e-ISSN: 3023-6487

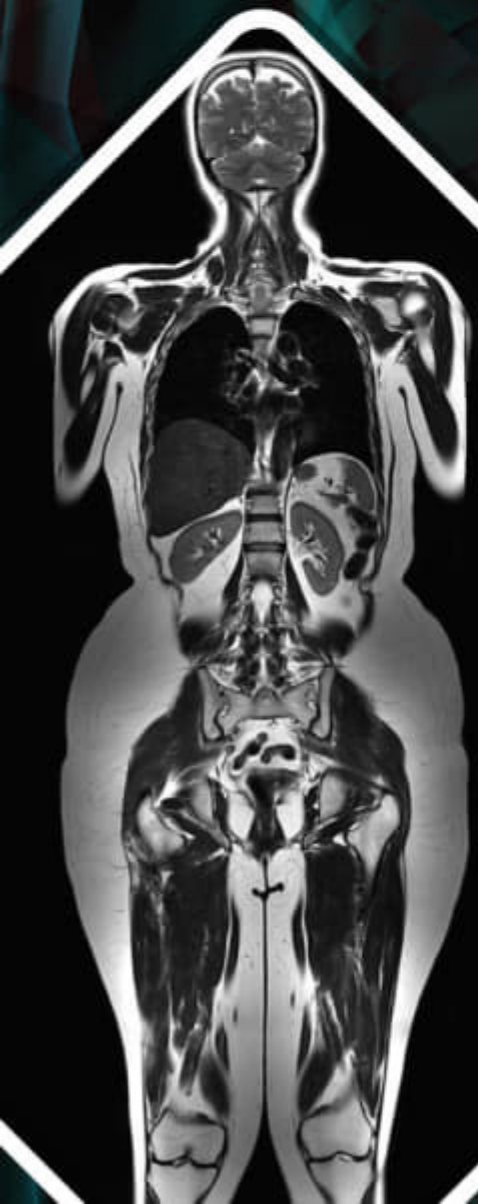
JRM

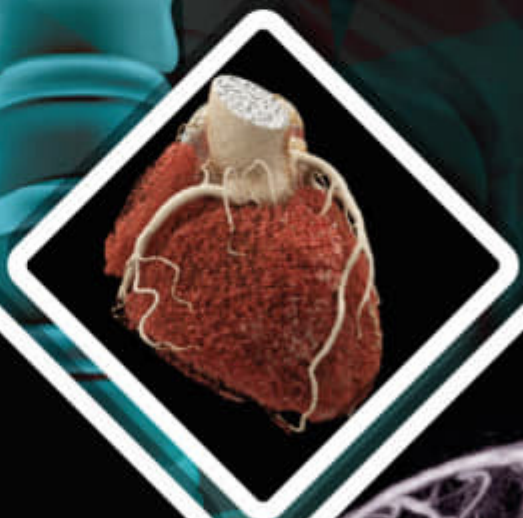
Journal of Radiology in Medicine

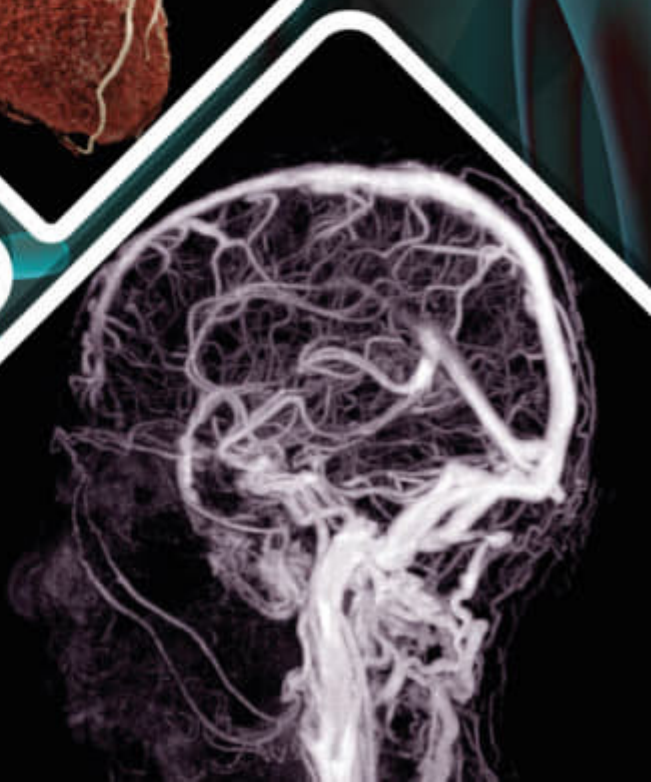
Volume: 2

Issue: 4

Year: 2025







MHA
MEDIHEALTH ACADEMY

Journal of Radiology in Medicine

Editorial Board

EDITORS-IN-CHIEF

EDITORIAL BOARD

ENGLISH LANGUAGE EDITOR

STATISTICS EDITOR

LAYOUT EDITOR

Journal of Radiology in Medicine

Volume: 2 Issue: 4 Year: 2025

Contents

ORIGINAL ARTICLES

REVIEW

Journal of Radiology Medicine in

Original Article

Doi: 10.51271/JRM-0036

Management of thyroid nodules initially classified as benign: ultrasonographic features as a criterion for repetition of fine-needle aspiration

Alev Dolu[1], Özlem Özkale Yavuz[2], Eşref Umut Özyer[3]

[1]Department of Radiology, Acıbadem Maslak Hospital, İstanbul, Turkiye
[2]Department of Radiology, Ankara Bilkent City Hospital, Ankara, Turkiye
[3]Department of Radiology, Medicana International Ankara Hospital, Ankara, Turkiye

Received: 23.06.2025 ◆ **Accepted**: 01.08.2025 ◆ **Published**: 15.10.2025

Cite this article: Dolu A, Özkale Yavuz Ö, Özyer EU. Management of thyroid nodules initially classified as benign: ultrasonographic features as a criterion for repetition of fine-needle aspiration. *J Radiol Med*. 2025;2(4):59-63.

Corresponding Author: Alev Dolu, alewdolu@gmail.com

ABSTRACT

Aims: Nodular thyroid disease is highly prevalent in the adult population. The primary objective in the management of thyroid nodules is to distinguish malignant lesions from benign ones. Following ultrasound (US)-guided fine-needle aspiration (FNA), malignancy is detected in only 3% to 12% of cases. However, in clinical practice, false-negative rates ranging from 1.3% to 11.5% in nodules classified as benign can introduce uncertainty into the management process. Early identification of distinct ultrasonographic features associated with malignancy may not only reduce unnecessary repeat FNA procedures but also help prevent diagnostic delays caused by false-negative results.

Methods: A retrospective review was performed of 513 nodules that underwent repeat FNA biopsies within 6 to 12 months over a 5-year period. US features (number and size of nodules, growth in nodule size, echotexture, echogenicity, presence of microcalcifications, cervical lymphadenopathy, and lymphadenopathy progression) were compared with cytology results to determine which sonographic findings warrant repeat biopsy and which support clinical follow-up.

Results: Repeat FNA was performed on 513 thyroid nodules initially classified as benign, with malignancy detected in 15 cases (2.9%). US features significantly associated with increased malignancy risk included microcalcifications, hypoechoic solitary nodules, enlarging hypoechoic solitary nodules, accompanying lymphadenopathy, and progressive lymphadenopathy. In contrast, nodule number, size, growth, and echotexture alone were not significant predictors of malignancy. US characteristics such as cystic composition, isoechogenicity, and hyperechogenicity were more commonly observed in benign nodules.

Conclusion: This study recommends repeat FNA for hypoechoic solitary nodules, enlarging hypoechoic solitary nodules, nodules with microcalcifications, and those associated with lymphadenopathy. In contrast, repeat FNA is not advised for cystic nodules or solid nodules without hypoechogenicity.

Keywords: Cone-beam computed tomography, mandible, mental foramen, premolar, radiology

INTRODUCTION

Nodular thyroid disease is very common among adults. The use of ultrasound (US) as a routine part of thyroid exams has led to detection rates as high as 68%.[1,2] The main purpose in the management of thyroid nodules is to differentiate malignancy. US-guided fine-needle aspiration (FNA) is the gold standard diagnostic tool to determine the management of the thyroid nodule.[3] After FNA, most nodules are diagnosed as benign, and only about 3–12% are found to be malignant.[4] Guidelines recommend clinical follow-up when a nodule is initially diagnosed as benign.[1,5] However, false-negative results ranging from 1.3% to 11.5% can make it difficult to manage benign nodules in clinical practice.[6,7] This study aims to clarify the necessity of FNA by identifying which US features are associated with an increased risk of malignancy in thyroid nodules initially diagnosed as benign. To this end, the US characteristics of the nodules and the results of repeated FNA were evaluated. This way, malignant nodules missed due to false-negative results can be detected with repeat FNA, and unnecessary repeat FNAs can be reduced.

METHODS

Ethics

All procedures complied with the Declaration of Helsinki and local ethical standards. This study has been approved by the Researches Ethics Committee of Başkent University Faculty of Medicine and Health Sciences (Date: 28.01.2015, Decision No: KAl 5/16).

Patients and Data Acquisition

A retrospective review was performed on patients who initially received a benign diagnosis and underwent repeat FNA between January 2009 and February 2014. Patients included in the study were those who underwent FNA of the same thyroid nodule, with detailed US reports available prior to both the initial and subsequent procedures. During the study period, a total of 531 nodules underwent repeat FNA. After excluding 18 nodules with inadequate cytology (n=11) or suspicious pathology (n=7) that did not undergo surgery or further FNA, 513 nodules were included in the final analysis. Exclusion criteria were: (i) nodules lacking adequate US documentation before FNA, (ii) patients with multinodular goiters where nodule correspondence between procedures was uncertain, (iii) patients who had repeat FNA from different nodules, and (iv) nodules with inadequate or suspicious cytology not followed by surgery or repeat aspiration.

Patient demographics, including gender, age and the interval between FNAs, were obtained from hospital records. US examinations were performed using high-resolution US machines by several radiologists with a minimum of 5 years of experience in thyroid imaging. Nodule characteristics were extracted from radiology reports and included: (i) size (<1 cm or ≥1 cm), (ii) internal structure (solid, cystic, or mixed), (iii) echogenicity (hypoechoic, isoechoic, hyperechoic, or mixed), (iv) number of nodules (solitary or multinodular), (v) presence of microcalcifications, (vi) associated cervical lymphadenopathy, defined as short-axis diameter >1 cm, and (vii) growth in lymph nodes, defined as ≥20% increase in at least two dimensions.[8]

Repeat FNA cytology results were classified into four categories: benign, suspicious, malignant, and inadequate specimens.[9] Nondiagnostic aspirates were excluded from final analysis unless repeated and adequate. The interval between the initial and repeat FNA, as well as any changes in sonographic features or cytology category, were recorded for all included cases.

