predicted probability. AUROC curves were compared using the non-parametric De-

Long method. Two-by-two contingency tables were used to calculate sensitivity, specificity, and positive and negative likelihood ratios (+LR and -LR). All estimates were reported with 95% confidence intervals (CI). Inter observer agreement between the two physicians was assessed using Cohen's kappa statistic. Sample size estimation was based on detecting a clinically meaningful improvement in sensitivity when using inverted grayscale radiography compared with conventional radiography. Using the sample size formula for comparing two proportions, with a sensitivity of 0.79 for conventional radiography (11), to test for an increase of 0.17 (from 0.79 to 0.9) in sensitivity, using a two sided = 0.05 and 80% power, 112 cases were required using the standard normal approximation. To calculate the diagnostic test performance indices, an online calculator, MedCalc, was used (12). All other statistical analyses were performed using RStudio version 3.0.3.1. A P-value<0.05 was considered statistically significant.

3. Results

The mean age of the study population was 44±15 years, and 72% of the population were male. For the first emergency physician, sensitivity was 54.84% (CI 95%: 36.00,72.70) with conventional radiographs and 35.48% (CI 95%: 19.20,54.60) with inverted grayscale images. For the second physician, sensitivity was identical for both modalities at 38.71% (CI 95%: 21.90,57.80). Specificity for the first physician was 85.2% (CI 95%: 75.6,92.1) with conventional radiographs and 83.95% (CI 95%: 74.10,91.20) with inverted grayscale images. For the second physician, specificity was 100.00% (CI 95%: 95.60,100.00) with conventional radiographs and 90.12% (CI 95%: 81.50,95.60) with inverted grayscale radiographs. The corresponding positive and negative likelihood ratios are summarized in table 1.

For the first physician, the AUROC was 0.70 (CI 95%: 0.60,0.78) with conventional radiographs and 0.59 (CI 95%: 0.50,0.68) with inverted grayscale images. For the second physician, AUROC values were 0.69 (CI 95%: 0.57,0.81) for conventional radiographs and 0.64 (CI 95%: 0.52,0.76) for inverted grayscale radiographs (Figure 2). Comparative analysis demonstrated significantly better diagnostic performance with conventional radiographs for the first physician (P=0.025), whereas no significant difference was observed for the second physician (P=0.18) (Table 2). Inter observer agreement between the two physicians was moderate for conventional radiographs (κ=0.51) and lower for inverted grayscale images (κ=0.38).

4. Discussion

In this study, evaluating the diagnostic accuracy of inverted grayscale radiography compared with standard radiography for detecting tibial plateau fractures on anteroposterior views, we found no improvement in sensitivity or specificity with grayscale inversion. In fact, diagnostic performance tended to decline when physicians interpreted inverted grayscale images.

A wide range of image processing techniques has been developed to enhance diagnostic performance, though their effectiveness varies considerably. These include gray level manipulations (e.g., histogram equalization), spatial corrections (e.g., geometric distortion correction), and preprocessing methods such as edge enhancement and temporal subtraction. Such techniques aim to restore image quality, reduce noise, enhance structural visualization, and facilitate segmentation of clinically relevant features (13). More recently, artificial intelligence–based tools have been explored as additional methods to improve diagnostic accuracy in medical imaging (14–17). Several studies have specifically examined the diagnostic utility of inverted grayscale radiography. Initial work focused on chest radiography for pulmonary nodule detection, followed by multiple subsequent studies evaluating the same pathology (8,18–20). Most of these studies reported improved detection of parenchymal nodules, a finding later supported by a meta analysis (7). However, studies evaluating other thoracic conditions—such as pneumothorax and rib fractures—did not demonstrate any diagnostic advantages (21,22). Looking at other areas, grayscale inversion improved detection of macrocalcifications in mammography (23) but did not enhance measurement accuracy in digital periapical radiographs (24). Several explanations have been proposed for these inconsistent findings. For example, Robinson et al. suggested that increased relative brightness of pulmonary nodules in inverted images may enhance their detectability (18) Conversely, other authors argued that grayscale inversion may impair visual perception and hinder accurate interpretation (21,24). Differences in reader experience may also contribute to the variability in reported results.

Within the musculoskeletal domain, only two studies have evaluated grayscale inversion. Ahmed et al. found that conventional radiography and inverted grayscale imaging were equally effective for the diagnosis of femoral neck fractures (25). Another study found that grayscale inversion improved the consistency of sagittal spinopelvic measurements in asymptomatic adults. This effect was most pronounced when pelvic anatomy was poorly visualized (26). Our study demonstrated lower diagnostic performance for detecting tibial plateau fractures using inverted grayscale radiography, a finding that should be interpreted in the context of the study population and methodological design. Our cohort consisted exclusively of patients who underwent CT imaging, inherently selecting cases with greater diagnostic uncertainty and more subtle radiographic findings. From a research standpoint, this represents a valuable opportunity to evaluate diagnostic performance in challenging cases. ALOS, this study only a single anteroposterior radiographic view was assessed, which might partially explain our findings. Prior study has shown that reducing the number of views from four (AP, lateral, and both obliques) to two (AP

and lateral) can decrease sensitivity from 85% to 79%(27). Additionally, images were interpreted by emergency physicians rather than radiologists, reflecting real world ED practice but potentially contributing to lower sensitivity.

