BENHA MEDICAL JOURNAL bmj

Print ISSN 1110-208X

Online ISSN 2357-0016

025

Diagnostic Validity and Reliability of Osseous Tumour Reporting and Data System (OT-RADS) in the Management of Primary Bone Tumour

Sherif A. Abd-Elsattar, Ahmed M. Alsowey, Hesham E. El Sheikh, Maha E. Bekheet

Department of Diagnostic and Interventional Radiology Faculty of Medicine; Benha University. Egypt.

Correspondence to: Maha E. Bekheet, Department of Diagnostic and Interventional Radiology Faculty of Medicine; Benha University. Egypt.

Email:

elsaeedmaha93@gmail.com

Received:

Accepted

Abstract:

Purpose: To validate an Osseous Tumour Reporting and Data System (OT-RADS) with the hypothesis that the suggested principles can be used as a trusted guidelines and differentiate benign and malignant bone tumours in a reliable way with a good area under the curve. **Methods:** This multi-reader cross-sectional validation study focused on evaluating osseous tumours using the OT-RADS classification based on magnetic resonance imaging features and expert consensus. The OT-RADS categories ranged from 0 (incomplete imaging) to VI (known biopsy-proven malignancy). Three blinded readers categorized each tumour, and inter-reader agreement was measured using Intra Class Correlation (ICC) and Conger κ. Diagnostic performance was assessed with the area under the receiver operating curve. The system was dichotomized into benign (I–III) and malignant (IV–V) categories to calculate sensitivity and specificity for accurate tumour classification. **Results:** Inter reader agreement for OT-RADS (ICC = 0.913), with confidence interval range (0.892:0.942) and binary distinction of benign versus malignant (κ = 0.58) were good to excellent. The sensitivities, specificities, and area under the receiver operating curve of the readers ranged from 81.3%-87.5%,76.5%-82.3% and 0.844%–0.886% respectively. **Conclusions:** Osseous Tumour Reporting and Data System lexicon has a great reliability and helps in stratification of tumours into benign and malignant categories. It can be prudently used for accurate characterization of bone tumours. The radiologists can use it in a practical way to guide the management of the patients, improve multidisciplinary communications and impacts. Referring consultants considered OT-RADS a helpful classification for clinical decision –making in bone tumours.

Keywords: OT-RADS, ICC, Conger κ, AUC.

Introduction:

To guide separating bony lesions- and because the discrepancy and high variability in bone tumour reporting- a structured bone tumour evaluation template is required, with great benefit to radiologists, and particularly non-musculoskeletal (MSK) radiologists [(1)].

Creation of a consensus document (structured reporting system) that classifies bone tumours in details- was done based on World Health Organization classification by 5 fellowship-trained attending radiologists, and was called Osseous Tumour Reporting and Data System (OT-RADS). We then conducted a three-reader blinded evaluation of osseous tumours. The hypothesis was that this standardized emerged powerful classification system for bone tumours can be used in separation of benign from malignant lesions with good to excellent inter-reader reliability. If successfully validated, it is simple and concise and could be used by musculoskeletal radiologists in a practical way to improve the communications and facilitate understanding the report with added inputs of impact outcomes and recommendations [(2).]

RADS is a known approach done on a systematic way for characterization of lesions and used recently for bone tumour characterization and management with a fair reproducibility allowing lesion stratification in classes of increasing malignancy frequency [(3).]

It has great reliability and ability to separate tumours into benign and malignant categories [(4).] It can be practically used by radiologists to guide management of the patient, improve communications, and impact [(5).]

To make this systematic approach more powerful, it is better to extend in a large field based on objective imaging signs. With large numbers of patients in different age groups and different countries and centres to be multi-centric and widespread, consequently can be the base of a successful RADS. This would help this system to be widely spread among the radiologists-community [(5).]

The categories are as follows: OT-RADS 0-incomplete imaging; OT-RADS I-negative; OT-RADS II-definitely benign; OT-RADS III-probably benign; OT-RADS IV-suspicious for malignancy or indeterminate; OT-RADS V-highly suggestive of malignancy; OT-RADS VI-known biopsy proven malignancy or recurrent malignancy in the tumour bed [(2).]