Technique

All FNAs were performed under US guidance using a 21-gauge needle, with 2–3 passes per nodule. In cystic nodules with a solid component, the solid portion was targeted for aspiration. The aspirates were processed using conventional smear techniques and stained with Papanicolaou and May-Grünwald-Giemsa stains.

Statistical Analysis

For the evaluation of data, descriptive statistics were used. Numerical variables were expressed as mean±standard deviation, median, minimum, and maximum values, while categorical variables were presented as frequencies and percentages. The rates of malignancy associated with US findings were analyzed using cross-tabulations and presented as counts and percentages. To assess the diagnostic performance of US features in predicting malignancy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. For statistically significant findings, Fisher's exact test and the Fisher-Freeman-Halton test were used. Variables that were significant in univariate analyses were included in a multivariate logistic regression using the Backward LR elimination method to determine the best model for predicting malignancy. Results of the logistic regression analysis were reported with odds ratios (ORs), p-values, and 95% confidence intervals (CIs). A p-value of <0.05 was considered statistically significant. All statistical analyses were performed using the SPSS 21.0 for Windows software package.

RESULTS

Of the 513 nodules initially diagnosed as benign, 15 (2.9%) were identified as malignant, while 498 (97.1%) remained benign. The mean interval between FNA procedures was 237 days (range: 181 to 341 days). The study group included 387 female and 126 male patients aged between 17 to 87 years (median, 62±15.4 years). Six nodules in male and 9 nodules in female group were malignant. Statistical analysis revealed that gender had no significant effect on malignancy (p=0.218).

The characteristics of the nodules in the study population and their US findings are summarized in **Table 1**.

Nodule size distribution showed 34.3% of nodules were ≤1 cm and 65.7% were >1 cm, with similar malignancy rates of 3.4% and 2.7%, respectively. There was no significant difference in malignancy when infracentimetric nodules were compared with supracentimetric nodules (p=0.783).

Solid nodules had higher rate of malignancy (3.5%) than mixed nodules (1.4%) however it was not significant (p=0.361). Cystic echo-structure was a significant feature for benign cytology (p<.001).

Among echogenicity types, hypoechoic nodules represented 48.9% of the total and showed the highest malignancy rate at 5.2% (p=0.021). Isoechoic and hyperechoic nodules were all benign (p<0.001), whereas nodules with mixed echogenicity had a malignancy rate of 2.1%.

Regarding nodule number, solitary nodules showed a higher malignancy rate (7.2%) than multiple nodules (1.9%) (p=0.014). Hypoechoic solitary nodules had a significantly higher rate of malignancy than hypoechoic nodules in multinodular thyroid gland or non-hypoechoic solitary nodules (p<0.005). Nodule growth was observed in 18.7% of the nodules, with a malignancy rate of 8.3%, which was higher compared to nodules without growth (1.7%). Among hypoechoic nodules, those with growth demonstrated an even higher malignancy rate of 23.3%, compared to 1.7% in hypoechoic nodules without growth. These findings suggest that nodule growth, particularly in hypoechoic nodules, is strongly associated with an increased risk of malignancy (p<0.005).

Table 1. Characteristics of thyroid nodules by benign and malignant status

Nodule feature	Benign, n (%)	Malignant, n (%)	Total, n (%)
All nodules	498 (97.1%)	15 (2.9%)	513 (100%)
Gender			
Male	120 (95.2%)	6 (4.8%)	126 (24.6%)
Female	378 (97.7%)	9 (2.3%)	387 (75.4%)
Nodule number			
Solitary	90 (92.8%)	7 (7.2%)	97 (18.9%)
Multiple	408 (98.1%)	8 (1.9%)	416 (81.1%)
Echo-structure			
Solid	385 (96.5%)	14 (3.5%)	399 (77.8%)
Cystic	43 (100%)	0 (0%)	43 (8.4%)
Mixed	70 (98.6%)	1 (1.4%)	71 (13.8%)
Echogenicity			
Hypoechoic	238 (94.8%)	13 (5.2%)	251 (48.9%)
Isoechoic	76 (100%)	0 (0%)	76 (17.5%)
Hyperechoic	90 (100%)	0 (0%)	90 (14.8%)
Mixed echogenicity	94 (97.9%)	2 (2.1%)	96 (18.7%)
Nodule size			
≤1 cm	170 (96.6%)	6 (3.4%)	176 (34.3%)
>1 cm	328 (97.3%)	9 (2.7%)	337 (65.7%)
Nodule growth			
Present	88 (91.7%)	8 (8.3%)	96 (18.7%)
Absent	410 (98.3%)	7 (1.7%)	417 (81.3%)
Microcalcification			
Present	127 (93.4%)	9 (6.6%)	136 (28.5%)
Absent	371 (98.4%)	6 (1.6%)	377 (73.5%)
Lymphadenopathy			
Present	25 (83.3%)	5 (16.7%)	30 (5.8%)
Absent	473 (97.9%)	10 (2.1%)	483 (94.2%)
Lymph node growth			
Present	17 (85%)	3 (15%)	20 (3.9%)
Absent	481 (97.6%)	12 (2.4%)	493 (96.1%)
Subgroups of hypoechoic nodules			
Solitary hypoechoic	45 (86.5%)	7 (13.5%)	52 (10.1%)
Multiple hypoechoic	453 (98.3%)	8 (1.7%)	461 (89.9%)
Solitary hypoechoic with growth	23 (76.7%)	7 (23.3%)	30 (5.8%)
Solitary hypoechoic without growth	475 (98.3%)	8 (1.7%)	483 (94.2%)

Notes: Percentage distribution of ultrasound findings and gender in benign and malignant groups

The presence of microcalcifications was detected in 28.5% of nodules and correlated with an increased malignancy rate of 6.6%, compared to 1.6% in nodules without microcalcifications (p<0.005). Cervical lymphadenopathy was identified in 5.8% of cases, with a malignancy rate of 16.7%, while lymph node growth was noted in 3.9% of nodules and associated with malignancy in 15% of these cases (p<0.005).

Sensitivity, specificity, PPV, NPV, p values and OR of the significant US findings are given in **Table 2**.

DISCUSSION

FNA is the most effective preoperative diagnostic method for determining thyroid malignancy.[1,4] Malignant or suspicious nodules and follicular neoplasm are proposed to surgery while benign nodules are followed-up.[10] However, benign cytology does not rule out the possibility of malignancy and FNA repetition has been advised for the detection of false negative cases.[7,11,12] Repetition was found to increase the accuracy, sensitivity and specificity of the FNA and reduce the false negative results from 6.7% to 0.8%. Our study revealed that 2.9 % of the nodules had a false negative initial FNA result and a final diagnosis of malignancy was made with FNA repetition after 6 months. This result was compatible with previous reports.[7,13]

Previous studies have demonstrated that repeat FNA can be effective in detecting occult malignancies.[7,11] However, it has also been reported that in nodules initially diagnosed as benign, repeat cytology yields benign results in over 90% of cases, suggesting limited additional diagnostic value.[4,11-14] Considering the low malignancy rates, repeat FNA should not be entirely disregarded in follow-up protocols. Nevertheless, to avoid subjecting a large patient population to unnecessary procedures, repeat FNA indications should be applied more selectively.

The most practical way to restrict the indications for repetition seems to be classifying the nodules according to the US features. In the population subject to FNA repetition, US features can be used to identify nodules with a higher probability of malignancy.[1,2] Thus, diagnosis of masked malignancy can be achieved without exposing a large proportion of patients to repeat FNAs.

Primary FNA indications based on ultrasonographic features have been clearly defined in the literature.[1,2] US

Table 2. Logistic regression analysis of significant ultrasound findings and risk of malignancy

Ultrasound finding	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	p value	Odds ratio
Hypoechoic nodule	86.7	35.8	5.2	99.5	0.021	3.2
Solitary nodule	46.7	81.9	7.2	98.1	0.014	3.96
Hypoechoic solitary nodule	46.7	91	13.5	98.3	0.001	5.01
Nodule growth present	53.3	82.3	8.3	98.3	0.003	5.31
Growing hypoechoic solitary nodule	46.7	95.4	23.3	98.3	<0.001	13.3
Microcalcification	60	74.5	6.6	98.4	0.005	4.86
Lymphadenopathy	33.3	95	16.7	97.9	0.001	10.96
Lymph node growth	20	96.6	15	97.5	0.017	–

Notes: Statistically significant results (p<0.05), Abbreviations: PPV: Positive predictive value, NPV: Negative predictive value

plays a key role in identifying nodules with an increased risk of malignancy. Solid hypoechoic structure, absence of a hypoechoic halo, irregular margins, a taller-than-wide shape, microcalcifications, marked intranodular vascularity, and accompanying cervical lymphadenopathy are among the main features associated with a higher likelihood of malignancy.[2,15] However, the significance of these US findings in distinguishing between benign and malignant nodules has been reported to vary across different studies.[16]

This study found that solid echo-structure was not significantly associated with malignancy, whereas cystic echo-structure was strongly correlated with benign cytology (p<0.001). Similarly, hypoechogenicity alone was a significant US feature for malignancy, while isoechogenicity and hyperechogenicity were significantly associated with benign cytology (p<0.001). Although some previous studies have reported hypoechogenicity as a significant predictor of malignancy, others have not.[12,13] Therefore, cystic nodules and non-hypoechoic solid nodules without additional suspicious US features may not require repeat FNA during follow-up.

Hypoechoic solitary nodules and growing hypoechoic solitary nodules both demonstrated a significantly increased risk of malignancy. Hypoechoic solitary nodules increased the specificity for malignancy to 91%. Although they comprised only 10.1% of all nodules in the study population, they accounted for 46.7% of malignant nodules and only 9% of benign nodules. When accompanied by growth, specificity further increased to 95.4%. These findings suggest that hypoechoic solitary nodules, and especially those that are growing, should be closely monitored with repeat FNA biopsy to avoid missing malignancies.

Consistent with previous studies, our findings demonstrated that the presence of microcalcifications is highly specific and significantly associated with an increased risk of malignancy.[1,2,12,13] Moreover, our results indicate that repeat FNA should be considered mandatory during the follow-up of initially benign nodules exhibiting microcalcifications to ensure early detection of malignant transformation.

Our study also showed that there was no significant difference among infracentimetric and supracentimetric nodules regarding malignancy in FNA repetition. US features of echogenicity, structure and microcalcification were superior criteria for FNA repetition compared to nodule size. This aligns with previous research showing no significant difference in size between benign and malignant nodules.[17,18] In contrast to nodule size, growth was significantly more common in malignant (53.3%) than benign nodules (17.7%) and was found to be highly specific for malignancy if present in hypoechoic solitary nodules. Like previous studies, our study showed that nodule growth without other specific US features was not a significant factor in increased malignancy risk.[5,12,13]

In our study, the presence of lymphadenopathy demonstrated high specificity (95%) and a significant association with malignancy (p=0.001), indicating its strong predictive value. Lymph node growth also showed high specificity (96.6%) and a significant correlation with malignancy (p=0.017), although its sensitivity was lower. These findings support the role of lymphadenopathy as an important US feature in evaluating malignancy risk in thyroid nodules. Similar results have been reported in the literature, emphasizing that lymph node enlargement and structural changes are closely linked to malignancy.[19,20] Therefore, patients with detected lymphadenopathy or lymph node growth should be closely monitored and considered for repeat FNA when necessary.