Most prior work on interobserver agreement with grayscale inversion has focused on chest pathologies. Pulmonary nodule detection showed fair agreement, with a slight decrease (from 0.396 to 0.332) in agreement with using inverted grayscale views regardless of radiologists' experience (19). Rib fracture assessments revealed slightly improved agreement (from 0.400 to 0.426) with inverted grayscale views, the highest reliability when both types of views were combined (0.435) (22). In our study, interobserver agreement decreased from moderate with conventional radiographs to fair with grayscale inversion (28). These findings raise important questions regarding the clinical utility of grayscale inversion.

5. Limitations

This study has several limitations. First, the relatively small number of true positive cases resulted in wide confidence intervals for sensitivity estimates; however, sensitivity estimates remain reasonably stable when more than 20–30 true positive cases are available (29). Second, the omission of clinical examination and orthogonal radiographic views—both essential components of comprehensive fracture evaluation—may have affected diagnostic accuracy and limited generalizability. Third, including additional readers and assessing intraobserver reliability would have yielded more precise reproducibility estimates, but this was not feasible. Fourth, because only patients who underwent both radiography and CT were included, probably the sample disproportionately consisted of cases with higher diagnostic uncertainty, rather than reflecting the full spectrum of knee trauma presentations in the emergency department. As a result, the sensitivity and specificity reported here may underestimate routine clinical performance in unselected ED patients. Finally, diagnostic confidence, interpretation time, and the potential benefit of combining conventional and inverted images were not assessed, which may further limit conclusions regarding the practical clinical utility of grayscale inversion.

6. Conclusion

In conclusion, inverted grayscale radiography did not improve diagnostic performance compared with conventional radiography.

7. Declarations

7.1. Acknowledgement

None.

7.2. Authors' contribution

SM: Writing - original draft, Writing - review & editing, Methodology, Data curation, Formal analysis; MM: Writing - original draft, Writing - review & editing, Methodology, Data curation; SP: Conceptualization, Supervision; PP: Data curation and validation; MP: Conceptualization; MA: Validation, Supervision; HM: Conceptualization, Validation, Writing - original draft. All authors reviewed the manuscript.

7.3. Conflict of interest

The authors declare no competing interests.

7.4. Funding

No funding or grants were received for performing this research.

7.5. Ethical approval and consent to participate

The study was approved by the institutional ethical review board at the university (IR.TUMS.IKHC.REC.1400.226) and conducted following the Declaration of Helsinki.

7.6. Informed consent

Written informed consent was obtained from selected individuals.

7.7. Data availability

De-identified case data are available from the corresponding author (H-Mirfazaelian@sina.tums.ac.ir) upon reasonable request.

References

1. Salduz A, Birisik F, Polat G, Bekler B, Bozdag E, Kilicoglu O. The effect of screw thread length on initial stability of Schatzker type 1 tibial plateau fracture fixation: a biomechanical study. J Orthop Surg Res. 2016;11(1):1462.
2. Malik S, Herron T, Mabrouk A, Rosenberg N. Tibial plateau fractures. Orthopedic trauma call for the attending surgeon. 2023. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026.
3. Rudran B, Little C, Wiik A, Logishetty K. Tibial plateau fracture: anatomy, diagnosis and management. Br J Hosp Med (Lond). 2020;81(10):1–9.
4. Milani L, Ferrari S. Importance of CT Scan in predicting the outcomes of tibial plateau fractures: a retrospective study of 216 patients over 10 Years' time. Indian J Orthop. 2021;56(3):377.
5. Bergeron C, Fleet R, Tounkara FK, Lavallée-Bourget I, Turgeon-Pelchat C. Lack of CT scanner in a rural emergency department increases inter-facility transfers: a pilot study. BMC Res Notes. 2017;10(1):772.
6. Lyon M, Sturgis L, Lendermon D, Kuchinski AM, Mueller T, Loeffler P, et al. Rural ED transfers due to lack of radiology services. Am J Emerg Med. 2015;33(11):1630–4.

