



Please summarize the BI-RADS classification, benign/malignant, and brief treatment recommendations based on the following ultrasound diagnostic results. The following is a description of breast ultrasound:

"Examination findings:

Both breasts have uneven echogenicity and disordered structure, with no mass or abnormal blood flow signals observed in the left breast.

In the right breast, multiple low-to-no echogenic nodules are found, with the largest one located at the 3 o'clock direction about 0.9cm from the nipple in the glandular layer, measuring approximately 0.7×0.3cm, having clear and regular boundaries, and no evident blood flow signals within.

No abnormally enlarged lymph nodes are detected in both axillae."



Based on the ultrasound findings, the largest nodule in the right breast is likely a BI-RADS category 4 lesion, specifically TI-RADS 4B, with a moderate to high probability of malignancy. However, without further information or biopsy, it is unclear if the nodules are benign or malignant.

Treatment recommendations will depend on the final diagnosis. If the nodule is confirmed to be malignant, treatment options may include surgical excision, radiation therapy, and/or systemic therapy. If the nodules are benign, observation or minimally invasive procedures, such as ultrasound-guided biopsy or monitoring with follow-up imaging, may be appropriate.

It is important to note that any abnormal breast findings should be promptly evaluated by a healthcare provider to ensure early detection and successful treatment of breast cancer.



**Figure S1** The procedure for posing questions to ChatGPT and gathering its responses. BI-RADS, Breast Imaging Reporting and Data System; TI-RADS, Thyroid Imaging Reporting and Data System.

**Table S1** Evaluation criteria for the performance of ChatGPT in generating breast cancer diagnosis reports

Criterion	Range	Description
Structure and organization (S&O)	1 to 5	Rate the structure of the report
Professional terminology and expression (PTE)	1 to 5	Rate the use of professional terminology in the report
BI-RADS classification	1 to 5	Rate the diagnostic accuracy of the report
Malignancy diagnosis (MD)	1 to 5	Rate the accuracy of malignancy diagnosis
Treatment recommendations (TR)	1 to 5	Rate the appropriateness of treatment recommendations
Clarity and comprehensibility (C&C)	1 to 5	Rate the clarity of expression in the report
Likelihood of being written by a physician (LWBP)	1 to 5	Rate the likelihood that the report was written by a doctor
Ultrasound doctor AI acceptance (UDAIA)	1 to 5	Rate the acceptance of AI-generated report by ultrasound doctors
Overall evaluation (OE)	1 to 5	Provide an overall rating for the entire report

For each criterion, a score of 1 represents the lowest rating (i.e., very poor, completely incorrect), while a score of 5 represents the highest rating (i.e., excellent, completely correct).

**Table S2** The score table for assessing the generated report.

Attribute	Value
ID	
Gender (male/female)	
Age (years)	
Reason for consultation	
Description	
Ultrasound conclusion	
Structure and organization	
Professional terminology and expression	
BI-RADS classification	
Malignancy diagnosis	
Treatment recommendations	
Clarity and comprehensibility	
Likelihood of being written by physician	
Ultrasound doctor AI acceptance	
Overall evaluation	
Structure and organization	

BI-RADS, Breast Imaging Reporting and Data System; AI, artificial intelligence.