Regarding to the future direction, there is super advance in software programs aiming to create user-friendly, content-rich interface for reporting in structured way. Recently, artificial intelligence (AI) is being tested for autonomous categorization of bone tumours according to variable imaging features required for pre-existing RADSs with encouraging initial results and its value for decision making in clinical setting [(5).]

Patients and methods

Before the study beginning, all patients have written consents and the study was

approved by the Institutional Review Board (IRB) of Faculty of medicine-Benha University institution (Reference No: MD 3-5-2023) and the study followed.

Study design and population:

This is a single centre prospective study. Approval was obtained from the institutional review board, and all participants provided written informed consent. We applied the ethical concepts during planning for this study. The study was carried out at Benha University Hospitals, Radio-Diagnosis Department during the period from May 2023 to May 2024.

The study included 50 patients with clinically suspected bone tumours, any age and any sex and patients have solitary bone tumours- **were included** in the study. Patients who refused to be included in the study, patients with metastatic bony lesions, patients with morbid obesity, and contraindication to magnetic resonance imaging (eg; metallic implants, pacemakers...) **were excluded**.

Patients were subjected to the following:

A. Clinical assessment:

1. **Full clinical history taking:**

- **Personal data include:** Patient name, age and sex.

- **Complaint and present history include:** painful mass and any primary malignancy or known medical disease.

2. **Clinical examination:** Patients assessed by the emergency physician then redirected to the radiology unit.

B. Radiological assessment:

MDCT examination*:*

Examination Technique:

CT examination for most patients included in this study.

All CT examinations were performed with a 128-channel MDCT scanner (GE128) using the following parameters: detector row configuration, 200-250mAs, 5mm slice thickness, field of view 500-600mm and voxel size 0.98x0.98x5mm.

An initial scout image is taken to position the table for desired axial coverage. The CT continuous scan acquisition is performed during breath hold and last approximately 10-20s depending on the axial coverage.

If patient injected with IV contrast material (non-ionic iodinated contrast Ultravist 300, 80-120ml) using an automatic injection with injection rate 3 ml/sec.

Post processing, the scans were reconstructed and reviewed. Multi-planar reconstructions (MRP) were acquired using the machine software in coronal and sagittal planes.

MRI study technique

The MRI was performed with a 1.5 -T superconductive magnet (Achieva, Philips Medical System) using a surface coil with respiratory triggering. Before acquisition of functional DWI sequence, a morphological study of the upper abdomen with four sequences was acquired:

- Axial T1-weighted turbo field echo (TFE)
- Axial T2-weighted single-shot turbo spin echo (TSE)
- Axial T2-weighted single-shot TSE with fat suppression [spectral Selection attenuated inversion recovery (SPAIR)]
- Axial T1 –weighted single-shot spin echo with fat sat after IV Gadolinium contrast media injection (T1W SE FS Gd-CM). With focus on maintaining repetition time and echo time on T1WI before and after Gd-CM.

The study was performed using 3D isotropic fast-field echo (FFE) T1-weighted sequences with SPAIR fat-signal suppression in the axial plane after intravenous administration of 0.10mmol/kg gadolinium-diethylenetriamine penta acetic acid (Gd-DTPA) contrast material with an injection rate of 2.0 ml/s, followed by 20 cc of saline solution with the same injection rate.

Histopathological correlation

The final diagnosis was obtained from the pathological results of the surgical specimen in patients with operable masses. Correlation of the findings obtained using CT, CE- MRI.

Statistical Design

Software Used: IBM SPSS 23.0 for Windows (SPSS Inc., Chicago, IL, USA).

Data Categorization:

Qualitative Data:

- **Number of each observation or sample size (n)- Percentage (%).**

Quantitative Data:

- **Mean-Median-Standard Deviation (SD)-Inter-Quartile Range (IQR)-Range**.

Data Analysis

Significance Testing:

- **P-value**: Used to determine statistical significance:
- **P > 0.05**: non-significant.
- **P ≤ 0.05**: Significant.
- **Logistic Regression**: Used to predict the presence or absence of an outcome based on independent variables. It's suited for categorical dependent variables.
- **Kappa**: Measures the level of agreement between two sets of observations. Interpretation of Kappa values:

- **≤ 0**: No agreement.
- **0.01–0.20**: None to slight agreement.
- **0.21–0.40**: Fair agreement.
- **0.41–0.60**: Moderate agreement.
- **0.61–0.80**: Substantial agreement.
- **0.81–1.00**: Almost perfect agreement.