Limitations

US evaluations were performed by different radiologists, which may have introduced interobserver variability in imaging findings. Furthermore, certain sonographic features strongly associated with malignancy—such as margin irregularity, taller-than-wide shape, and intranodular vascularity—were not assessed due to inconsistent availability in retrospective records. Malignancy diagnoses were based solely on FNA cytology, and histopathological confirmation through surgical specimens was not available for all cases. Lastly, the low number of malignant cases (n=15; 2.9%) limited the statistical power, necessitating cautious interpretation of the results.

CONCLUSION

This study suggests that FNA repetition to all nodules with initially benign cytology is not justified. FNA repetition should not be performed in cystic, isoechoic and hyperechoic nodules. Solitary hypoechoic nodules, nodules with microcalcification and accompanying cervical lymphadenopathy have increased risk of malignancy and should always undergo FNA repetition within 12 months despite initially benign cytological diagnosis.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study has been approved by the Researches Ethics Committee of Başkent University Faculty of Medicine and Health Sciences (Date: 28.01.2015, Decision No: KAI 5/16).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Tessler FN, Middleton WD, Grant EG, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *J Am Coll Radiol*. 2017;14(5):587-595. doi:10.1016/j.jacr.2017.01.046
2. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016;26(1):1-133. doi:10.1089/thy.2015.0020
3. Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JW, Wiersinga W. European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *Eur J Endocrinol*. 2006;154(6):787-803. doi:10.1530/eje.1.02158
4. Glynn N, Hannon MJ, Lewis S, et al. Utility of repeat cytological assessment of thyroid nodules initially classified as benign: clinical insights from multidisciplinary care in an Irish tertiary referral centre. *BMC Endocr Disord*. 2016;16(1):45. doi:10.1186/s12902-016-0125-7
5. Rosario PW, Calsolari MR. What is the best criterion for repetition of fine-needle aspiration in thyroid nodules with initially benign cytology? *Thyroid*. 2015;25(10):1115-1120. doi:10.1089/thy.2015.0253
6. Singh Ospina N, Sebo TJ, Morris JC, Castro MR. The value of repeat thyroid fine-needle aspiration biopsy in patients with a previously benign result: how often does it alter management? *Thyroid*. 2015;25(10):1121-1126. doi:10.1089/thy.2015.0146
7. Hwang SH, Sung JM, Kim EK, Moon HJ, Kwak JY. Imaging-cytology correlation of thyroid nodules with initially benign cytology. *Int J Endocrinol*. 2014;2014:491508. doi:10.1155/2014/491508
8. van den Brekel CJ, Castelijns JA. The size of lymph nodes in the neck on sonograms as a radiologic criterion for metastasis: how reliable is it? *AJNR Am J Neuroradiol*. 1998;19:695-700.
9. The Papanicolaou Society of Cytopathology Task Force on Standards of Practice. Guidelines of the Papanicolaou Society of Cytopathology for fine-needle aspiration procedure and reporting. *Diagn Cytopathol*. 1997;17(4):239-247. doi:10.1002/(SICI)1097-0339(199710)17:4<239:AID-DC1>3.0.CO;2-7
10. Cooper DS, Doherty GM, Haugen BR, et al. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2006;16(2):109-142. doi:10.1089/thy.2006.16.109
11. Furlan JC, Bedard YC, Rosen IB. Single versus sequential fine-needle aspiration biopsy in the management of thyroid nodular disease. *Can J Surg*. 2005;48:12-18.
12. Shin JH, Han BK, Ko K, Choe YH, Oh YL. Value of repeat ultrasound-guided fine-needle aspiration in nodules with benign cytological diagnosis. *Acta Radiol*. 2006;47(5):469-473. doi:10.1080/02841850600635921
13. Rosario PW, Purisch S. Ultrasonographic characteristics as a criterion for repeat cytology in benign thyroid nodules. *Arq Bras Endocrinol Metabol*. 2010;54(1):52-55. doi:10.1089/thy.2015.0253
14. Chehade JM, Silverberg AB, Kim J, Case C, Mooradian AD. Role of repeated fine-needle aspiration of thyroid nodules with benign cytologic features. *Endocr Pract*. 2001;7(4):237-243. doi:10.4158/EP.7.4.237
15. Rago T, Vitti P, Chiovato L, et al. Role of conventional ultrasonography and color flow-doppler sonography in predicting malignancy in 'cold' thyroid nodules. *Eur J Endocrinol*. 1998;138(1):41-46. doi:10.1530/eje.0.1380041
16. Gharib H. Changing concepts in the diagnosis and management of thyroid nodules. *Endocrinol Metab Clin North Am*. 1997;26(4):777-800. doi:10.1016/s0889-8529(05)70282-6
17. Frates MC, Langer JE. Biopsy of thyroid nodules: comparison of three sets of guidelines. *AJR Am J Roentgenol*. 2010;195(6):W472. doi:10.2214/AJR.10.4755
18. Butros R, Boyvat F, Ozyer U, et al. Management of infracentimetric thyroid nodules with respect to ultrasonographic features. *Eur Radiol*. 2007;17(5):1358-1364. doi:10.1007/s00330-006-0413-0
19. Yoo RE, Kim JH, Bae JM, et al. Ultrasonographic indeterminate lymph nodes in preoperative thyroid cancer patients: malignancy risk and ultrasonographic findings predictive of malignancy. *Korean J Radiol*. 2020;21(5):598-604. doi:10.3348/kjr.2019.0755
20. Chasen NN, Wang JR, Gan Q, Ahmed S. Imaging of cervical lymph nodes in thyroid cancer: ultrasound and computed tomography. *Neuroimaging Clin N Am*. 2021;31(3):313-326. doi:10.1016/j.nic.2021.04.002

Journal of Radiology Medicine in

Original Article

Doi: 10.51271/JRM-0037

Integration of CT radiomics and clinical factors to predict pathological response to FLOT chemotherapy in gastric cancer

Görkem Özdemir[1], Çağatay Neftali Tulu[2], Ömer Işık[2], Tolga Ölmez[1], Alper Sözütek[1]

[1]Department of Gastroenterological Surgery, Adana City Training and Research Hospital, Adana, Turkiye
[2]Department of Software Engineering, Adana Alparslan Türkeş Science and Technology University, Adana, Turkiye

Received: 04.08.2025 ◆ **Accepted**: 29.09.2025 ◆ **Published**: 15.10.2025

Cite this article: Özdemir G, Tulu ÇN, Işık Ö, Ölmez T, Sözütek A. Integration of CT radiomics and clinical factors to predict pathological response to FLOT chemotherapy in gastric cancer. *J Radiol Med*. 2025;2(4):64-70.

Corresponding Author: Görkem Özdemir, dr.gorkemozdemir@yahoo.com.tr

ABSTRACT

Aims: Neoadjuvant chemotherapy (NAC) has become a standard of care for locally advanced gastric cancer (LAGC), but a significant portion of patients do not respond to treatment, leading to unnecessary toxicity and delayed definitive surgery. Non-invasive, predictive tools are urgently needed to identify these non-responders prior to treatment. This study aimed to develop and validate a predictive model using an integrated approach of clinical factors and radiomic features extracted from pretreatment computed tomography (CT) scans to classify patients with LAGC as responders or non-responders to the neoadjuvant FLOT protocol.

Methods: This retrospective study utilized 163 patients. A comprehensive feature set, including 11 patient biomarkers and 101 radiomic features, was used to train various machine learning algorithms via the pyCaret library. The Light Gradient Boosting Machine (LightGBM) was identified as the top-performing model.

Results: On the internal validation, the LightGBM model achieved an accuracy of 80% with an AUC of 0.7926. Key features contributing to the model's prediction included albumin, CA 19-9, age, and specific radiomic features such as Robust Mean Absolute Deviation. However, the model's performance decreased in the independent external validation, with an accuracy of 54.59% and an AUC of 0.5409, highlighting challenges in generalizability.

Conclusion: This study demonstrates the potential of an integrated clinical and radiomic approach using a LightGBM model to non-invasively predict treatment response to FLOT in LAGC. While the model showed strong internal performance, the decreased performance on an external validation set underscores the critical need for larger, multi-institutional, prospective studies with standardized data collection to establish clinical generalizability and widespread utility.

Keywords: Gastric cancer, neoadjuvant chemotherapy, FLOT protocol, radiomics, predictive model, machine learning

INTRODUCTION

Gastric cancer is a serious global health challenge, ranking among the leading causes of cancer-related mortality worldwide, with its incidence being particularly high in Asia and Eastern Europe.[1] Locally advanced gastric cancer (LAGC) presents a substantial clinical challenge, characterized by a heterogeneous prognosis and a significant risk of recurrence even after aggressive therapeutic interventions.[2] The standard of care for resectable LAGC has evolved to a multimodal approach that incorporates neoadjuvant chemotherapy (NAC) as a pivotal component. The primary goals of NAC include improving surgical outcomes, achieving tumor downstaging, eradicating micrometastases, and ultimately enhancing resection rates and long-term patient survival.[2-4]

Among various regimens, the FLOT (Fluorouracil, Leucovorin, Oxaliplatin, and Docetaxel) protocol has emerged as a preferred NAC regimen, consistently demonstrating superior survival outcomes compared to other approaches in randomized clinical trials.[3,4] However, a significant proportion of patients exhibit a poor or non-existent response to FLOT, rendering them susceptible to unnecessary toxicities and detrimental delays in definitive

surgical intervention.[5] Identifying these non-responders before treatment initiation is of paramount importance for optimizing patient care. Such a predictive tool would enable clinicians to personalize therapeutic strategies, potentially directing non-responsive patients to alternative therapies or appropriate clinical trials, thereby optimizing outcomes and reducing treatment-related morbidity.[6,7]

