Clinical Imaging 137 (2026) 110879

Contents lists available at ScienceDirect

Clinical Imaging

journal homepage: www.elsevier.com/locate/clinimag

Breast Imaging

Intra- and inter-reader agreement of the Node-RADS v1.0 score for axillary lymph node assessment on preoperative breast MRI

P. Minichetti [a], L. Cereser [a,b,*], A.P. Pace [b], F. Francioso [b], E. Versienti [b], F. Sparascio [a], R. Girometti [a,b], C. Zuiani [a,b]

[a] Institute of Radiology, University Hospital S. Maria della Misericordia, Azienda Sanitaria Universitaria Friuli Centrale (ASUFC), Udine, Italy
[b] Institute of Radiology, Department of Medicine (DMED), University of Udine, Italy

ARTICLE INFO

Keywords:
Breast cancer
Axillary lymph nodes
Node-RADS
Breast MRI
Computed tomography
Reproducibility

ABSTRACT

Purpose: To assess intra- and inter-reader agreement of the Node-RADS v1.0 score for axillary lymph node evaluation in breast cancer (BC) patients undergoing preoperative breast MRI (bMRI), and to explore its patient-level association with pathological nodal involvement. Contrast-enhanced CT was evaluated as a subgroup exploratory analysis.

Methods: This retrospective study included 100 women with biopsy-proven BC who underwent preoperative bMRI; 25 also underwent contrast-enhanced CT. Four readers (two expert breast radiologists and two junior readers) independently assigned Node-RADS v1.0 scores to the most suspicious axillary lymph node on MRI (n = 100) and CT (n = 25). Thirty bMRI examinations were re-evaluated after a 3-week washout to assess intra-reader agreement. Agreement was assessed using percent agreement and quadratic-weighted Cohen's κ. Exploratory patient-level diagnostic performance for pathological nodal involvement was evaluated using ROC analysis and AUC values.

Results: Inter-reader agreement on bMRI was moderate among junior readers (κ = 0.59; 95%CI: 0.39–0.79) and almost perfect among expert readers (κ = 0.98; 95%CI: 0.96–1.00). Intra-reader agreement ranged from substantial to almost perfect. Diagnostic performance on bMRI was moderate for most readers (AUC 0.62–0.64), with higher AUC in one junior reader (AUC 0.79). In the clinically selected CT subgroup, AUC values ranged from 0.74 to 0.90. A Node-RADS threshold >2 showed the most balanced sensitivity-specificity trade-off in this cohort.

Conclusion: Node-RADS v1.0 showed good to excellent reproducibility for axillary lymph node assessment, particularly among expert readers. Diagnostic performance was exploratory and variable, reflecting the patient-level reference standard and the one-node-per-axilla approach. A Node-RADS threshold >2 may represent a pragmatic exploratory cut-off for patient-level risk stratification.

1. Introduction

Breast cancer (BC) staging is based on the TNM classification system developed by the American Joint Committee on Cancer (AJCC),[1] which defines disease extent according to the size of the primary tumor (T), regional lymph nodes involvement (N), and distant metastases (M). Among these parameters, nodal status is a key prognostic factor and a major determinant in therapeutic decision-making, particularly regarding surgical planning and systemic treatment.

While axillary ultrasound (US) remains the first-line imaging modality for lymph node evaluation and biopsy guidance, breast MRI (bMRI), used for locoregional staging, and computed tomography (CT), typically employed in patients with advanced-stage BC, provide a more standardized, operator-independent, and wide-field assessment of axillary and internal mammary lymph nodes.[2] However, despite the routine description of lymph node findings in bMRI and CT reports, interpretation criteria have historically been heterogeneous and subjective.[2,3]

To address this gap, the Node Reporting and Data System (Node-RADS v1.0)[3] was introduced in 2021, proposing a structured 5-point scale to estimate the likelihood of nodal malignancy on CT and bMRI. The system is based on two descriptors, “size” and “configuration”, whose combination yields a 5-point Node-RADS score, as shown in

* Corresponding author at: Institute of Radiology, Department of Medicine - University of Udine, University Hospital S. Maria della Misericordia, 33100, Udine, Italy.
E-mail address: lorenzo.cereser@uniud.it (L. Cereser).

https://doi.org/10.1016/j.clinimag.2026.110879
Received 18 February 2026; Received in revised form 18 June 2026; Accepted 22 June 2026
Available online 24 June 2026