Results

The current study enrolled 50 patients with clinically diagnosed suspected bone tumour. We successfully performed all CT and MRI examination without any side effects. Based on the results, we have 34 benign and 16 malignant osseous tumours. On patient - based analysis, the diagnostic performance for predicting significant OT-RADS, the sensitivity, specificity, PPV, NPP, and AUC depending on the reviewers are summarized in **Table (1), Figure (1).**

We analysed the data set of the diagnostic performance of OT-RADS to determine the cut-off value for separation of benign from malignant tumours depending on the reviewers (Fig 1). Based on ROC analysis, all reviewers agreed that the optimal cut-off value for predicting classify osseous tumour was > or =OT-RADS 4. The use of this cut-off value was associated with AUC range from 0.844-0.886 (CI range between 0.854-0.942, *p* value < 0.001)

The inter-reviewer agreement (IRA) for OT-RADS categories results and the sub-analysis of IRA by categories are presented in(**Table (2)).**

Case Presentation:

Case 1:

Clinical history: 50-year-old male did MRI due to knee pain.

MRI revealed: MRI of the RT leg revealed: Rather defined oblong shaped exophytic osseous lesion seen at the anterior border of the upper third of the tibia. The lesion is seen indenting the cortex with possible disruption. Complementary CT was done and revealed that lytic with calcification shows endosteal scalloping about 50 %, no pathological fracture.

The lesion is reported as enchondroma by one reader and so classified it as OT-RADS 2, however, two readers reported it as atypical cartilaginous bony tumours/ malignant transformation of enchondroma and so classified it as OT-RADS 4, (Figure 2).

Case (2)

Clinical history: 36 –year old female did MRI RT forearm for painful mass at lower RT forearm.

MRI revealed: MRI revealed an ill-defined abnormal signal intensity intra-medullary expansile lesion involving distal radius with sub-particular location, relative narrow zone of transition of the surrounding bone, the lesion shows deep endosteal scalloping >2/3 cortical thickness. The lesion causes cortical destruction (at the volar

aspect), with extra-osseous soft tissue component.

The lesion is reported as giant cell tumour by two readers and so classified as OT-RADS 3, however, one reader reported the lesion as low-grade chondrosarcoma OT-RADS 4, (Figure 3).

Table 1: sensitivity, specificity, PPV, NPP, and AUC for the 4 readers

Reader	MSK experience y	sensitivity %	specificity %	PPV%	NPP%	AUC
reader1	20	81.25%	79.40%	65%	90%	0.864
reader2	12	87.50%	82.35%	70%	93.30%	0.886
reader3	7	81.25%	76.47%	61.90%	89.90%	0.844

Table 2: The agreement between the readers.

OT-RAD	reader1	reader2	reader3	ICC	CI	*P* value
I	2	1	2	0.905	0.854	<0.001
II	19	21	19		0.892	<0.001
III	9	8	10		0.901	<0.001
IV	20	20	19		0.942	<0.001

The intra observer agreement between the three readers, with intra class correlation =0.913, with confidence interval range (0.892: 0.942), *P* value < 0.001.

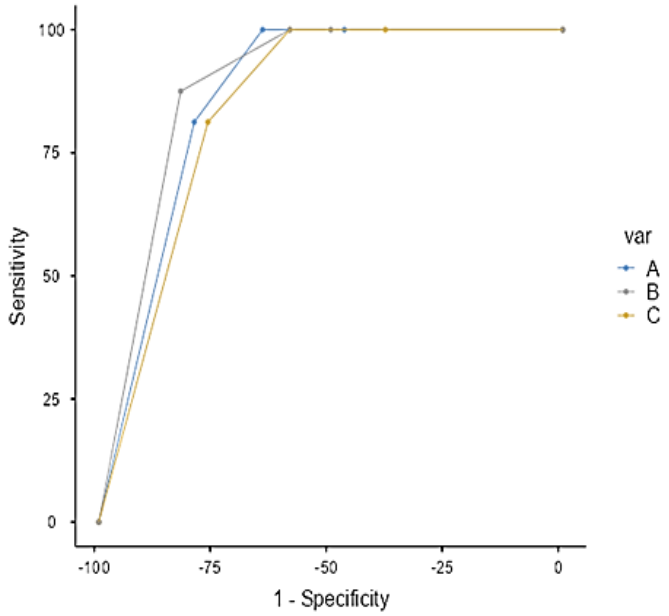

Figure 1: ROC curves for osseous tumours (AUC=0.844-0.886) shows the diagnostic accuracy of three readers in detecting tumours.