Traditional methods for assessing treatment response are broadly categorized into radiological and pathological approaches. Radiological assessments, performed non-invasively, have historically relied on methods such as the response evaluation criteria in solid tumors (RECIST). However, RECIST, which primarily focuses on macroscopic tumor size changes, has demonstrated significant limitations in accurately predicting patient outcomes. For instance, it has been shown to fail in identifying a substantial portion of patients who progress to metastatic disease and cannot reliably stratify patients by long-term survival.[8,9] A more effective radiological approach involves TNM re-staging, which compares the baseline CT stage to the post-NAC pathological stage (downstaging). This method has been demonstrated to successfully stratify patients into distinct groups with different long-term survival rates and is recommended for routine clinical use.[10] Pathological evaluation of the surgical specimen provides the most definitive assessment of treatment response. Key pathological indicators include pathological complete response (pCR) and graded histological regression (GHR). Achieving pCR after NAC is a promising prognostic factor, strongly correlated with significantly improved overall survival and disease-free survival compared to non-pCR patients.[3,11] The survival benefit is particularly pronounced in patients who also achieve ypN0 status (no residual nodal disease).[5] Furthermore, a GHR of ≥50% is considered a reliable criterion for predicting a favorable prognosis and can inform postoperative chemotherapy decisions.[3] The post-NAC pathological stage (ypTNM) and Lauren classification also serve as independent prognostic factors.[3] Indeed, response to NAC is widely considered the strongest independent predictor of survival after curative resection for gastric cancer.[12]

Given the inherent limitations of conventional radiological downsizing methods and the pressing need for more precise, non-invasive predictive tools, there is growing interest in integrating advanced computational approaches. Radiomics, a promising field, extracts high-dimensional quantitative features from standard medical images such as CT scans, which can capture subtle phenotypic characteristics of tumors not discernible by the naked eye.[13,14] When combined with sophisticated machine learning algorithms, radiomics-based models have shown considerable promise in non-invasively predicting NAC response and survival, often with improved accuracy when integrated with clinical data.[14,15] Furthermore, the rise of advanced machine learning algorithms, such as the Light Gradient Boosting Machine (LightGBM), provides a powerful framework for integrating and analyzing these complex, high-dimensional, multi-modal datasets. These models can uncover intricate, non-linear relationships between a multitude of features and the final treatment outcome, surpassing the limitations of traditional statistical methods.[16-18]

Therefore, this study aims to develop and thoroughly validate a promising predictive model using the LightGBM algorithm to accurately classify LAGC patients as responders or non-responders to the neoadjuvant FLOT protocol. By comprehensively integrating patient clinical, demographic, and laboratory data with quantitative features derived from pre-treatment CT scans, we seek to provide a novel, data-driven tool. This tool is intended to improve personalized treatment decision-making, enhance risk stratification, and ultimately optimize the clinical management of LAGC.

METHODS

Ethics

The study was approved by the Adana City Training and Research Hospital Clinical Researches Ethics Committee (Date: 07.11.2024, Decision No: 220) and was conducted in accordance with the Declaration of Helsinki.

Study Design and Patient Cohorts

This retrospective analysis included a total of 163 consecutive patients with LAGC who received neoadjuvant FLOT chemotherapy followed by total or subtotal gastrectomy with D2 lymph node dissection.

The inclusion criteria for the study were as follows: a diagnosis of LAGC, the availability of a pre-treatment (baseline) contrast-enhanced CT examination, and the receipt of the neoadjuvant FLOT protocol followed by surgical resection.

Patients were excluded if they did not complete the FLOT regimen (n=12), did not undergo surgical resection (n=10), or had CT examinations with insufficient image quality (n=20).

The final study population of 163 eligible patients was divided into two distinct cohorts:

Internal cohort: A total of 126 patients were used for model training and internal validation. The training cohort consisted of 46 responders and 54 non-responders, while the internal validation test cohort consisted of 10 non-responders and 16 responders.

External cohort: A total of 37 patients were used for independent external validation. This cohort consisted of 18 responders and 19 non-responders. These patients met the inclusion criteria but their pre-treatment CT scans were performed at external imaging centers, providing a source of independent data for testing the model's generalizability.

Data Collection and Preprocessing

Baseline assessments were conducted for all patients before the initiation of neoadjuvant therapy. Data were collected on patient age, gender, tumor grade, carcinoembryonic antigen (CEA) level, CA 19-9 level, albumin and hemoglobin levels, tumor location, dysphagia, vomiting, and weight loss.

The final pathological response was determined by a pathologist's examination of the surgical specimen using the Ryan Tumor Regression Grade system (RRR). Patients were categorized into two groups for the purpose of this study: responders (partial or complete response, RRR=1) and non-responders (no response, RRR=0).

Radiomics Feature Extraction

Radiomics features were extracted from the pre-treatment contrast-enhanced CT scans. Tumor regions of interest (ROIs) were manually delineated by a single experienced gastroenterological surgeon on all relevant axial slices. It is important to note that a formal assessment of inter-observer and intra-observer variability was not conducted. To facilitate the feature extraction process, these delineated ROIs were used to create a binary mask for each image.

Prior to feature extraction, the CT images and corresponding masks were processed using the pyRadiomics open-source package (version 3.0). No image-specific preprocessing steps, such as intensity normalization or resampling, were applied.

A total of 101 radiomics features were extracted using the default settings of the pyRadiomics package. These features included:

- First-order statistics (e.g., Mean, Skewness, Kurtosis)
- Shape features
- Gray-Level Co-occurrence Matrix (GLCM) features
- Gray-Level Run-Length Matrix (GLRLM) features
- Gray-Level Size Zone Matrix (GLSZM) features
- Gray-Level Dependence Matrix (GLDM) features

These 101 radiomics features were then combined with the 11 patient-specific clinical and laboratory biomarkers, creating a comprehensive feature set of 112 variables for model training and evaluation.

Feature Selection

To mitigate the risk of overfitting due to high dimensionality, we applied a feature selection step prior to model development. A classical recursive feature elimination (RFE) strategy was used, employing LightGBM as the base estimator to iteratively remove features with the lowest importance weights. The selection process was guided by cross-validation performance, and the procedure ultimately identified 21 features (a combination of clinical biomarkers and radiomic descriptors) as the optimal subset for model training. These selected features were then used to construct the predictive models.

Predictive Model Development and Evaluation

A LightGBM model was trained on the internal cohort to classify patients as responders or non-responders. The model's performance was evaluated using standard metrics, including:

Accuracy: The proportion of correctly classified patients.

AUC: The area under the receiver operating characteristic curve.

Recall: The ability of the model to correctly identify responders.

Precision: The proportion of positive identifications that were actually correct.

F1 score: The harmonic mean of precision and recall.

Feature importance was calculated to determine the relative influence of each variable on the model's predictions. The final trained model was then applied to the independent external validation cohort to assess its generalizability.

RESULTS

Performance of the Predictive Models

After training, various machine learning models were evaluated on a dedicated test dataset. The LightGBM model demonstrated the highest overall performance, achieving an accuracy of 0.80 and an AUC of 0.7926. The model's recall was 0.825, its precision was 0.8461, and its F1-score was 0.8354. Performance metrics for all models are summarized in **Table 1**.

Table 1. Performance metrics of the models on the test dataset after training

Model	Accuracy	AUC	Recall	Prec.	F1
Light gradient boosting machine	0.8	0.7926	0.825	0.8461	0.8354
Gradient boosting classifier	0.7538	0.7852	0.7875	0.8077	0.7975
Ada boost classifier	0.7308	0.736	0.7	0.8358	0.7619
Extreme gradient boosting	0.6538	0.7492	0.6625	0.7465	0.702
Extra trees classifier	0.5692	0.6652	0.5125	0.7069	0.5942
Decision tree classifier	0.5462	0.5562	0.5125	0.6721	0.5816
SVM-Linear Kernel	0.6308	0.5875	0.775	0.6739	0.7209
Logistic regression	0.6692	0.6812	0.775	0.7126	0.7425
K neighbors classifier	0.5692	0.6347	0.4375	0.7609	0.5556
Linear discriminant analysis	0.6462	0.6842	0.7	0.7179	0.7089
Quadratic discriminant analysis	0.6615	0.662	0.8	0.6957	0.7442
Dummy classifier	0.3846	0.5	0	0	0
Naive bayes	0.6615	0.7547	0.9	0.6667	0.766

AUC: Area under the curve

Feature Importance Analysis

Following training with the LightGBM method, a feature importance analysis was conducted to identify the most influential variables in the model's decision-making process. The top five features, ranked by their importance scores, were albumin, CA 19-9, age, hemoglobin, and CEA. These are often clinically relevant markers, which add credibility to the model's findings. Other features, including weight loss, grade, and several image diagnostics features, also showed some level of influence. The detailed feature importance scores are presented in **Table 2**.

External Validation Results

The pre-trained LightGBM model was then tested on the independent external validation dataset (n=37). The model achieved an accuracy of 0.5459, an AUC of 0.5409, a recall of 0.5312, and a precision of 0.5666. The F1-score was 0.5483.

Table 2. Feature importances for the LightGBM model after training

Feature	Importance score
Albumin	0.1484
CA 19-9	0.1391
Age	0.1251
Hemoglobin	0.1158
CEA	0.1033
Weight loss	0.0420
Grade	0.0420
Robust mean absolute deviation (first-order)	0.0404
Mean intensity (image diagnostics)	0.0311
Zone variance (GLSZM)	0.0256
Skewness (first-order)	0.0241
Gray level non-uniformity normalized (GLRLM)	0.0218
Dependence non-uniformity normalized (GLDM)	0.0202
Gray level non-uniformity (GLRLM)	0.0186
Gray level non-uniformity (GLSZM)	0.0186
Tumor location	0.0186
Dependence variance (GLDM)	0.0171
Minimum intensity (first-order)	0.0163
Gender	0.0140
Kurtosis (first-order)	0.0093
Energy (first-order)	0.0085

LightGBM: Light Gradient Boosting Machine, CEA: Carcinoembryonic antigen

DISCUSSION

The accurate prediction of pathological response to NAC is a critical step towards personalized medicine in LAGC. Our study aimed to address this challenge by developing a machine learning model, specifically a LightGBM classifier, which integrates both clinical and quantitative radiomic features from pre-treatment CT scans to predict treatment response to the neoadjuvant FLOT regimen. The model demonstrated promising predictive performance in internal validation, suggesting potential value of a multi-modal approach, though this was not replicated externally in identifying patients most likely to benefit from NAC.The prognostic significance of pathological response, particularly pCR and good histologic response, is well-established in gastric cancer literature, correlating strongly with improved long-term outcomes.[19] However, conventional radiological response assessments, such as RECIST, often show poor consistency with pathological findings.[20,21] Furthermore, their predictive value for overall survival can be limited.[21] Our study bypasses this limitation by directly predicting pathological response, which is a more reliable prognostic marker. By identifying patients who will not respond to treatment preoperatively, our model could potentially guide clinicians to alternative strategies, such as upfront surgery or different neoadjuvant regimens, thereby avoiding unnecessary treatment toxicity and delaying definitive surgery.