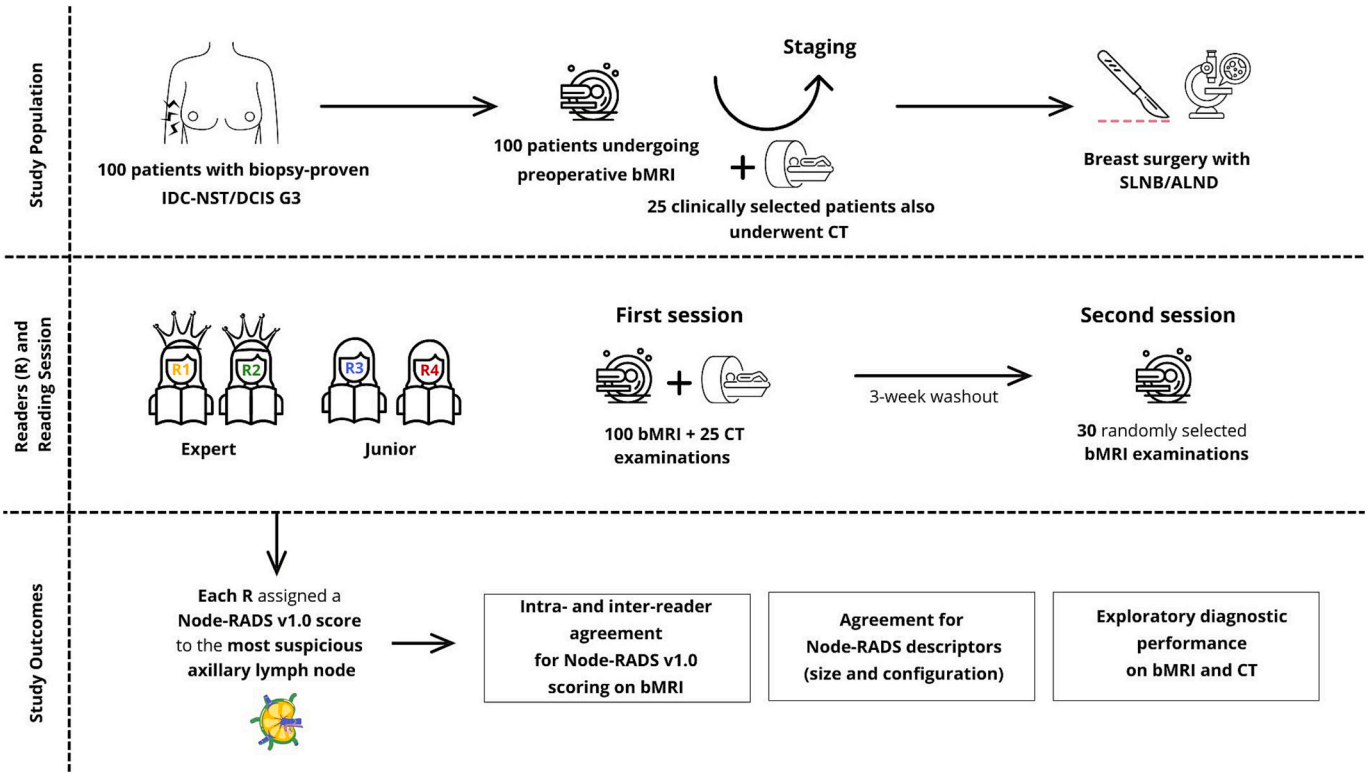


Fig. 1. The study design. NST: No Special Type; DCIS: Ductal Carcinoma In-situ; bMRI: breast Magnetic Resonance Imaging, CT: Computed tomography; SLNB: sentinel lymph node biopsy; ALND: axillary lymph node dissection. Figure created using Canva.

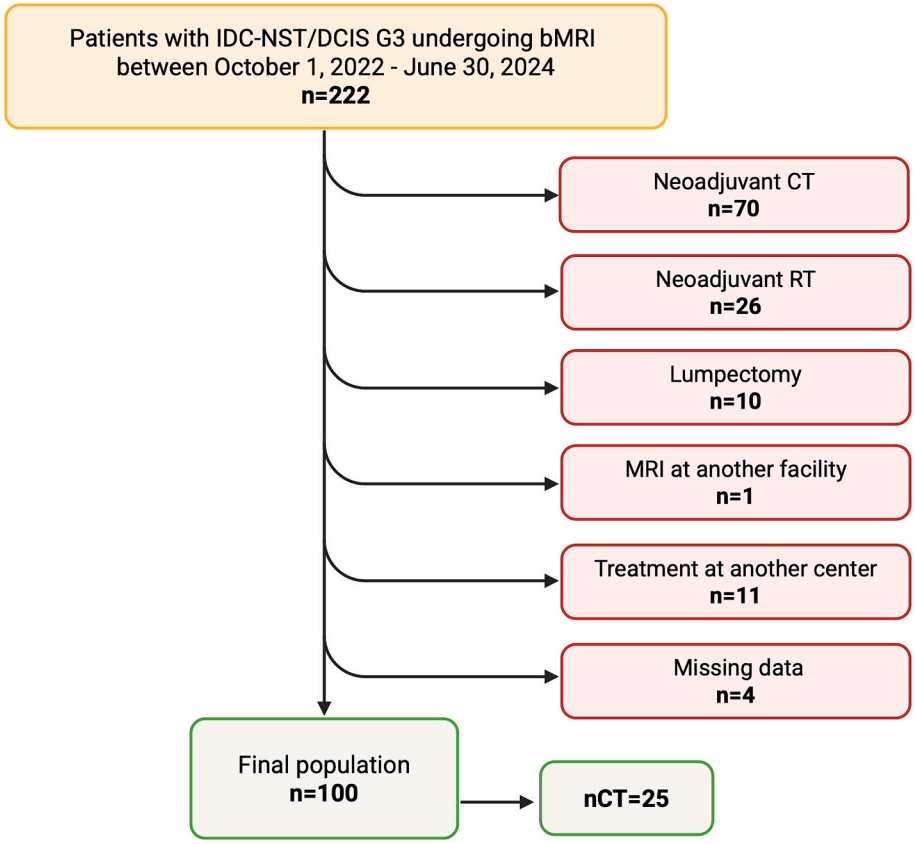
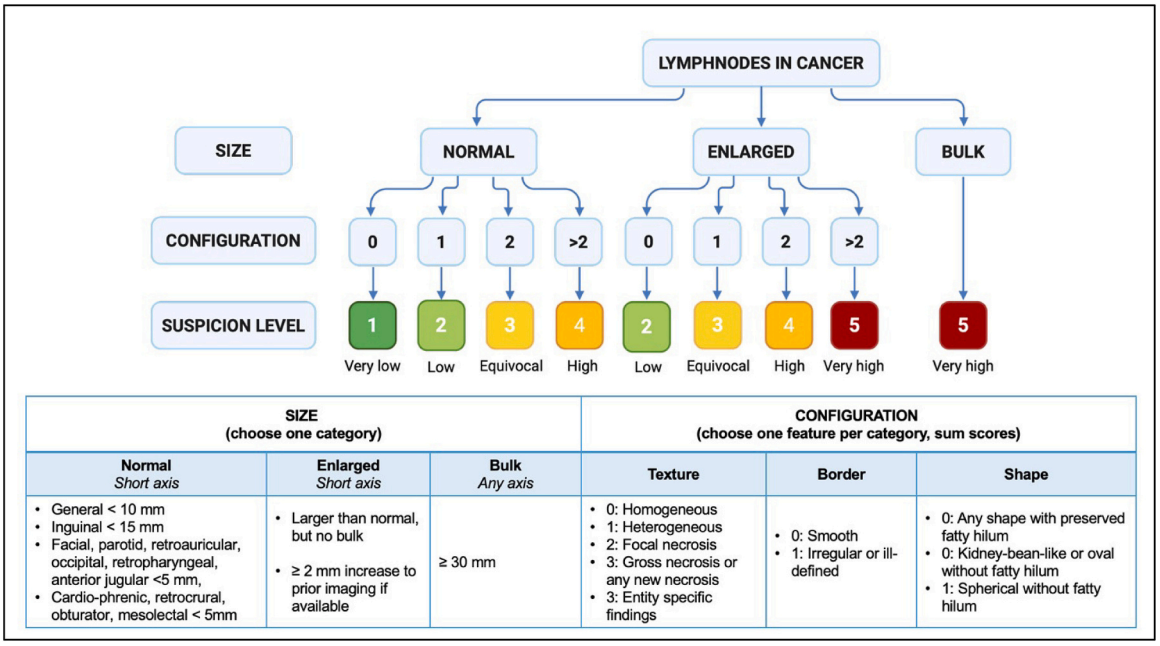


Fig. 2. Inclusion and exclusion criteria flow-chart. IDC-NST: Invasive Ductal Carcinoma – No Special Type; BC: breast cancer; DCIS: Ductal Carcinoma In-situ; bMRI: breast MRI; CT: chemotherapy; RT: radiotherapy; nCT: number of patients who underwent computed tomography. Figure created with BioRender.com.