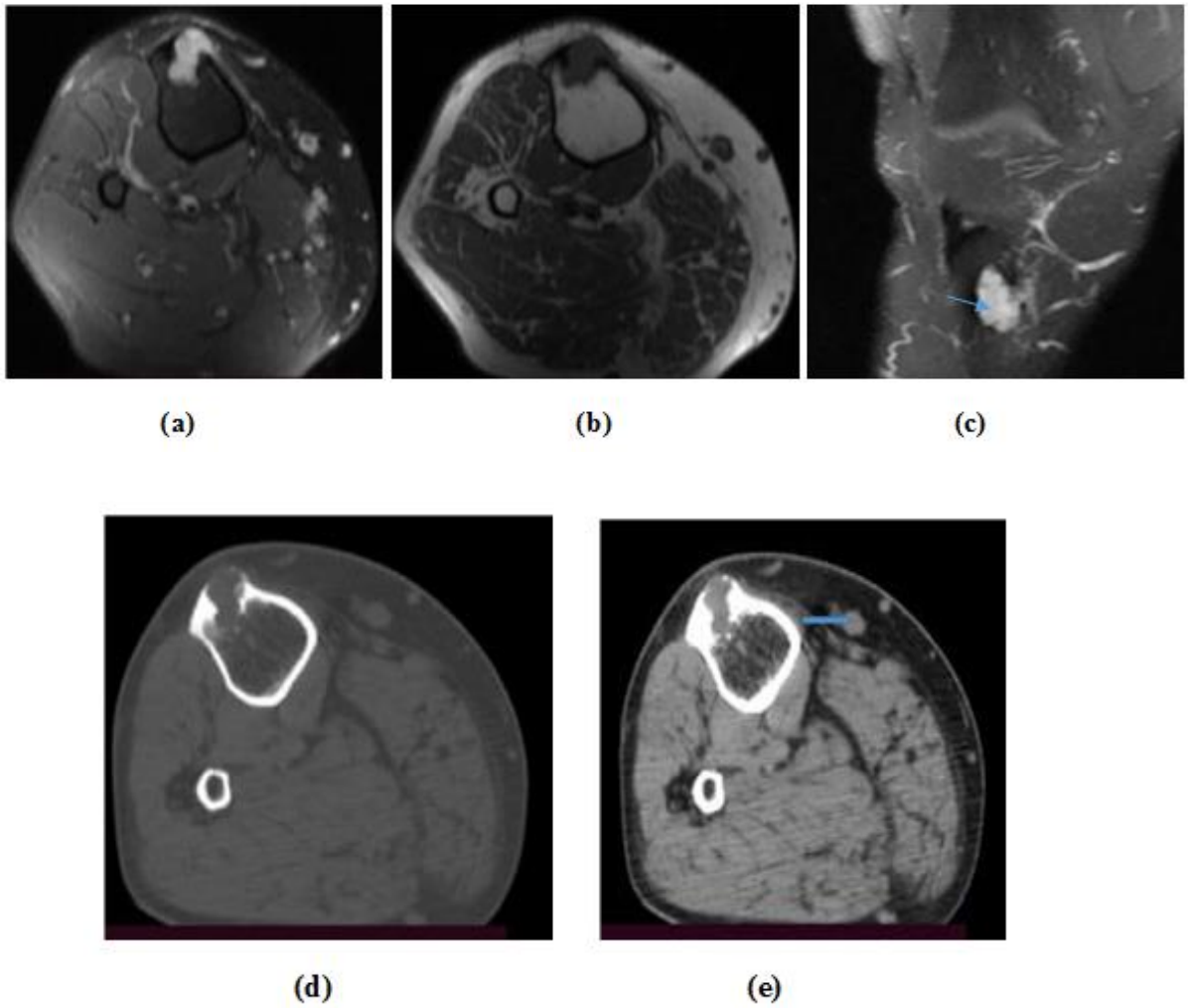

Figure 2 : A-E, a &b) axial T1WI pre and post contrast, c) coronal PD, d&e) CT axial bone window& soft tissue window images show rather defined oblong shaped exophytic osseous lesion seen at the anterior border of the upper third of the tibia displayed mixed signal intensity being iso to low signal intensity at T1WI with multiple internal signal voids and high signal intensities at T2WI and STIR corresponding to the chondroid matrix (thin arrow) , with moderate post contrast enhancement, the lesion is seen indenting the cortex with possible disruption (thick arrow).