The integration of radiomic features into our model proved to be highly effective. Radiomics allows for the extraction of high-dimensional, quantitative data from medical images that are imperceptible to the human eye, providing a "non-invasive biopsy" of the tumor's phenotype and microenvironment.[22] The fact that features such as mean intensity and zone variance were identified as important by our model highlights their potential as imaging biomarkers. Several studies have similarly demonstrated the power of radiomics in predicting treatment response and survival in gastric cancer, often outperforming traditional methods.[20,23] Moreover, our use of LightGBM, a powerful and efficient gradient-boosting framework, is also noteworthy. It is well-suited for handling complex, high-dimensional data and has shown superior performance compared to other machine learning algorithms in similar predictive tasks.[24]

Our findings are consistent with previous reports suggesting that multi-modal, AI-driven approaches may have potential to may enhance precision medicine in oncology, although our failed external validation emphasizes that their clinical utility remains unproven.

For instance, recent studies have shown that machine learning models combining clinical data, radiomics, and even genomic biomarkers can accurately predict pathological response to both chemotherapy and immunotherapy in gastric cancer.[24] This multi-omics approach represents a promising future direction for a more comprehensive understanding of tumor biology and patient-specific response. While our study focused on a single institution and a specific treatment regimen, the methodology provides a strong foundation for future research, including prospective, multicenter trials needed to validate the model's clinical utility and establish its role in routine clinical practice.

A key methodological consideration in our study is the manual delineation of the tumor regions, which was performed by a single gastroenterological surgeon. The rationale for this approach was to develop a practical predictive tool specifically tailored for the primary treating physician in high-volume clinical settings, where the opportunity for extensive multidisciplinary council discussions on every patient may be limited. While this method benefits from expert clinical knowledge, it inherently introduces the potential for inter-observer and intra-observer variability. We acknowledge that a formal assessment of this variability was not conducted, which represents a limitation on the reproducibility and generalizability of our findings. This potential for inconsistency in radiomics feature extraction may have contributed to the performance discrepancy between our internal and external cohorts. This limitation highlights a critical area for future research, where the implementation of standardized, multi-reader delineation protocols or the development of robust automated segmentation algorithms could significantly improve reproducibility. For the present study, the manual delineation represents a potential source of bias, reinforcing the need for further validation with data from diverse institutions and standardized imaging workflows.

The notable decrease in predictive performance observed during the independent external validation, as evidenced by the drop in accuracy from 0.80 to 0.5459 and AUC from 0.7926 to 0.5409, represents a critical challenge for the broader applicability of our model. This significant drop in predictive performance is a well-recognized issue in machine learning and is often attributed to several factors. The most

likely reason is that the external dataset of 37 patients is fundamentally different from the training cohort.

This could be due to differences in CT image acquisition protocols. Moreover, demographics, disease stage distribution, or comorbidities may affect the results. Radiomic features are highly sensitive to variations in image acquisition. Even subtle differences in the CT scanner manufacturer, model, slice thickness, reconstruction algorithm, or contrast injection protocol between the original institution and the source of the external dataset could lead to different feature values. This "batch effect" can severely degrade a model's performance. Our model, trained on data from a single institution, may have learned patterns specific to that population and failed to generalize to the new, unseen cohort. Furthermore, the model may have overfit to the training data. While it performed well on the internal validation set, it might have learned noise or specific characteristics of the training data rather than true underlying biological signals. The external dataset then exposes this weakness, revealing that the model is not robust enough for a different population. Additionally, with only 37 patients in the external validation set, the performance metrics (accuracy, AUC, etc.) are highly susceptible to individual patient outcomes. This small sample size makes it difficult to draw a definitive conclusion about the model's true performance on a larger, more representative external population. The reported values might not be a stable or reliable measure of the model's capability.

A critical consideration in evaluating the generalizability of our predictive model is the relatively small size of the independent external validation cohort. While the inclusion of an external validation set is a significant strength of this study, a cohort of this size is inherently susceptible to individual patient variations, which can disproportionately influence performance metrics such as accuracy and AUC. This limited sample size complicates drawing definitive conclusions about the model's true generalizability and robust performance across diverse populations. Consequently, these external validation results, while indicative, underscore the crucial need for further validation in significantly larger, multi-institutional, and prospectively collected cohorts to establish the model's widespread clinical utility and broader applicability.

While acknowledged contributing factors include inherent dataset shift, variations in clinical practices, and potential overfitting to the internal cohort, this finding underscores the paramount importance of data harmonization. Moving forward, standardization of imaging protocols (e.g., scanner type, acquisition parameters, contrast injection timing) and detailed clinical definitions across different institutions will be crucial to mitigate these discrepancies. Furthermore, this finding strongly reinforces the necessity for larger, multi-institutional, prospective studies, which are essential to thoroughly validate the model's generalizability, minimize the impact of small sample size variability, and establish its consistent clinical utility across diverse patient populations and healthcare settings. Such collaborative efforts will be fundamental in transitioning this promising predictive tool into routine clinical practice. While our model demonstrated robust performance within the internal cohort, its failure to replicate this success on the external dataset can be attributed to several factors. These likely include subtle but impactful differences in CT imaging protocols across institutions, variations in clinical practices, and inherent dataset shift.

The feature importance analysis revealed that clinical variables such as albumin, CA 19-9, age, hemoglobin, and CEA were consistently ranked higher than most radiomic features. This pattern suggests that the predictive power of the model may be primarily driven by easily accessible clinical markers, with radiomics contributing minimally. While this does not reduce the potential role of radiomics, it highlights the need to systematically evaluate whether their integration provides significant incremental value over clinical features alone. Future work should include direct comparisons between models based solely on clinical data and those that integrate radiomic features, to determine the true added value of radiomics in predictive modeling for gastric cancer.

Limitations

This study has several key limitations. Its retrospective, single-center design may introduce selection bias and limit generalizability, and the relatively small sample size,

increases the risk of overfitting. Manual segmentation by a single observer also represents a potential source of bias, as radiomics features are sensitive to variations in both delineation and imaging protocols. Future multicenter studies with larger cohorts, standardized imaging, and automated or multi-reader segmentation approaches are needed to improve reproducibility and validation.

Another limitation of our study is the absence of a formal intra-observer and inter-observer reliability analysis of manual tumor segmentation. The manual delineations were performed by a single experienced gastroenterological surgeon, which provided clinical expertise but potentially introduces the possibility of intra-observer variability. Since repeated segmentations by the same observer or a radiologist were not available, an intraclass correlation coefficient (ICC) analysis could not be performed. This methodological limitation underscores the need for future research employing standardized multi-reader protocols or automated segmentation tools, which could enhance reproducibility and minimize observer-related variability in radiomics studies.

Another limitation of our study is the absence of image preprocessing or harmonization across datasets. The external validation cohort included scans from multiple centers with varying scanners, acquisition parameters (slice thickness, kVp, mAs), and contrast protocols, but detailed parameter-level information was not available for all patients. This heterogeneity likely contributed to the performance decline in external validation. While such variability reflects real-world clinical conditions, future studies should adopt standardized acquisition protocols or apply harmonization techniques (e.g., ComBat, intensity normalization) to improve reproducibility and generalizability of radiomics-based models.

A further methodological challenge concerns the generalizability of the selected features. Although we applied recursive feature elimination (RFE) with a LightGBM estimator and reduced the initial 112 features to 21, the external validation still showed a notable drop in accuracy

and AUC. This indicates that while feature selection reduced overfitting within the internal cohort, the chosen features may not have been sufficiently reliable across different institutions and imaging protocols. Future studies with larger, multi-institutional datasets and standardized imaging workflows are needed to validate the stability of selected features and improve reproducibility.

Furthermore, the predictive model is specific to the neoadjuvant FLOT regimen, and its applicability to other treatment protocols is not established. Lastly, while the LightGBM model is effective, its "black box" nature offers limited biological insight, suggesting a need for future studies that integrate radiomic findings with molecular data and confirm clinical utility through prospective trials.

CONCLUSION

This study explored the development of a LightGBM-based predictive model that showed potential in internal validation, but failed to generalize to an external cohort in classifying LAGC patients as responders or non-responders to neoadjuvant FLOT chemotherapy. These findings highlight both the opportunities and challenges of radiomics-based predictive modeling, and underscore the necessity of larger multicenter prospective validation before clinical application.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was approved by the Adana City Training and Research Hospital Clinical Researches Ethics Committee (Date: 07.11.2024, Decision No: 220).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
2. Ajani JA, D'Amico TA, Bentrem DJ, et al. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2022; 20(2):167-192. doi:10.6004/jnccn.2022.0008
3. Al-Batran SE, Homann N, Pauligk C, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet*. 2019; 393(10184):1948-1957. doi:10.1016/S0140-6736(18)32557-1
4. van Laarhoven H. Management of locally advanced gastric cancer. *Lancet Oncol*. 2025;26(3):267-269. doi:10.1016/S1470-2045(24)00727-7
5. Paszt A, Simonka Z, Budai K, et al. Impact of neoadjuvant FLOT treatment of advanced gastric and gastroesophageal junction cancer following surgical therapy. *Front Surg*. 2023;10:1148984. doi:10.3389/fsurg.2023.1148984
6. Sun F, Gao X, Wang W, Zhao X, Zhang J, Zhu Y. Predictive biomarkers in the era of immunotherapy for gastric cancer: current achievements and future perspectives. *Front Immunol*. 2025;16:1599908. doi:10.3389/fimmu.2025.1599908
7. Zhou H, Gu Q, Bao R, et al. Machine learning based models for predicting presentation delay risk among gastric cancer patients. *Front Oncol*. 2025;14:1503047. doi:10.3389/fonc.2024.1503047
8. Zubarayev M, Min EK, Son T. Clinical and molecular prognostic markers of survival after surgery for gastric cancer: tumor-node-metastasis staging system and beyond. *Transl Gastroenterol Hepatol*. 2019;4:59. doi:10.21037/tgh.2019.08.05
9. Kuhl CK, Alparslan Y, Schmoee J, et al. Validity of RECIST version 1.1 for response assessment in metastatic cancer: a prospective, multireader study. *Radiology*. 2019;290(2):349-356. doi:10.1148/radiol.2018180648
10. Chan AK, Wong A, Jenken D, Heine J, Buie D, Johnson D. Posttreatment TNM staging is a prognostic indicator of survival and recurrence in tethered or fixed rectal carcinoma after preoperative chemotherapy and radiotherapy. *Int J Radiat Oncol Biol Phys*. 2005;61(3):665-677. doi:10.1016/j.ijrobp.2004.06.206
11. Lowy AM, Mansfield PF, Leach SD, Pazdur R, Dumas P, Ajani JA. Response to neoadjuvant chemotherapy best predicts survival after curative resection of gastric cancer. *Ann Surg*. 1999;229(3):303-308. doi:10.1097/00000658-199903000-00001
12. Chen YH, Xiao J, Chen XJ, et al. Nomogram for predicting pathological complete response to neoadjuvant chemotherapy in patients with advanced gastric cancer. *World J Gastroenterol*. 2020;26(19):2427-2439. doi:10.3748/wjg.v26.i19.2427
13. Sun KY, Hu HT, Chen SL, et al. CT-based radiomics scores predict response to neoadjuvant chemotherapy and survival in patients with gastric cancer. *BMC Cancer*. 2020;20(1):468. doi:10.1186/s12885-020-06970-7
14. Adili D, Mohetaer A, Zhang W. Diagnostic accuracy of radiomics-based machine learning for neoadjuvant chemotherapy response and survival prediction in gastric cancer patients: a systematic review and meta-analysis. *Eur J Radiol*. 2024;173:111249. doi:10.1016/j.ejrad.2023.111249
15. Song R, Cui Y, Ren J, et al. CT-based radiomics analysis in the prediction of response to neoadjuvant chemotherapy in locally advanced gastric cancer: a dual-center study. *Radiother Oncol*. 2022;171: 155-163. doi:10.1016/j.radonc.2022.04.023
16. Wang HN, An JH, Wang FQ, Hu WQ, Zong L. Predicting gastric cancer survival using machine learning: a systematic review. *World J Gastrointest Oncol*. 2025;17(5):103804. doi:10.4251/wjgo.v17.i5.103804
17. Akcay M, Etiz D, Celik O. Prediction of survival and recurrence patterns by machine learning in gastric cancer cases undergoing radiation therapy and chemotherapy. *Adv Radiat Oncol*. 2020;5(6):1179-1187. doi:10.1016/j.adro.2020.07.007
18. Wang W, Sheng R, Liao S, et al. LightGBM is an effective predictive model for postoperative complications in gastric cancer: a study integrating radiomics with ensemble learning. *J Imaging Inform Med*. 2024;37(6):3034-3048. doi:10.1007/s10278-024-01172-0
19. Wang S, Zheng G, Cheng B, et al. Circulating tumor cells (CTCs) detected by RT-PCR and its prognostic role in gastric cancer: a meta-analysis of published literature. *PLoS One*. 2014;9(6):e99259. doi:10.1371/journal.pone.0099259
20. Sun C, Niu P, Zhang X, et al. Concurrent clinical and pathological response predicts favorable prognosis of patients with gastric cancer after neoadjuvant therapy: a real-world study. *BMC Cancer*. 2023;23(1):996. doi:10.1186/s12885-023-11508-8
21. D'Alessio A, Lombardi P, Stefanini B, et al. Use of pathological response versus RECIST assessment in predicting relapse-free survival following neoadjuvant immunotherapy for hepatocellular carcinoma. Meeting Abstract: 2025 ASCO Gastrointestinal Cancers Symposium. *J Clin Oncol*. 2025;43(Suppl 4). doi:10.1200/JCO.2025.43.4_suppl.633