SIZE (choose one category)			CONFIGURATION (choose one feature per category, sum scores)		
Normal *Short axis*	**Enlarged** *Short axis*	**Bulk** *Any axis*	**Texture**	**Border**	**Shape**
• General < 10 mm • Inguinal < 15 mm • Facial, parotid, retroauricular, occipital, retropharyngeal, anterior jugular <5 mm, • Cardio-phrenic, retrocrural, obturator, mesolectal < 5mm	• Larger than normal, but no bulk • ≥ 2 mm increase to prior imaging if available	≥ 30 mm	• 0: Homogeneous • 1: Heterogeneous • 2: Focal necrosis • 3: Gross necrosis or any new necrosis • 3: Entity specific findings	• 0: Smooth • 1: Irregular or ill-defined	• 0: Any shape with preserved fatty hilum • 0: Kidney-bean-like or oval without fatty hilum • 1: Spherical without fatty hilum

Fig. 3. The Node Reporting and Data System (Node-RADS v1.0) *Adapted from Elsholtz FHJ* et al. *Introducing the Node Reporting and Data System 1.0, 2021.* Figure created with BioRender.com.

Fig. 3. A score of 1 indicates a "very low probability," and a score of 5 indicates a "very high probability" of nodal involvement by malignancy. The aim is to improve reproducibility and reporting consistency across imaging modalities and clinical scenarios both in primary staging and in treatment response assessment.

More recently, the American College of Radiology (ACR) introduced in the Sixth Edition of the Breast Imaging Reporting and Data System (BI-RADS), released as the BI-RADS® v2025 Manual,[4] a dedicated section for lymph node assessment across breast imaging modalities, including bMRI. Node-RADS may provide a complementary categorical framework to support standardized and reproducible nodal assessment within existing bMRI reporting. In the preoperative bMRI setting, its intended role is to support nodal risk stratification, rather than to serve as a stand-alone diagnostic test for nodal metastasis.

Although Node-RADS v1.0 has potential applications in several clinical settings,[5] its clinical validation in BC remains limited. To date, only one study has specifically investigated Node-RADS on bMRI in BC patients with a primary focus on inter-reader agreement among expert readers.[6] Building on this preliminary evidence, our investigation further evaluates reader agreement across different experience levels and explores the applicability of Node-RADS to contrast-enhanced CT.

The primary aim of this study was to evaluate intra- and inter-reader agreement in assigning Node-RADS v1.0 scores for axillary lymph node assessment in BC patients undergoing preoperative bMRI for local staging. Secondary aims were to assess inter-reader agreement for the individual Node-RADS descriptors (size and configuration) and to explore the patient-level association between Node-RADS and pathological nodal involvement on bMRI and contrast-enhanced CT, when available.

2. Materials and methods

The study was approved by the Institutional Review Board, and the requirement for written informed consent was waived due to the retrospective study design. The study workflow is summarized in Fig. 1.

2.1. Study population and reference standard

Consecutive patients with biopsy-proven invasive ductal carcinoma of no special type (IDC-NST) or grade 3 ductal carcinoma in situ (DCIS) who underwent preoperative bMRI at our university hospital between October 1, 2022 and June 30, 2024 were retrospectively included. Indications for preoperative imaging followed the European Society of Breast Cancer Specialists (EUSOMA) criteria[7] or were based on clinical assessment. At our institution, preoperative bMRI is not routinely performed in all BC patients but is requested selectively after multidisciplinary or breast radiology evaluation, mainly in cases requiring more accurate local staging or assessment of disease extent. Exclusion criteria included prior neoadjuvant chemotherapy or radiotherapy, breast surgery performed before bMRI, bMRI performed at another institution, definitive treatment carried out at another center, or other reasons leading to incomplete imaging or histopathological data, as detailed in the study flowchart (Fig. 2). The final study cohort comprised 100 women. Among them, a subgroup of 25 patients also underwent contrast-enhanced CT. CT was performed based on clinical indication, mainly for suspected advanced disease or staging work-up according to institutional practice. This subgroup was therefore clinically selected and enriched for nodal involvement, and CT analyses were considered exploratory rather than a head-to-head comparison with bMRI.

The reference standard consisted of histopathological assessment obtained after sentinel lymph node biopsy (SLNB) or axillary lymph node dissection (ALND). In patients with grade 3 DCIS, axillary surgery was not routinely performed but was reserved for selected cases according to multidisciplinary evaluation and institutional surgical pathways, including patients considered at risk of occult invasive disease or undergoing mastectomy. Histopathological evaluation was performed by one of three board-certified pathologists with 5–25 years of subspecialty experience in breast pathology, in accordance with College of American Pathologists guidelines.[8] Histopathological results were available at the patient level and served as the reference for imaging-based nodal assessment, reflecting routine clinical practice. Lymph node involvement was defined by the presence of micrometastases (0.2–2 mm) or macrometastases (>2 mm), whereas isolated tumor cells (<0.2 mm) were considered node-negative for the purposes of this study.[9]

2.2. Breast MRI and CT technique

Breast MRI examinations were performed on one of two 1.5-T scanners (MAGNETOM Aera or MAGNETOM Avanto; Siemens Healthcare, Erlangen, Germany) using a bilateral 16-channel breast coil, with patients in the prone position. The standard bMRI protocol included diffusion-weighted imaging (DWI), T2-weighted imaging, and non–fat-suppressed T1-weighted dynamic contrast-enhanced (DCE) imaging, with subtraction images reconstructed from pre- and post-contrast acquisitions. Gadoteridol (ProHance; Bracco Imaging S.p.A., Milan, Italy) was administered intravenously at a dose of 0.1 mmol/kg and an injection rate of 2 mL/s, followed by a 20-mL saline flush. Post-contrast T1-weighted acquisitions were obtained at 1, 3, 5, 7, and 9 min after contrast administration, according to the institutional protocol (Supplementary Table 1).

All CT examinations were performed using 64-row multidetector CT scanners (Discovery CT750 HD or Optima CT660; GE Healthcare, Milwaukee, WI, USA), with patients in the supine position during full inspiration. Images were acquired in the venous phase after intravenous administration of iodinated contrast agent, either iomeprol (Iomeron; Bracco Imaging S.p.A., Milan, Italy) or iobitridol (Xenetix; Guerbet, Villepinte, France), with an iodine concentration of 350–400 mg/mL, according to the institutional protocol (Supplementary Table 2).

2.3. Image analysis

A study coordinator organized two independent reading sessions separated by a 3-week washout interval and recorded, in a dedicated worksheet, the axillary side for which histopathological reference data were available.

Image analysis was performed independently by four radiologists (R1–R4) with different levels of experience in breast imaging. Readers R1 and R2 were classified as expert radiologists, defined as board-certified radiologists meeting the criteria recommended by the European Society of Breast Cancer Specialists (EUSOMA) and the European Cancer Concord (ECC) for international Breast Unit certification,[10] each with 5 years of continuous clinical activity in breast imaging. Readers R3 and R4 were classified as junior readers (radiology residents), each with 2 years of experience in breast imaging and approximately 200 previously interpreted bMRI examinations. All readers were blinded to clinical, histopathological, and follow-up information, except for knowledge of the laterality of BC.