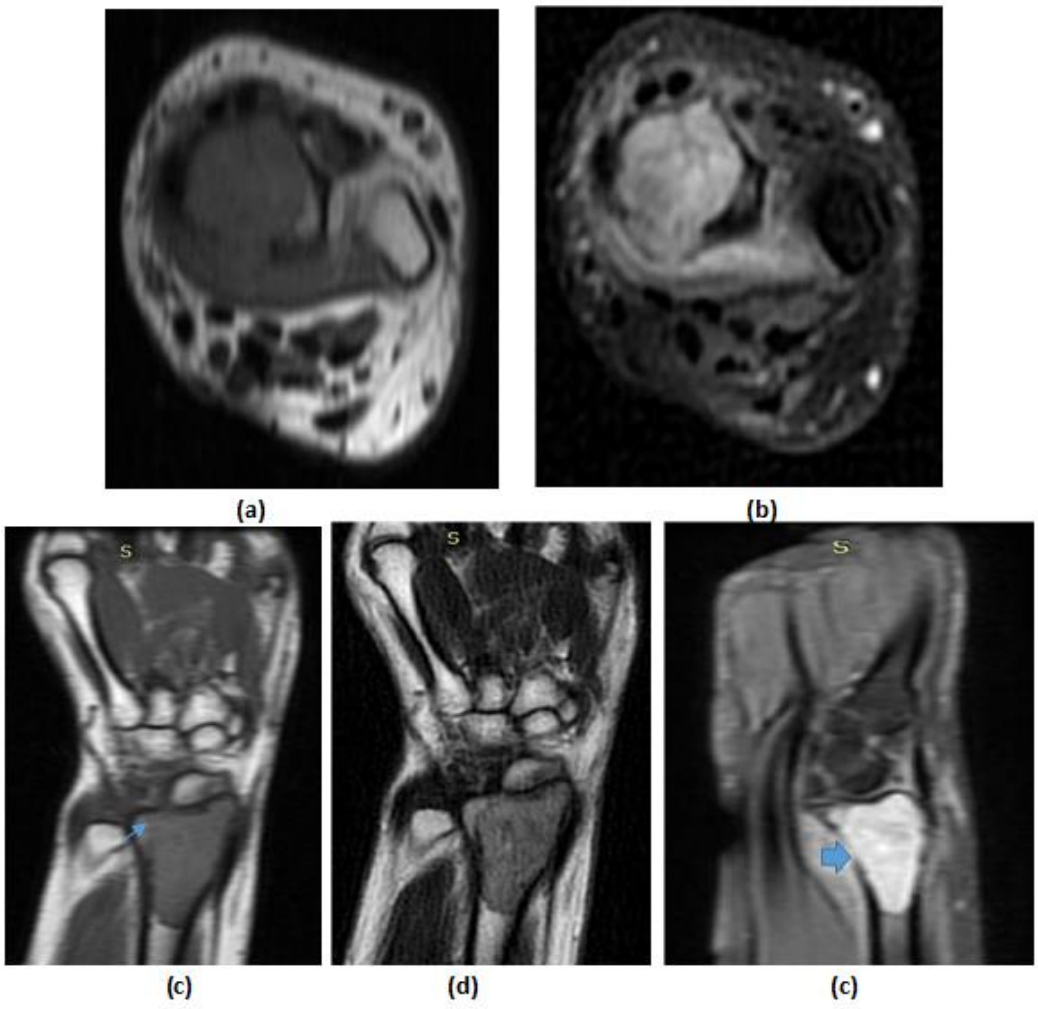

Figure 3: A-E, axial T2WI (A), axial PD(B), coronal T1WI (C), coronal T2WI(D) and sagittal PD(E) images show an ill-defined abnormal signal intensity intra medullary expansile lesion involving distal radius with sub-articular location(thin arrow), relative narrow zone of transition of the surrounding bone, the lesion shows deep endosteal scalloping >2/3 cortical thickness. The lesion causes cortical destruction (at the volar aspect), (thick arrow), with extra-osseous soft tissue component. The lesion displays low SI at T1WI, intermediate SI at T2WI and bright at STIR with moderate post contrast enhancement.

Discussion

Various studies have highlighted the importance of the analysis of medical information in structured way and use it in decision-making and to enhance the impact and quality of imaging. Consequently, we wish to develop and validate an Osseous Tumour Reporting and Data System (OT-RADS) with the hypothesis that it has great reliability and the guidelines can accurately separate benign from malignant bone tumours with excellent performance for categorization. However, more importantly in our opinion, the classification system is being appreciated by the treating consultants as the report become easier with clear

findings and the decision will be taken with more confidence [6].

Three general radiologists (**F.A,** reader #1 and **E. S.**, reader #2 and **A.M**., reader #3) with **20, 12 and 7** years of experience in MSK imaging with CT and MRI, respectively- independently- reviewed the images of the 50 participants included. The readers were blinded to the histopathological diagnosis. All studies were interpreted with a digital structured report created with a Microsoft® Excel spreadsheet using macro-programming. The report file was composed of dropdown menus and free text cells that allowed an automatic generation of an image finding database. Each image feature was evaluated by a specific modality (CT or MRI). Demographic and clinical data were also registered.