7. Eftekhari M, Jalili M, Karim A, Doosti–irani A, Mirfazaelian H. Diagnostic accuracy of inverted grayscale mode in radiographs: a systematic review and meta-analysis. Front Emerg Med. 2023;7(3):e29–e29.
8. Lungren MP, Samei E, Barnhart H, McAdams HP, Leder RA, Christensen JD, et al. Gray-scale inversion radiographic display for the detection of pulmonary nodules on chest radiographs. Clin Imaging. 2012;36(5):515–21.
9. BLACKWELL HR. Contrast thresholds of the human eye. J Opt Soc Am. 1946;36(11):624–43.
10. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6(11):e012799.
11. Gray SD, Kaplan PA, Dussault RG, Omary RA, Campbell SE, Chrisman HB, et al. Acute knee trauma: how many plain film views are necessary for the initial examination? Skeletal Radiol. 1997;26(5):298–302.
12. Diagnostic test evaluation | MedCalc [Internet]. [cited 2026 Jun 11]. Available from: https://www.medcalc.org/en/calc/diagnostic_test.php.
13. Armato SG, Van Ginneken B. Anniversary paper: image processing and manipulation through the pages of Medical Physics. Med Phys. 2008;35(10):4488–500.
14. Mohammadi M, Parviz S, Parvaz P, Pirmoradi MM, Afzalimoghaddam M, Mirfazaelian H. Diagnostic performance of ChatGPT in tibial plateau fracture in knee X-ray. Emerg Radiol. 2025;32(1):59–64.
15. Liu P ran, Zhang J yao, Xue M di, Duan Y yu, Hu J lang, Liu S xiang, et al. Artificial intelligence to diagnose tibial plateau fractures: an intelligent assistant for orthopedic physicians. Curr Med Sci. 2021;41(6):1158–64.
16. Bousson V, Attané G, Benoist N, Perronne L, Diallo A, Hadid-Beurrier L, et al. Artificial intelligence for detecting acute fractures in patients admitted to an emergency department: real-life performance of three commercial algorithms. Acad Radiol. 2023;30(10):2118–39.
17. Mert S, Stoerzer P, Brauer J, Fuchs B, Haas-Lützenberger EM, Demmer W, et al. Diagnostic power of ChatGPT 4 in distal radius fracture detection through wrist radiographs. Arch Orthop Trauma Surg. 2024;144(5):2461–7.
18. Robinson JW, Ryan JT, Mcentee MF, Lewis SJ, Evanoff MG, Rainford LA, et al. Grey-scale inversion improves detection of lung nodules [Internet].
19. Kirchner J, Gadek D, Goltz JP, Doroch-Gadek A, Stückradt S, Liermann D, et al. Standard versus inverted digital luminescence radiography in detecting pulmonary nodules: a ROC analysis. Eur J Radiol. 2013;82(10):1799–803.
20. Sheline ME, Brikman I, Epstein DM, Mezrich JL, Kundel HL, Arenson RL. The diagnosis of pulmonary nodules: comparison between standard and inverse digitized images and conventional chest radiographs. AJR Am J Roentgenol. 1989;152(2):261–3.
21. Musalar E, Ekinci S, Ünek O, Arş E, Eren HŞ, Gürses B, et al. Conventional vs invert-grayscale X-ray for diagnosis of pneumothorax in the emergency setting. Am J Emerg Med. 2017;35(9):1217–21.
22. Park JB, Cho YS, Choi HJ. Diagnostic accuracy of the inverted grayscale rib series for detection of rib fracture in minor chest trauma. Am J Emerg Med. 2015;33(4):548–52.
23. Altunkeser A, Körez MK. Usefulness of grayscale inverted images in addition to standard images in digital mammography. BMC Med Imaging. 2017;17(1):1–6.
24. Farhadi N, Shokraneh A, Mehdizadeh M. Effect of contrast inversion enhancement on the accuracy of endodontic file length determination in digital radiography. J Clin Diagn Res. 2015;9(5):ZC102–5.
25. Ahmed M, Saad A, Bani-Khalid A, Sonsale P, Iyengar KP, Botchu R. The role of grey scale inversion imaging (GSII) as a diagnostic tool of neck of femur fractures: is it more effective? Emerg Radiol. 2023;30(4):419–23.
26. Sun W, Zhou J, Qin X, Xu L, Yuan X, Li Y, et al. Grayscale inversion radiographic view provided improved intra- and inter-observer reliabilities in measuring spinopelvic parameters in asymptomatic adult population. BMC Musculoskelet Disord. 2016;17(1):1–6.
27. Gray SD, Kaplan PA, Dussault RG, Omary RA, Campbell SE, Chrisman HB, et al. Acute knee trauma: how many plain film views are necessary for the initial examination? Skeletal Radiol. 1997;26(5):298–302.
28. Byrt T. How good is that agreement? Epidemiology. 1996;7(5):561.
29. Fenn Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med. 1996;3(9):895–900.