Image interpretation followed the Node-RADS v1.0 recommendations and was supported by a three-step flowchart and a concise checklist summarizing the scoring criteria (Fig. 3). When multiple lymph nodes were present in the same axilla, readers assessed the node expected to receive the highest Node-RADS v1.0 category after combined evaluation of size and configuration. In case of equivalent suspicion, the node with the most suspicious configuration was selected; if still equivalent, the largest node was selected. Diagnostic performance analyses compared this single-node imaging assessment with patient-level pathological nodal status and were therefore considered exploratory. During the first reading session, each reader assigned a Node-RADS score to the most suspicious lymph node on all bMRI examinations (n = 100) and, when available, on contrast-enhanced CT examinations (n = 25). In the second reading session, each reader re-evaluated 30 of the 100 bMRI examinations, randomly selected by the study coordinator using Randomizer.org, to assess intra-reader agreement.

All bMRI and CT examinations were reviewed on a dedicated picture archiving and communication system (PACS) workstation (Suitestensa; Ebit Srl, an Esaote Group company, Genoa, Italy).

2.4. Data analysis

Continuous variables are reported as mean ± standard deviation or median and interquartile range (IQR), as appropriate, after assessment of normality using the Kolmogorov–Smirnov test. Categorical variables are presented as counts and percentages.

Node-RADS v1.0 score assignments were summarized descriptively for each reader. Inter- and intra-reader agreement was assessed using percent agreement and quadratic-weighted Cohen's κ statistics with 95% confidence intervals (CI), accounting for the ordinal nature of the Node-RADS categories.[11,12] Agreement strength was interpreted according to the Landis and Koch criteria.[13]

Secondary analyses evaluated inter-reader agreement for the individual Node-RADS descriptors ("size" and "configuration") using the same statistical approach. The "size" descriptor was categorized as normal, enlarged, or bulky, whereas "configuration" was defined based on combined assessment of internal texture, nodal borders, and shape, according to Node-RADS recommendations.[3]

Diagnostic performance for predicting pathological lymph node involvement was assessed by calculating sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) across all Node-RADS thresholds. Cut-off analysis was performed by progressively grouping Node-RADS categories: score 1 versus scores 2–5 (cut-off >1), scores 1–2 versus scores 3–5 (cut-off >2), scores 1–3 versus scores 4–5 (cut-off >3), and scores 1–4 versus score 5 (cut-off >4). For each threshold, likelihood ratios (LR+ and LR–) were also calculated.

Receiver operating characteristic (ROC) curve analysis was used to assess overall diagnostic performance, with area under the curve (AUC) values calculated for each reader. Threshold-specific sensitivity, specificity, PPV, NPV, and LR were calculated descriptively across Node-RADS cut-offs. At the exploratory Node-RADS >2 threshold, exact binomial 95% confidence intervals were calculated for sensitivity and specificity. All reader comparisons were exploratory and intended to describe between-reader variability rather than support formal hypothesis testing.

All statistical analyses were performed using Stata/BE, version 19.5 (StataCorp LLC, College Station, TX, USA).

Table 1
The main histological and surgical characteristics of the 100 women included in the study.

Features	N (%; 95% CI)	
Histological type	IDC-NST G1	9/100 (9; 3–15)
	IDC-NST G2	61/100 (61; 51–71)
	IDC-NST G3	11/100 (11; 4–17)
	DCIS G3	19/100 (19; 11–27)
Breast surgical treatment	Mastectomy	24/100 (24; 16–32)
	Quadrantectomy	76/100 (76; 68–84)
Axillary surgical treatment	ALND	19/100 (19; 11–27)
	SLNB	81/100 (81; 73–89)

Table 2
Distribution of Node-RADS v1.0 scores according to the readers in the first reading session.

	Expert readers		Junior readers	
	R1	R2	R3	R4
Node-RADS v1.0 score 1	70/100 (70%)	71/100 (71%)	66/100 (66%)	75/100 (75%)
Node-RADS v1.0 score 2	15/100 (15%)	14/100 (14%)	22/100 (22%)	11/100 (11%)
Node-RADS v1.0 score 3	8/100 (8%)	7/100 (7%)	8/100 (8%)	6/100 (6%)
Node-RADS v1.0 score 4	3/100 (3%)	4/100 (4%)	3/100 (3%)	5/100 (5%)
Node-RADS v1.0 score 5	4/100 (4%)	4/100 (4%)	1/100 (1%)	3/100 (3%)

Footnotes: R1, reader 1; R2, reader 2; R3, reader 3; R4, reader 4.

Table 3
Intra-reader agreement results for Node-RADS v1.0.

		PA	95% CI	Interpretation	Cohen's Kappa	95% CI	Interpretation
Expert readers	R1	0.94	0.80–0.99	Almost perfect	0.64	0.28–1.00	Substantial
	R2	0.95	0.83–1.00	Almost perfect	0.68	0.34–1.00	Substantial
Junior readers	R3	1.00	1.00–1.00	Almost perfect	1.00	1.00–1.00	Almost perfect
	R4	0.98	0.95–1.00	Almost perfect	0.86	0.70–1.00	Almost perfect

Footnotes: R1, reader 1; R2, reader 2; R3, reader 3; R4, reader 4; PA, percent agreement; 95% CI, 95% confidence interval.

Table 4
Inter-reader agreement results for the Node-RADS v1.0 descriptors.

		PA	95% CI	Interpretation	Cohen's Kappa	95% CI	Interpretation
Size							
Expert readers	R1	0.99	0.99–1	Almost perfect	0.90	0.75–1	Almost perfect
	R2						
Junior readers	R3	0.97	0.94–1	Almost perfect	0.75	0.48–1	Substantial
	R4						
Configuration							
Expert readers	R1	0.99	0.99–1	Almost perfect	0.99	0.97–1	Almost perfect
	R2						
Junior readers	R3	0.93	0.80–0.94	Almost perfect	0.46	0.27–0.66	Moderate
	R4						

Footnotes: R1, reader 1; R2, reader 2; R3, reader 3; R4, reader 4; PA, percent agreement; 95% CI, 95% confidence interval.

3. Results

3.1. Study population and lymph node status

The final study cohort comprised 100 women (median age, 50.7 years; range, 29–79 years). The main histological and surgical characteristics of the study population are summarized in Table 1. Axillary lymph node involvement was identified in 27/100 patients in the bMRI cohort (27%) and in 20/25 patients in the CT subgroup (80%). The mean number of lymph nodes removed during sentinel lymph node biopsy was 2.8 ± 1.8, whereas the mean number excised during axillary lymph node dissection was 15.0 ± 3.9.

3.2. Inter- and intra-reader agreement of Node-RADS v1.0 scores

Table 2 summarizes the distribution of Node-RADS v1.0 scores assigned by each reader during the first reading session.

Inter-reader agreement for Node-RADS v1.0 scoring was almost perfect among expert readers (R1–R2; $\kappa = 0.98$; 95% CI: 0.96–1.00) and moderate among junior readers (R3–R4; $\kappa = 0.59$; 95% CI: 0.39–0.79).