All 50 cases were pathologically diagnosed and the histopathological result used as a reference then the radiological assessment is done for all cases, benign-malignant distribution in the study sample was (68-32%) which is matched with (67–33%) distribution reviewed by **Ribeiro et al**, however, the study done by **Chhabra et al.,2021** reported more malignant percentage the benign –malignant distribution was (56.6-43.4%). This difference in the benign –malignant distribution may be due to the difference in age group and different demography.

OT-RADS is reported for each case by the three readers.

For assessment of OT-RADS: MRI was done for all cases, complementary CT was done for most of cases, using benign and malignant indicators and specific MRI imaging characteristics for each tumor type to reach definitive diagnosis. Different lesions were fit under OT-RADS category from OT-RADS1 to OT-RADS 6.

This assessment was done for each case and revealed that: Among the benign lesions studied: 38.2-47% (26-32/68) as 2; 8.8-17.6 % (6-12/68) as 3. With total **55.8-64.7% (**38-44) as 2 &3, **20.5–23.5%** (14-16/ 68) as 4 and **14.7-17.6%** (10-12/68) as 5.

No malignant tumours were classified as OT-RADS 2; **0-6.2%** (0-2/32) were classified as 3; **6.2–12.5%** (2-4/32) were classified as 4; and **87.5%** (28/32) were classified as 5.

To the best of our knowledge, there is a paucity of previously conducted studies evaluating the bone tumours using OT-RADS reporting system. However, RG conducted a relevant study analysing the bone tumours using another framework called BT-RADS. He found **58.7-65%** (91-101) as BTI-RADS total I& II, this corresponds to OT-RADS total 2& 3" benign lesions", **19.3–25.2%** (30–39/155) as BTI-RADS III this corresponds to OT-RADS4 " indeterminate lesions" ; and **14.2–21.3%** (22–33/155) as BTI-RADS IV this corresponds to OT-RADS 5" highly suspicious malignant". No malignant tumours were classified as BTI-RADS I; **1.3–4%** (1–3/75) were classified as II; **9.3–13.3%** (7–10/75) were classified as BTI-RADS III corresponding to OT-RADS 4; and **84–88%** (63–66/75) were

classified as BTI-RADS IV corresponding to OT-RAS 5.

OT-RADS agreed with gold standard results in (83.3%) true malignant cases and (95.4%) of true benign cases, while there were (4.5%) false negative cases (diagnosed as benign lesions by OT-RADS but were proven to be malignant by histopathology) and (72.2%) false positive cases (diagnosed as malignant masses by OT-RADS but were proven to be benign by histopathology).