22. Wang Y, Jin ZY. Radiomics approaches in gastric cancer: a frontier in clinical decision making. *Chin Med J (Engl)*. 2019;132(16):1983-1989. doi:10.1097/CM9.0000000000000360
23. Yang L, Chu W, Li M, et al. Radiomics in gastric cancer: first clinical investigation to predict lymph vascular invasion and survival outcome using 18F-FDG PET/CT images. *Front Oncol*. 2022;12:836098. doi:10.3389/fonc.2022.836098
24. Huang ZN, Zhang HX, Sun YQ, et al. Multi-cohort study in gastric cancer to develop CT-based radiomic models to predict pathological response to neoadjuvant immunotherapy. *J Transl Med*. 2025;23(1):362. doi:10.1186/s12967-025-06363-z

Journal of Radiology in Medicine

Original Article

Doi: 10.51271/JRM-0038

Magnetic resonance imaging (MRI) based cerebellum and brain volume measurement using cavalier principle in epilepsy patients on long-term phenytoin and carbamazepine therapy

Orhan Oyar[1], Mustafa Yalçın[2]

[1]Department of Radiology, Faculty of Medicine, İzmir Katip Çelebi University, İzmir, Turkiye
[2]Department of Radiology, Antalya Training and Researh Hospital, Antalya, Turkiye

Received: 30.07.2025 ◆ **Accepted**: 06.10.2025 ◆ **Published**: 15.10.2025

Cite this article: Oyar O, Yalçın M. Magnetic resonance imaging (MRI) based cerebellum and brain volume measurement using cavalier principle in epilepsy patients on long-term phenytoin and carbamazepine therapy. *J Radiol Med.* 2025;2(4):71-75.

Corresponding Author: Mustafa Yalçın, myalcin21@hotmail.com

ABSTRACT

Aims: This study aims to evaluate cerebellar atrophy findings in epilepsy patients using long-term antiepileptic drugs by calculating cerebellum/brain volume ratios with the Cavalieri principle and comparing this method with other volumetric measurement techniques.

Methods: A total of 30 patients diagnosed with epilepsy and receiving long-term phenytoin or carbamazepine treatment between January 2004 and May 2005, along with a control group of 15 individuals, were included in the study. Magnetic resonance imaging (MRI) was performed using a 0.2 Tesla open MRI system. Brain volumes and cerebellum/brain volume ratios were calculated using the point-counting method based on the Cavalieri principle.

Results: No statistically significant difference was found between the calculated volume values of the phenytoin, carbamazepine, and control groups ($p>0.05$). However, in one patient using phenytoin, the cerebellum volume was found to be significantly low, consistent with cerebellar atrophy.

Conclusion: Larger studies are needed to determine whether long-term use of phenytoin and carbamazepine leads to cerebellar atrophy. The point-counting method based on the Cavalieri principle may offer a practical and applicable alternative for routine clinical use.

Keywords: Cerebellar atrophy, epilepsy, phenytoin, stereology, Cavalieri principle

INTRODUCTION

Epilepsy is one of the most common chronic neurological disorders affecting millions of people worldwide. Antiepileptic drugs (AEDs) form the cornerstone of epilepsy treatment; however, their long-term use has been associated with certain neurotoxic effects. Some studies suggest that AEDs, such as phenytoin and carbamazepine, may contribute to cerebellar atrophy (shrinkage of the cerebellum), leading to motor coordination deficits and balance issues. However, it remains unclear whether cerebellar atrophy is solely related to drug use or whether the frequency and duration of seizures also play a role.

The primary aim of this study is to assess the presence and degree of cerebellar atrophy in epilepsy patients on long-term phenytoin and carbamazepine therapy. To achieve this, cerebellar volume and cerebellum-brain ratios were calculated using magnetic resonance imaging (MRI) combined with the Cavalier principle-based point-counting technique and compared to a control group. This study aims to objectively assess the effects of epilepsy medications on cerebellar volume and contribute to more reliable diagnostic methods in clinical practice. Furthermore, the findings may highlight the importance of personalized approaches in epilepsy treatment.

METHODS

Ethics

This article is derived from the medical specialty thesis of Dr. Mustafa Yalçın, conducted at Isparta Süleyman Demirel

University in 2005. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

This retrospective study examines epilepsy patients who have been on phenytoin or carbamazepine therapy for at least five years. Individuals without other neurological disorders, metabolic conditions, or structural brain pathologies, and who had complete MRI scans, were included. A total of 45 individuals participated: 15 on phenytoin therapy, 15 on carbamazepine therapy, and 15 healthy controls. Participants were matched for age, gender, and body-mass index (BMI).

Imaging Protocol

All patients and the control group underwent sagittal T1-weighted MRI scans using a 0.2 Tesla open MRI system. A standard neuroimaging protocol was followed during the MRI scans. Head supports were used to standardize head positioning, and participants were instructed to remain still during the scan. Individuals with motion artifacts in their scans were excluded from the study. Additionally, all participants underwent a neurological examination prior to imaging and were assessed for signs of cerebellar dysfunction.

Volume Calculation and Applied Formulas

Cerebellum and brain volumes were calculated using the Cavalier principle-based point-counting method. The brain and cerebellum areas were analyzed across predefined slice intervals to obtain the total volumes. The point-counting method relies on counting the number of grid points that intersect the target region in each slice. The point-counting method based on the Cavalieri principle was applied (**Figure 1**).

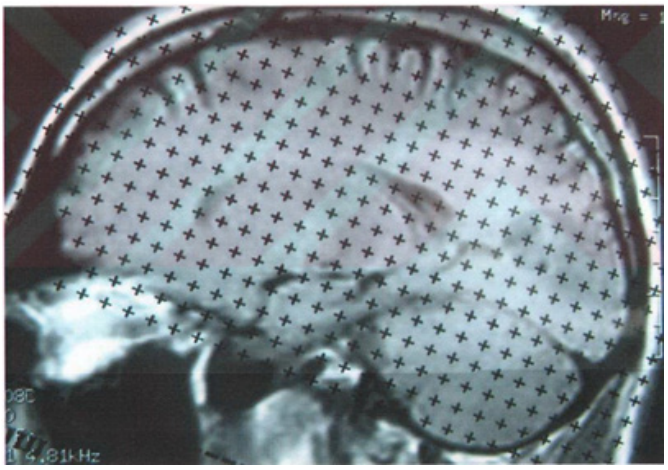


Figure 1. Random point grid superimposed on cross-sections for area estimation

The formula used for volume calculation is as follows: $V=t\times\sum P\times a(p)$

Where:

•V=Volume (mm^3)

•t=Slice thickness and slice interval (mm)

•P=Number of points in each slice

•a(p)=Area represented by each point (mm^2)

This formula allows the volume estimate to account for the gap between slices and the total area calculated in each slice. The point-counting process involves determining the number of grid points in each slice that cover the cerebellum.

The cerebellum-brain volume ratio was calculated as: Ratio=V_cerebellum/V_brain×100

This ratio was used as a comparative metric in cerebellar atrophy assessment and helped identify differences among patients.

Measurements were performed by two independent observers, and intra-observer and inter-observer reliability were assessed. Bland-Altman analysis was applied to determine statistical consistency.

Statistical Analysis

The data analyses were performed using SPSS 22.0 software. Volume differences between groups were analyzed using the Kruskal-Wallis test. Pairwise comparisons were made using the Mann-Whitney U test. The correlation between cerebellar volume and cerebellum-brain ratios was evaluated using Spearman correlation analysis. A p-value of <0.05 was considered statistically significant. Additionally, multivariate logistic regression analysis was conducted to control for the effects of age and gender.

RESULTS

Cerebral and cerebellar volumes, as well as cerebellum-to-brain volume ratios, were calculated in patients receiving phenytoin or carbamazepine and in healthy controls using the point-counting method based on the Cavalieri principle. The mean cerebellar and total brain volumes, along with cerebellum-to-brain ratios and standard deviations, are presented below. The calculated cerebellar and total brain volumes of patients receiving phenytoin are presented in **Table 1**.