Intra-reader agreement results for Node-RADS v1.0 scoring are reported in Table 3. Agreement ranged from substantial to almost perfect for both expert and junior readers, with quadratic-weighted Cohen's κ values between 0.64 and 1.00.

3.3. Inter- reader agreement of Node-RADS v1.0 descriptors

During the first reading session, inter-reader agreement for the individual Node-RADS v1.0 descriptors varied according to reader experience and descriptor type. For the size descriptor, agreement was almost perfect among expert readers ($\kappa = 0.90$; 95% CI: 0.75–1.00) and substantial among junior readers ($\kappa = 0.75$; 95% CI: 0.48–1.00). For the configuration descriptor, agreement was almost perfect among expert readers ($\kappa = 0.99$; 95% CI: 0.97–1.00) but moderate among junior readers ($\kappa = 0.46$; 95% CI: 0.27–0.66). Detailed inter-reader agreement results for each descriptor are summarized in Table 4.

3.4. Diagnostic performance and cut-off analysis of the Node-RADS score

On bMRI, ROC analysis showed moderate exploratory patient-level diagnostic performance, with AUC values ranging from 0.62 to 0.79 across readers (Table 5, Fig. 4a). In this cohort ($n = 100$; 27 node-positive and 73 node-negative patients), a Node-RADS threshold >2 showed the most balanced sensitivity-specificity trade-off. Detailed metrics at this threshold are reported in Table 5. These findings indicate that, despite good reproducibility, the diagnostic value of Node-RADS on bMRI was limited and variable at the patient level.

In the clinically selected CT subgroup, in which pathological nodal involvement was present in 20/25 patients (80%), AUC values ranged

Table 5
Diagnostic Performance of the Node-RADS v1.0 score for bMRI and CT per Reader.

Modality	Metric	Expert readers		Junior readers	
		R1	R2	R3	R4
bMRI	AUC (95% CI)	0.62 (0.50–0.73)	0.64 (0.53–0.75)	0.64 (0.53–0.76)	0.79 (0.70–0.89)
	Sensitivity at >2, % (95% CI)	26 (11–46)	26 (11–46)	26 (11–46)	44 (26–65)
	Specificity at >2, % (95% CI)	89 (80–95)	89 (80–95)	93 (85–98)	97 (91–100)
	TP/FP/TN/FN at >2	7/8/65/20	7/8/65/20	7/5/68/20	12/2/71/15
CT	AUC (95% CI)	0.78 (0.59–0.96)	0.79 (0.59–0.98)	0.74 (0.53–0.94)	0.90 (0.79–1.00)
	Sensitivity at >2, % (95% CI)	65 (41–85)	60 (36–81)	30 (12–54)	75 (51–91)
	Specificity at >2, % (95% CI)	80 (28–100)	80 (29–100)	100 (48–100)	100 (48–100)
	TP/FP/TN/FN at >2	13/1/4/7	12/1/4/8	6/0/5/14	15/0/5/5

Footnotes: R1, reader 1; R2, reader 2; R3, reader 3; R4, reader 4; AUC, area under the curve; 95% CI, 95% confidence interval; bMRI, breast magnetic resonance imaging; CT, computed tomography; TP, true positive; FP, false positive; TN, true negative; FN, false negative.
At the >2 threshold, scores 3–5 were considered positive.

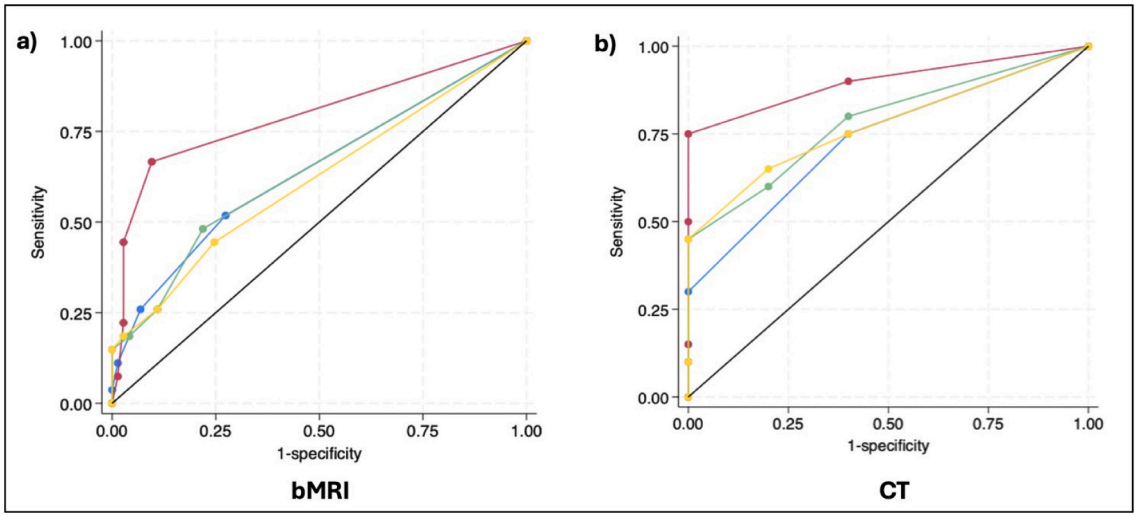


Fig. 4. ROC curves and AUC in a) bMRI and b) CT for Expert Radiologist R1 (yellow), Expert Radiologist R2 (green), Radiologist in training R3 (blue), Radiologist in training R4 (red). ROC: receiver operating characteristic; AUC: area under the curve. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