There was good agreement between OT-RADS and gold standard results (K=0.58). The value was highly significant (*P* value <0.001).

Our study successfully validated the reliability of OT-RADS. The inter observer variability and the interclass correlation (ICC) and binary distinction of benign versus malignant (κ) evaluated by the three readers. (ICC) values were interpreted as follow:

Excellent agreement, 0.75 to 1.00; good agreement, 0.60–0.75; fair agreement, 0.40– to .60; and poor agreement, less than, 0.40.

The results were good to excellent ROC (AUC=0,844-0,886) and excellent intra class correlation (ICC =0.913), confidence interval (CI=0.854-0.942), (*P*<0.001). This inter observer agreement is more than the study reported by **Chhabra et al.,2021** who reported good to excellent ICC =0.78. But the study done by **Ribeiro et al.,2021** reported fair, that difference may be due to high level of experience of our readers and smaller sample size.

We found good diagnostic accuracy for identifying malignant from benign bony lesions with moderate sensitivity (81.3-87.5%) and moderate specificity (76.5-82,3%), PPV (62-70%) and NPP (90-93%). This percentage of sensitivity is less than the study done by **Chhabra et al.,2021** that reported 0.93-1.0 sensitivity, but is similar to the percentage of the reported specificity which is ranged from0.71-0.86, this study reported 0.7-0.8 range of PPV and 0.9-1 range of NPP.

There were some limitations to our study since benign-appearing bony lesions are not routinely biopsied, reference standard for these cases was a minimum 2 year follow up in addition to double reads by our senior radiologist. Cases without pathological diagnosis or fewer than 2 years of follow up were excluded from the cohort, in addition, this was a single centre study, but we recruited attending radiologists with a wide range of experience (7 to 20 years) in osseous tumour reading to characterize these tumours. Finally, CSI and DWI were not evaluated as a part of this initial validation work because they are not widely used. This work was intended to reproduce routine practices of osseous tumour imaging.

Conclusion

The OT-RADS classification will be a dynamic document, it has great value for enhancing structural reports with high accuracy in diagnosis and high reproducibility and better used by the consultants in their management decision.

References

1- **Dähnert W.** Radiology Review Manual. Philadelphia: Lippincott Williams & Wilkins; 2011,30:122-124.

2- **Ribeiro, G. J., Gillet, R., Blum, A., & Teixeira, P. A. G.** Imaging report and data system (RADS) for bone tumours: where do we stand and future directions. Skeletal Radiology, 2023; 52(2): 151-156.

3- **Ribeiro, G. J., Gillet, R., Hossu, G., Trinh, J. M., Euxibie, E., Sirveaux, F., et al.,** Solitary bone tumour imaging reporting and data system (BTI-RADS): initial assessment of a systematic imaging evaluation and comprehensive reporting method. European Radiology, 2021: 1-16.

4- **Caracciolo, J. T., Ali, S., Chang, C. Y., Degnan, A. J., Flemming, D. J., Henderson, E. R., et al.,** Bone Tumour Risk Stratification and Management System: A Consensus Guideline from the ACR Bone Reporting and Data System Committee. Journal of the American College of Radiology, 2023; 20(10): 1044-1058.

5- **Abdella, A. E. F., Elshafey, K. I., Sherif, M. F., & Nagy, H. A.** Diagnostic performance of PET/CT in primary malignant bone tumours. Egyptian Journal of Radiology and Nuclear Medicine, 2021; 52: 1-13.

6- **Chhabra, A., Gupta, A., Thakur, U., Pezeshk, P., Dettori, N., Callan, A., et al.,** Osseous tumor reporting and data System—Multi-reader validation study. Journal of computer assisted tomography, 2021; 45(4): 571-585.

To cite this article: Sherif A. Abd-Elsattar, Ahmed M. Alsowey, Hesham E. El Sheikh, Maha E. Bekheet . Diagnostic Validity and Reliability of Osseous Tumour Reporting and Data System (OT-RADS) in the Management of Primary Bone Tumour. BMFJ XXX, DOI: 10.21608/bmfj.2025.357161.2320