Table 1. Calculated cerebellar and total brain volumes and cerebellum/brain ratios for phenytoin users

Patient	Cerebellum volume (mm^3)	Total brain volume (mm^3)	Cerebellum/ brain ratio, (%)
1	165	1684	9.79
2	147	1287	11.43
3	133	1100	12.11
4	130	1308	9.96
5	123	1313	9.34
6	150	1349	11.16
7	137	1406	9.76
8	120	1195	10.05
9	125	1283	9.74
10	119	1135	10.51
11	96	1048	9.19
12	123	909	13.52
13	130	1135	11.46
14	155	1202	12.86
15	124	1082	11.46

Comparative analysis revealed no statistically significant differences in volume parameters among the phenytoin, carbamazepine, and control groups ($p>0.05$). However, one

individual in the phenytoin group demonstrated a markedly reduced cerebellum-to-brain ratio (7.7%), the lowest observed across all cohorts and consistent with radiological evidence of atrophy. Visual inspection of MR images demonstrated widened cerebrospinal fluid spaces consistent with cerebellar atrophy. Corresponding values for carbamazepine users are shown in **Table 2**.

Table 2. Cerebellar and brain volumes and their ratios for carbamazepine users

Patient	Cerebellum volume (mm^3)	Total brain volume (mm^3)	Cerebellum/ brain ratio (%)
1	132	1119	11.80
2	119	1156	10.32
3	118	1134	10.44
4	140	1153	12.14
5	138	1390	9.94
6	134	1218	10.99
7	127	1097	11.55
8	123	1230	9.98
9	131	1221	10.71
10	123	1101	11.00
11	121	1165	10.41
12	105	1254	8.37
13	94	853	11 03
14	101	1028	9.81
15	107	1050	10.17

No statistically significant differences were found between the carbamazepine group and either the phenytoin or control groups (p>0.05). Group comparisons of cerebellar volume, brain volume, and cerebellum-to-brain ratio are summarized in **Figure 2**.

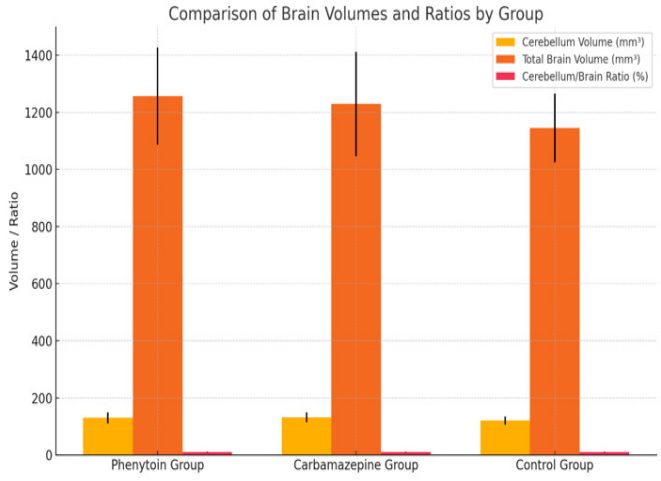


Figure 2. Mean cerebellar and brain volumes, and cerebellum/brain volume ratios with standard deviations across groups

Comparison of cerebellar volume, total brain volume, and cerebellum-to-brain ratio among the study groups. The chart displays mean values±standard deviations for each parameter across the phenytoin group, carbamazepine group, and healthy controls. While total and regional brain volumes differ among groups, the cerebellum-to-brain volume ratio remains relatively stable, indicating proportional preservation despite overall volumetric variations.

There were no statistically significant differences among groups in cerebellum-to-brain volume ratios (p>0.05).

Cerebellar atrophy is a frequent finding in chronic epilepsy patients and has been associated with long-term phenytoin usage. Multiple pathophysiological mechanisms have been proposed, including direct toxicity to Purkinje cells, degeneration of cerebellar afferent pathways, and seizures themselves leading to hypoxia-induced injury. Some studies suggest a synergistic role of genetics in increasing susceptibility to these effects.[1-3]

Imaging modalities such as CT, MRI, MR volumetry, MR spectroscopy, and diffusion tensor imaging have all been used to assess cerebellar atrophy.[4-7] Notably, cerebellar atrophy is considered a marker of poor prognosis in epilepsy and may influence surgical decision-making, including temporal lobectomy.[8,9]

In our study, no significant differences were observed between the phenytoin, carbamazepine, and control groups. However, one subject on phenytoin therapy exhibited both a reduced cerebellum/brain volume ratio and radiological signs of cerebellar atrophy, in the absence of confounding variables. This supports the potential of phenytoin to induce isolated cases of cerebellar atrophy, in line with previous literature.[10-13]

Multivariate logistic regression analysis, adjusted for age and gender, showed no significant association between antiepileptic drug type (phenytoin vs. carbamazepine) and cerebellum-to-brain volume ratio (p>0.05). Similarly, neither age nor gender independently predicted cerebellar volume reduction in this cohort.

Carbamazepine showed no such effect, aligning with prior findings that it does not independently cause cerebellar atrophy, although caution is advised in patients with pre-existing cerebellar degeneration.[8-14]

Our results reinforce the utility of the Cavalieri principle in delivering reliable volumetric assessments, with coefficients of error remaining below 5% in all cases. The technique offers a cost-effective, efficient alternative for tracking volumetric changes in neurological patients.

An important limitation of the present study is the relatively small sample size (n=45). This limited number of participants reduces the statistical power of the analyses and may have obscured subtle volumetric differences among the groups. Therefore, the absence of significant findings should be interpreted with caution. Larger, prospective, multicenter studies are warranted to provide more definitive conclusions.

DISCUSSION

This study aimed to evaluate the effects of commonly used AEDs, phenytoin and carbamazepine, on cerebellar volume. The results revealed no statistically significant difference between groups. However, cerebellar atrophy was detected in a phenytoin user, indicating potential long-term neurotoxic effects of phenytoin. The outlier case with a markedly low

cerebellum-to-brain ratio (**Table 1**) supports the potential for phenytoin-related neurotoxicity.

Previous studies have shown that phenytoin use may contribute to dose-dependent cerebellar atrophy. Long-term phenytoin use can lead to Purkinje cell loss, significantly reducing cerebellar volume. Animal studies suggest that phenytoin may trigger neuronal degeneration by increasing oxidative stress levels. Human studies have reported that neurodegenerative effects are dose-dependent.

Cerebellar atrophy in epilepsy patients may not solely be due to medication use; the frequency and duration of seizures could also play a significant role. Chronic seizure activity may lead to neuronal loss in cerebellar structures through hypoxia and inflammation. Some studies suggest that although cerebellar volume reduction is observed in epilepsy patients, these changes may be more strongly associated with drug effects rather than seizure frequency.

In this study, the Cavalier principle-based point-counting technique was used for cerebellar volume assessment. This technique provides more precise volume measurements compared to manual or visual evaluations. Both phenytoin and carbamazepine users were analyzed separately and compared with the control group, enabling a more comprehensive evaluation of the effects of different drugs on cerebellar structures. The analysis of patients who have been on AEDs for at least five years is crucial for determining the long-term neurological effects of these medications.

When comparing our findings to the existing literature, some inconsistencies were noted. For instance, earlier research has demonstrated cerebellar volume loss associated with long-term phenytoin use, although our findings did not show a statistically significant difference among groups. This discrepancy may be due to sample size, MRI resolution, and individual patient variability. Other studies have also examined the effects of multiple AEDs on cerebellar volume, with carbamazepine being suggested to have less neurotoxic effect.

Recent evidence has reinforced the association between long-term antiepileptic drug exposure and cerebellar structural alterations. Prolonged use of phenytoin has been linked to irreversible cerebellar damage, particularly due to Purkinje cell vulnerability.[15] Mechanistic pathways underlying chronic drug exposure and cerebellar degeneration have also been described.[16] From a methodological standpoint, normative modeling approaches have been introduced to improve volumetric MRI biomarkers, thereby enhancing sensitivity in heterogeneous epilepsy cohorts.[17] Furthermore, structural and functional cerebellar abnormalities have been demonstrated in temporal lobe epilepsy using volumetric, diffusion, and connectivity-based analyses.[18] A systematic review additionally emphasized that phenytoin exposure, poor seizure control, and long disease duration remain consistent risk factors for cerebellar degeneration.[19] Taken together, these contemporary findings suggest that our results should be interpreted within the context of technological advances and larger datasets that have emerged since the time of data collection. Although the use of a low-field MRI scanner and limited sample size may have prevented the detection of subtle changes, the Cavalieri principle remains a valuable volumetric method that can complement modern imaging approaches.

The use of a low-field MRI scanner in our study may have limited the detection of subtle morphological changes. Future studies using high-field MRI systems could provide clearer insights into the relationship between cerebellar atrophy and AED use. Moreover, long-term prospective studies are needed to assess the progressive nature of cerebellar atrophy over time.

Another limitation of this study is the sample size. Larger, multicenter studies would improve the generalizability of the findings. Furthermore, future research should consider the detailed neurological histories of patients, including seizure type and duration, to evaluate their effects on cerebellar atrophy more comprehensively. Genetic factors may also contribute to cerebellar atrophy in epilepsy patients, and pharmacogenetic studies could help identify individuals more susceptible to phenytoin's neurotoxic effects.

Limitations

A major limitation of our study is that the data were collected nearly two decades ago using low-field MRI technology. Since then, substantial progress has been made in neuroimaging, including automated volumetric methods, radiomics, and ultra–high-field (7T) MRI, which enable more sensitive and reliable assessment of subtle volumetric changes. Nevertheless, by applying the Cavalieri principle to this historical dataset, our study provides methodological insights that remain relevant, especially for centers where advanced imaging resources are limited. When considered in the context of more recent investigations,[15,16,18] our findings underscore the need for large-scale, prospective studies to confirm the long-term neurotoxic effects of phenytoin and to further elucidate the role of the cerebellum in epilepsy.

Although there is strong evidence supporting phenytoin's neurotoxic effects, the absence of cerebellar atrophy in some patients suggests that individual genetic and metabolic differences may play a role. Future research focusing on pharmacogenetics could help identify patients at higher risk for neurological side effects and enhance personalized epilepsy treatment. As this study is retrospective, long-term prospective follow-up studies are needed to evaluate the The retrospective nature of our study constitutes another limitation, as it prevents us from establishing causal relationships and from monitoring progressive changes over time. Prospective, longitudinal follow-up studies would be more suitable for clarifying the temporal relationship between long-term antiepileptic drug exposure and cerebellar atrophy. Progressive nature of cerebellar atrophy over time.

CONCLUSION

The Cavalier principle-based point-counting technique is an objective and feasible method for assessing cerebellar atrophy in epilepsy patients. However, further large-scale, long-term studies using high-field MRI scanners are necessary to clarify the relationship between AED use and cerebellar atrophy.

Future research should also examine the effects of seizure duration and frequency on cerebellar atrophy. Personalized drug selection and seizure management are critical strategies for minimizing cerebellar degeneration in epilepsy patients.

Finally, it should be acknowledged that this study was conducted nearly 20 years ago using a low-field MRI scanner, which limits its applicability to current clinical practice. Advances in high-field MRI, automated volumetry, and radiomics now provide superior methods for assessing cerebellar atrophy. Nonetheless, our findings demonstrate the feasibility and reproducibility of the Cavalieri principle and emphasize the importance of historical data in shaping contemporary research directions.