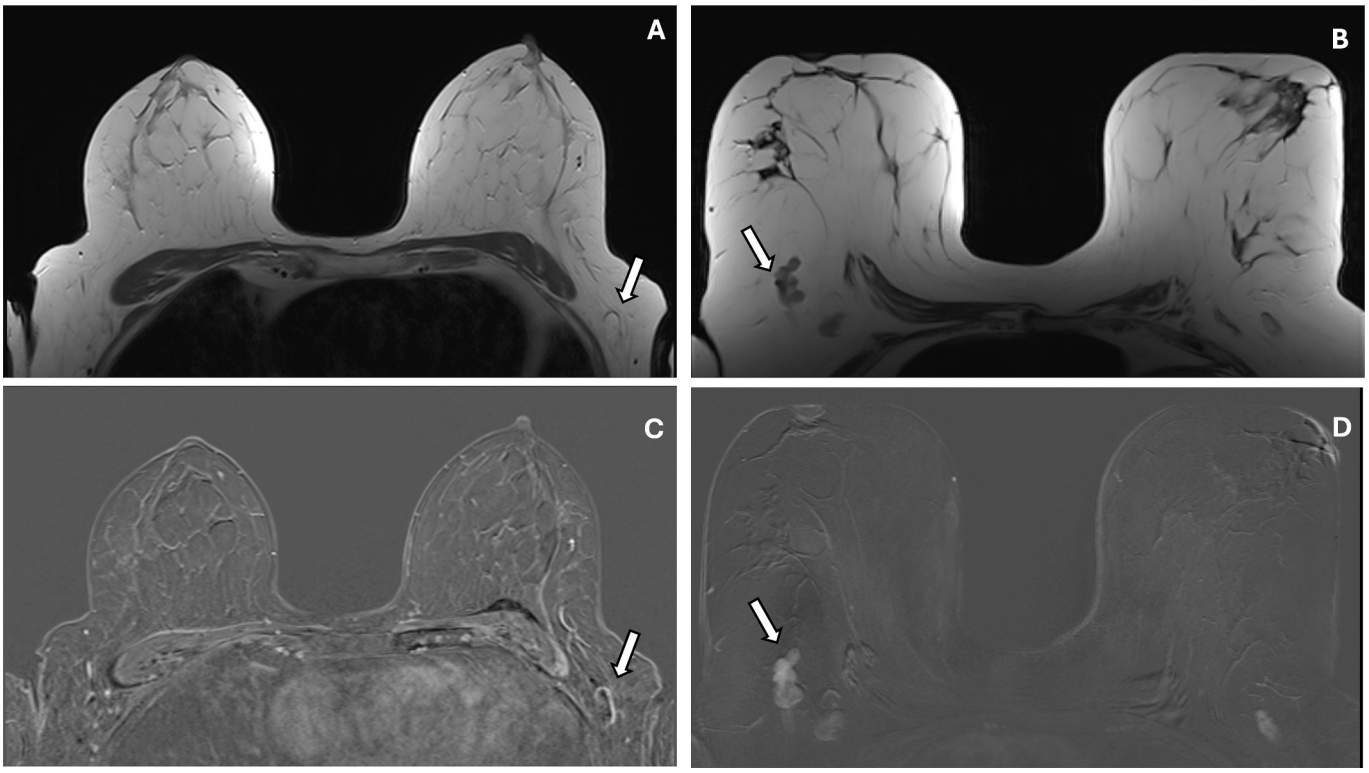


Fig. 5. *Node-RADS 1* (A, C) – 44-year-old woman with grade 1 invasive breast carcinoma of no special type (IBC-NST) in the left breast undergoing preoperative MRI. (A) Axial T2-weighted image shows an oval lymph node with smooth borders and preserved fatty hilum (arrow); (C) Axial subtracted contrast-enhanced T1-weighted image confirms normal size and shows homogeneous enhancement (arrow). These features are consistent with Node-RADS 1.
Node-RADS 5 (B, D) – 56-year-old woman with multifocal grade 2 IBC-NST in the right breast undergoing preoperative MRI. (B) Axial T2-weighted image shows a bulky lymph nodal mass (arrow); (D) Axial subtracted contrast-enhanced T1-weighted image confirms these findings (arrow). These features are consistent with Node-RADS 5.

from 0.74 to 0.90 across readers (Table 5, Fig. 4b). Within this subgroup, a Node-RADS threshold >2 showed the most balanced sensitivity-specificity trade-off (Table 5).

Cut-off analysis showed the expected threshold-dependent pattern: lower thresholds were associated with high sensitivity and low specificity, whereas higher thresholds increased specificity and positive likelihood ratios at the expense of sensitivity. Detailed cut-off analyses are reported in Supplementary Tables 3 and 4.

Representative examples of Node-RADS v1.0 score assignments are shown in Figs. 5, 6, and 7.

4. Discussion

Overall, Node-RADS v1.0 demonstrated high reproducibility for

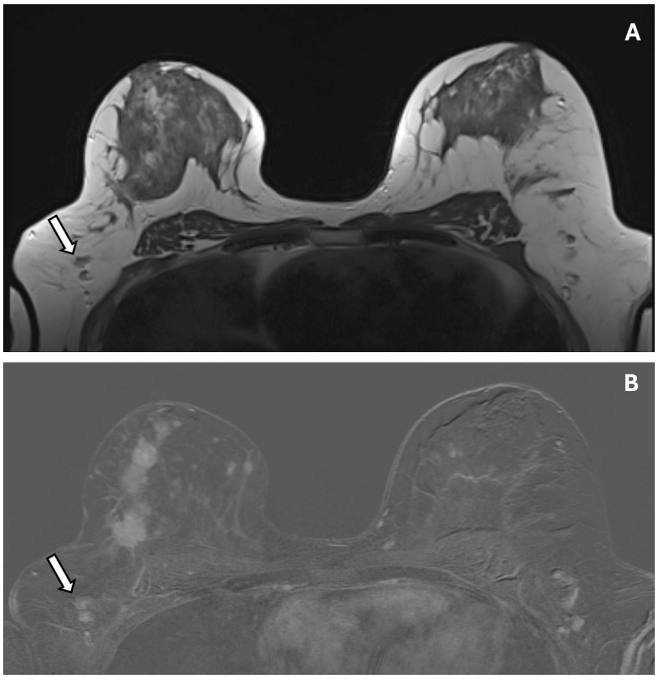


Fig. 6. *Node-RADS 3*–52-year-old woman with multicentric grade 3 breast carcinoma of no special type (IBC-NST) in the right breast undergoing preoperative MRI. (A) Axial T2-weighted image shows a kidney-bean-like lymph node with irregular borders, without fatty hilum (arrow); (B) Axial contrast-enhanced T1-weighted image confirms normal size and shows heterogeneous enhancement (arrow). These features are consistent with Node-RADS 3.

axillary lymph node assessment on preoperative bMRI, with almost perfect inter- and intra-reader agreement among expert readers and moderate to substantial agreement among junior readers. These findings are consistent with previously published data on Node-RADS reproducibility in bMRI,[6] while indicating that reader experience may influence agreement, particularly at the descriptor level. The reproducibility observed among junior readers suggests that the structured and simplified architecture of Node-RADS may facilitate consistent nodal classification even in less experienced hands, supporting its role as a standardized reporting framework.

Agreement for the individual Node-RADS descriptors, namely lymph node size and configuration, was generally high, although it varied by descriptor and reader experience. Agreement was consistently higher for the size descriptor than for configuration, particularly among junior readers, suggesting greater interpretative variability when more complex morphological features are considered. Overall, the use of a limited set of predefined descriptors appears to support reproducible nodal assessment, while still leaving room for an effect of reader experience.

A key observation of this study is the dissociation between the high reproducibility of Node-RADS scoring and its more modest discriminative performance on bMRI. Despite good inter- and intra-reader agreement, diagnostic accuracy remained only moderate for most readers, indicating that reproducibility alone does not ensure clinically sufficient diagnostic performance when relying on morphological criteria. Accordingly, ROC analysis on bMRI yielded modest AUC values for three readers (0.62–0.64), with higher accuracy observed for one junior reader (AUC 0.79). Compared with previous reports,[6] these findings likely reflect the lower prevalence of nodal disease in our cohort, differences in MRI protocol and field strength (1.5 T vs 3 T), variability in reader experience, and differences in study design.

Our results are consistent with a meta-analysis reporting high specificity but only moderate sensitivity for Node-RADS, together with substantial heterogeneity across studies and imaging modalities.[14] A comparative study reported lower sensitivity but preserved specificity of Node-RADS compared with conventional MRI-based nodal assessment, supporting a potential role as a standardized rather than sensitivity-driven tool.[15] In addition, a single-center study applying Node-RADS in BC imaging reported a mismatch between reader agreement and diagnostic accuracy, highlighting that these should be interpreted as related but distinct properties of the score.[16] This variability has also been emphasized in a narrative review summarizing the heterogeneous methodologies and partially conflicting results reported across Node-RADS studies, particularly regarding diagnostic performance and inter-observer agreement.[17]

Cut-off analysis helped characterize the behavior of the Node-RADS score across its categories. On bMRI, lower thresholds were associated with high sensitivity and low specificity, whereas higher thresholds progressively increased specificity and positive likelihood ratios at the expense of sensitivity. A similar pattern was observed on contrast-enhanced CT, where Node-RADS >2 maintained high sensitivity with modest specificity, while higher thresholds achieved very high specificity for most readers. In this study population, a Node-RADS threshold >2 represented the most balanced sensitivity-specificity trade-off and should be considered a hypothesis-generating, population-specific operating point rather than a definitive diagnostic boundary. Like previous studies,[6] we adopted a patient-level analytical approach based on the highest Node-RADS score per patient, reflecting routine clinical practice. The CT analysis was performed in a small, clinically selected subgroup enriched for nodal involvement and should therefore be regarded as exploratory rather than as a direct comparison with bMRI.

From a practical perspective, the combination of high reproducibility and an exploratory Node-RADS >2 operating point suggests that Node-RADS may facilitate a more standardized approach to axillary lymph node assessment within the preoperative bMRI pathway. In this context, scores >2 may help identify patients who could require targeted axillary evaluation according to existing diagnostic workflows.

This study has several limitations. First, the lack of a node-by-node analysis precluded a direct correlation between individual Node-RADS categories and histologically proven metastatic involvement at the single-lymph-node level. Second, the analysis focused on a single, most suspicious lymph node per axilla, which may not fully capture the complexity of nodal disease in patients with multiple abnormal lymph nodes. Third, the CT subgroup was small and clinically selected, as CT was performed only when clinically indicated, typically in patients with suspected advanced disease. This resulted in enrichment for nodal involvement and precluded direct comparison with bMRI. Fourth, the retrospective, single-center design may limit the robustness and generalizability of the results and may have introduced selection bias. Fifth, invasive lobular carcinoma (ILC) was not represented in our cohort because the study population was restricted to IDC-NST and grade 3 DCIS. Recent evidence has shown that the diagnostic performance of Node-RADS on bMRI may vary according to tumor histology, with reduced sensitivity reported in invasive lobular carcinoma, highlighting potential biological constraints of morphology-based nodal assessment.[18]

5. Conclusion

Node-RADS v1.0 showed high reproducibility for preoperative axillary lymph node assessment on bMRI in patients with IDC-NST and grade 3 DCIS, with almost perfect agreement among expert readers and moderate to substantial agreement among junior readers. In this study population, a Node-RADS threshold >2 represented a pragmatic, hypothesis-generating cut-off balancing sensitivity and specificity. This study also explored the application of Node-RADS on contrast-enhanced CT in a small clinically selected subgroup; these findings should be considered preliminary. Overall, Node-RADS may support more standardized nodal assessment in the preoperative bMRI setting.

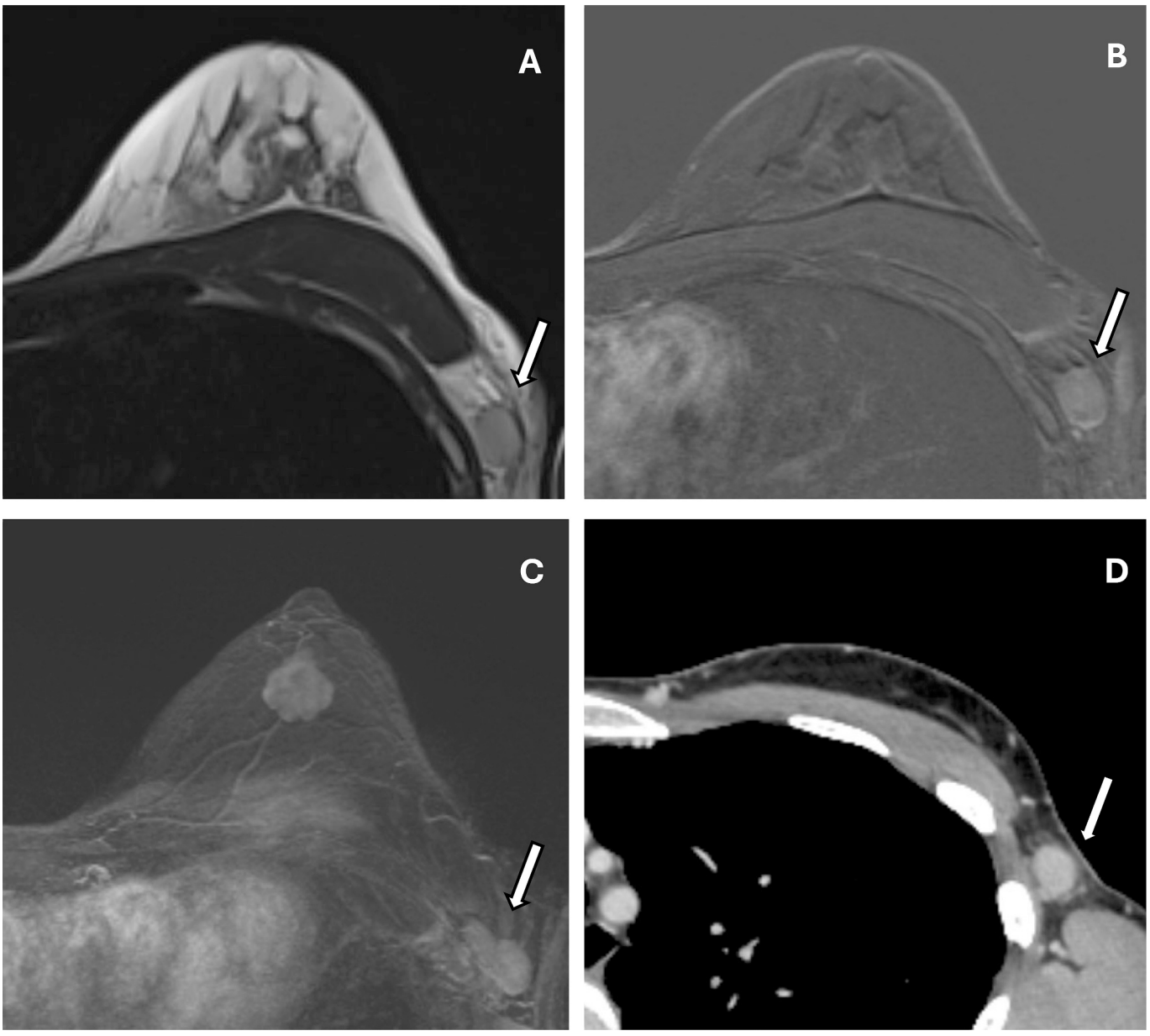


Fig. 7. *Node-RADS MRI* vs *CT* – 39-year-old woman with unifocal grade 2 breast carcinoma of no special type (IBC-NST) in the left breast undergoing preoperative MRI and CT staging. (A) Axial T2-weighted image shows a spherical, enlarged lymph node with smooth borders, without fatty hilum (arrow); (B, C) Axial contrast-enhanced T1-weighted image and MIP projection show homogeneous enhancement (arrow). These features are consistent with Node-RADS 3. (D) Axial venous phase shows the same features as above, but enhance the irregular border, upgrading Node-RADS score from 3 to 4.

Statistic and biometry

Lorenzo Cereser has significant statistical expertise and performed the statistical analysis.

Study subjects or cohorts overlap

None.

Methodology

- Retrospective
- Single-center
- Observational diagnostic imaging study
- Inter-reader agreement and diagnostic performance analysis
- Breast cancer population
- Reference standard: histopathology

CRediT authorship contribution statement

P. Minichetti: Writing – review & editing, Writing – original draft, Supervision, Project administration. **L. Cereser:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization. **A.P. Pace:** Writing – review & editing, Investigation, Data curation, Conceptualization. **F. Francioso:** Writing – review & editing, Software, Data curation. **E. Versienti:** Writing – review & editing, Software, Data curation. **F. Sparascio:** Writing – review & editing, Formal analysis. **R. Girometti:** Writing – review & editing, Methodology, Investigation. **C. Zuiani:** Writing – review & editing, Supervision.

Informed consent

Acquisition of the informed consent was waived by the IRB because of the retrospective design.

Ethical approval

This retrospective study has been approved by the Institutional Review Board (IRB) of the Department of Medicine, University of Udine. Study protocol code: 049/2025. Approval date: 20/02/2025.

Guarantor

The scientific guarantor of this publication is Lorenzo Cereser.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT (OpenAI) to support language editing, improve clarity, and refine the structure of selected sections of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

No funding was received for this study.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgments

The authors acknowledge the use of ChatGPT (OpenAI) for refining the English language and improving the overall readability of the manuscript. However, all data analysis, scientific reasoning, and the full content of the study are entirely the responsibility of the authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinimag.2026.110879.

References

1 Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin 2017 Mar;67(2):93–9. https://doi.org/10.3322/caac.21388.
2 Chang JM, Leung JWT, Moy L, Ha SM, Moon WK. Axillary nodal evaluation in breast cancer: state of the art. Radiology 2020 Jun;295(3):500–15. https://doi.org/10.1148/radiol.2020192534.
3 Elsholtz FHJ, Asbach P, Haas M, et al. Introducing the Node Reporting and Data System 1.0 (Node-RADS): a concept for standardized assessment of lymph nodes in cancer. Eur Radiol. 2021 Aug;31(8):6116–6124. doi:10.1007/s00330-020-07572-4. Epub 2021 Feb 14. Erratum in. Eur Radiol 2021 Sep;31(9):7217. https://doi.org/10.1007/s00330-021-07795-z.
4 Newell MS, Destounis SV, Leung JWT, DeMartini WB, Lee CH, Eby PR. ACR BI-RADS v2025 Manual. Reston, Virginia: American College of Radiology; 2025.
5 Zhong J, Mao S, Chen H, et al. Node-RADS: a systematic review and meta-analysis of diagnostic performance, category-wise malignancy rates, and inter-observer reliability. Eur Radiol 2025 May;35(5):2723–35. https://doi.org/10.1007/s00330-024-11160-1.
6 Pediconi F, Maroncelli R, Pasculli M, et al. Performance of MRI for standardized lymph nodes assessment in breast cancer: are we ready for Node-RADS? Eur Radiol 2024 Dec;34(12):7734–45. https://doi.org/10.1007/s00330-024-10828-y.
7 Sardanelli F, Boetes C, Borisch B, et al. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. Eur J Cancer 2010 May;46(8):1296–316. https://doi.org/10.1016/j.ejca.2010.02.01.
8 Fitzgibbons PL, Connolly JL, Bose S, et al. Protocol for the Examination of Specimens From Patients With Invasive Carcinoma of the Breast. Northfield (IL): College of American Pathologists; 2020.
9 De Boer M, van Dijck JA, Bult P, Borm GF, Tjan-Heijnen VC. Breast cancer prognosis and occult lymph node metastases, isolated tumor cells, and micrometastases. J Natl Cancer Inst 2010 Mar 17;102(6):410–25. https://doi.org/10.1093/jnci/djq008.
10 Biganzoli L, Cardoso F, Beishon M, et al. The requirements of a specialist breast centre. Breast 2020 Jun;51:65–84. https://doi.org/10.1016/j.breast.2020.02.003.
11 Benomar A, Zarour E, Létourneau-Guillon L, Raymond J. Measuring interrater reliability. Radiology 2023 Dec;309(3):e230492. https://doi.org/10.1148/radiol.230492.
12 Geijer M, Båth M, Wessman C. Some common statistical methods for assessing rater agreement in radiological studies. Acta Radiol 2025 Jun;66(6):675–83. https://doi.org/10.1177/02841851251319666.
13 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159–74.
14 Lu F, Zhao Y, Wang Z, Feng N. Diagnostic accuracy of node-RADS for the detection of lymph node invasion: a systematic review and meta-analysis. Eur Radiol 2025 Aug;35(8):4429–37. https://doi.org/10.1007/s00330-025-11387-6.
15 Liao J, Liu C, Huang X, et al. Evaluating the clinical utility of node-RADS 1.0 for axillary lymph node assessment in breast cancer: comparison with conventional MRI diagnosis. Br J Radiol 2025 Dec 6:tqaf304. https://doi.org/10.1093/bjr/tqaf304.
16 Çubuk HS, Erkent M, Yağdıran B, Yıldırım S. Predictive value of node-RADS scoring system for axillary lymph node metastasis in breast cancer. BMC Med Imaging 2026 Jan 28. https://doi.org/10.1186/s12880-026-02181-x.
17 Parillo M, Quattrocchi CC. Node reporting and data system 1.0 (Node-RADS) for the assessment of oncological patients' lymph nodes in clinical imaging. J Clin Med 2025 Jan 5;14(1):263. https://doi.org/10.3390/jcm14010263.
18 Kim HJ, Chae EY, Eom HJ, et al. Node reporting and data system evaluation of axillary nodes in invasive ductal and lobular carcinoma. Radiology 2025 Sep;316(3):e243823. https://doi.org/10.1148/radiol.243823.