This study demonstrated that long-term use of phenytoin and carbamazepine did not result in statistically significant differences in cerebellar or total brain volumes when compared to healthy controls. However, an isolated case of cerebellar atrophy was observed in a patient receiving phenytoin, suggesting that individual susceptibility or drug-specific neurotoxic effects may still play a role. The Cavalieri principle-based point-counting method proved to be a reliable and practical approach for volumetric assessment, with low measurement error and consistent reproducibility. While our findings do not support a generalized association between long-term carbamazepine or phenytoin use and cerebellar volume reduction, they highlight the importance of monitoring individual patients for potential neuroanatomical changes. Larger, prospective, and high-resolution imaging studies are warranted to clarify the long-term neurological effects of AEDs and to further evaluate the contribution of seizure frequency, genetic predisposition, and treatment duration. Personalized treatment strategies remain essential in minimizing neurological side effects and optimizing outcomes in patients with chronic epilepsy.

Another limitation of our study is the use of a low-field (0.2 Tesla) MRI scanner, which inherently restricts spatial resolution and sensitivity to subtle volumetric alterations. This technical limitation may have reduced the ability to detect mild degrees of cerebellar or cerebral volume loss. Future investigations should utilize high-field MRI systems (≥1.5 Tesla or 3 Tesla), which provide superior image quality and volumetric precision.

ETHICAL DECLARATIONS

Ethics Committee Approval

This article is derived from the medical specialty thesis of Dr. Mustafa Yalçın, conducted at Isparta Süleyman Demirel University in 2005.

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. De Marco FA, Ghizoni E, Kobayashi E, Li LM, Cendes F. Cerebellar volume and long-term use of phenytoin. *Seizure*. 2003;12(5):312-315. doi:10.1016/s1059-1311(02)00267-4
2. Lindvall O, Nilsson B. Cerebellar atrophy following phenytoin intoxication. *Ann Neurol*. 1984;16(2):258-260. doi:10.1002/ana.410160218
3. Luft AR, Skalej M, Schulz JB, et al. Patterns of age-related shrinkage in cerebellum and brainstem observed in vivo using three-dimensional MRI volumetry. *Cereb Cortex*. 1999;9(7):712-721. doi:10.1093/cercor/9.7.712
4. Hayakawa K, Kumagai H, Suzuki Y, et al. MR imaging of chronic alcoholism. *Acta Radiol*. 1992;33(3):201-206.
5. Savoiardo M, Bracchi M, Passerini A, et al. Computed tomography of olivopontocerebellar degeneration. *AJNR Am J Neuroradiol*. 1983;4(3):509-512.
6. Ramos A, Quintana F, Diez C, et al. CT findings in spinocerebellar degeneration. *AJNR Am J Neuroradiol*. 1987;8(4):635-640.
7. Tedeschi G, Bertolino A, Massaquoi SG, et al. Proton magnetic resonance spectroscopic imaging in patients with cerebellar degeneration. *Ann Neurol*. 1996;39(1):71-78. doi:10.1002/ana.410390111
8. Specht U, May TW, Rohde M, et al. Cerebellar atrophy decreases the threshold of carbamazepine toxicity in patients with chronic focal epilepsy. *Arch Neurol*. 1997;54(4):427-431. doi:10.1001/archneur.1997.00550160063017
9. Specht U, May T, Schulz R, et al. Cerebellar atrophy and prognosis after temporal lobe resection. *J Neurol Neurosurg Psychiatry*. 1997;62(5):501-506. doi:10.1136/jnnp.62.5.501
10. Ney GC, Lantos G, Barr WB, Schaul N. Cerebellar atrophy in patients with long-term phenytoin exposure and epilepsy. *Arch Neurol*. 1995;52(3):232. doi:10.1001/archneur.1994.00540200043014
11. Guerrero AL, Paniagua JA, Diaz Cascajo P, et al. Temporal cerebellar atrophy following phenytoin therapy. *Neurologia*. 1997;12(6):259-261.
12. Savic I, Thorell JO. Localized cerebellar reductions in benzodiazepine density in human partial epilepsy. *Arch Neurol*. 1996;53(7):656-662. doi:10.1001/archneur.1996.00550070094016
13. Luef G, Burtscher J, Kremser C, et al. Magnetic resonance volumetry of the cerebellum in epileptic patients after phenytoin overdosages. *Eur Neurol*. 1996;36(5):273-277. doi:10.1159/000117271
14. Luef G, Chemelli A, Birbamer G, et al. Phenytoin overdosage and cerebellar atrophy in epileptic patients: clinical and MRI findings. *Eur Neurol*. 1994;34(2):79-81. doi:10.1159/000119516
15. Ferner RE. Phenytoin and damage to the cerebellum. *Expert Opin Drug Saf*. 2022;21(7):801-804. doi:10.1080/14740338.2022.2058487
16. Elder C. The cerebellum in epilepsy. *Epilepsia*. 2025;66(1):34-45. doi:10.1111/epi.18316
17. Mito R, Cole JH, Genc S, Jackson GD, Zalesky A. Towards precision MRI biomarkers in epilepsy with normative modelling. *Brain*. 2025;148(7):2247-2261. doi:10.1093/brain/awaf090
18. Wang G, Liu X, Zhang M, et al. Structural and functional changes of the cerebellum in temporal lobe epilepsy. *Front Neurol*. 2023;14:1213224. doi:10.3389/fneur.2023.1213224
19. Ibdali M, Hadjivassiliou M, Grünewald RA, Shanmugarajah PD. Cerebellar degeneration in epilepsy: a systematic review. *Int J Environ Res Public Health*. 2021;18(2):473. doi:10.3390/ijerph18020473

Journal of Radiology Medicine in

Original Article

Doi: 10.51271/JRM-0039

Radiological evaluation of the impact of chronic obstructive pulmonary disease on ascending aortic diameter in hypertensive patients

Can Kerestecioğlu[1], Ümmühan Tuğba Tümüklü[2], Duygu Felek[3], Mustafa Fatih Erkoç[4]

[1]Department of Cardiovascular Surgery, Faculty of Medicine, Yozgat Bozok University, Yozgat, Turkiye
[2]Department of Pulmonology, Faculty of Medicine, Yozgat Bozok University, Yozgat, Turkiye
[3]Department of Internal Medicine, Faculty of Medicine, Yozgat Bozok University, Yozgat, Turkiye
[4]Department of Radiology, Faculty of Medicine, Yozgat Bozok University, Yozgat, Turkiye

Received: 08.08.2025 ◆ **Accepted**: 06.10.2025 ◆ **Published**: 15.10.2025

Cite this article: Kerestecioğlu C, Tümüklü ÜT, Felek D, Erkoç MF. Radiological evaluation of the impact of chronic obstructive pulmonary disease on ascending aortic diameter in hypertensive patients. *J Radiol Med*. 2025;2(4):76-80.

Corresponding Author: Can Kerestecioğlu, cankerestecioglu@hotmail.com

ABSTRACT

Aims: This study aims to radiologically evaluate the impact of chronic obstructive pulmonary disease (COPD) on the diameter of the AsAo in individuals diagnosed with hypertension and to assess potential structural effects of COPD on the cardiovascular system.

Methods: This retrospective cross-sectional study was conducted at Yozgat Bozok University Research and Training Hospital. Clinical and radiological data of 60 hypertensive individuals who underwent thoracic computed tomography (CT) between January 1, 2023, and June 1, 2025, were analyzed. Participants were divided into two groups based on the presence or absence of a COPD diagnosis. AsAo and aortic arch (Arcus Ao) diameters were measured from CT images. Clinical variables, comorbidities, and lipid profiles were statistically evaluated.

Results: The mean age of participants was significantly higher in the COPD group (73.8±7.92 years) compared to the control group (67.93±12.59 years) (p=0.036). The mean AsAo diameter was 40.40±4.90 mm in the COPD group and 36.01±3.79 mm in the control group, revealing a statistically significant difference (p=0.00028). No significant difference was found between the groups regarding Arcus Ao diameter (p=0.103). Among lipid parameters, there were no significant differences in HDL, LDL, or total cholesterol levels; however, triglyceride levels showed a trend toward significance (p=0.069).

Conclusion: The findings suggest that COPD is not merely a respiratory condition but also a systemic disease with notable vascular implications. The significant increase in AsAo diameter in the COPD group may be attributed to chronic inflammation, endothelial dysfunction, and degeneration of elastic fibers associated with the disease. The absence of a significant difference in Arcus Ao diameter may be related to the anatomically stable structure of this segment. The higher comorbidity burden and older age observed in COPD patients further support the systemic nature of the disease. Therefore, COPD should be regarded as a potential vascular risk modifier to be considered in hypertensive individuals.

Keywords: Ascending aortic diameter, chronic obstructive pulmonary disease (COPD), hypertension

INTRODUCTION

Hypertension (HT) is recognized as one of the leading causes of cardiovascular morbidity and mortality worldwide. According to the 2018 guidelines jointly issued by the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH), HT is defined as a systolic blood pressure ≥140 mmHg and/or a diastolic blood pressure ≥90 mmHg.[1] Over time, HT may lead to degeneration of elastic fibers and accumulation of collagen in the aortic wall, resulting in structural changes—particularly in the ascending aorta—which can ultimately cause aortic dilatation. This pathological process increases the risk of life-threatening complications such as aortic dissection and rupture.[2]

A meta-analysis has reported the prevalence of proximal aortic dilatation in hypertensive individuals to range between 6.1% and 19%.[2] Furthermore, a Mendelian randomization study revealed that genetically predicted increases in systolic and diastolic blood pressure were significantly associated

with greater ascending thoracic aorta diameter.[3] This finding highlights the pronounced effect of HT on aortic morphology. A 2025 position paper also emphasized that HT may lead to remodeling of the proximal aortic wall, which is closely linked to clinically significant outcomes such as dissection and aneurysm formation.[4]

Chronic obstructive pulmonary disease (COPD), particularly in individuals with emphysematous phenotype, can also contribute to structural changes in the aortic wall. This is primarily due to systemic inflammation, oxidative stress, and increased proteolytic enzyme activity, which together result in elastic fiber loss and vascular remodeling.[5] In the large-scale, multi-center MESA (Multi-Ethnic Study of Atherosclerosis) study conducted in 2021, individuals with COPD were shown to have significantly larger ascending aortic diameters compared to those without the disease.[5]

The coexistence of COPD and HT is commonly observed in clinical practice. A large cohort study covering the years 1999 to 2018 reported a HT prevalence of 58.5% among individuals with COPD, with a particularly higher rate among adults under 60 years of age and heavy smokers.[6] In addition, frequently observed features in COPD—such as arterial stiffness, endothelial dysfunction, and increased pulmonary vascular resistance—may contribute to a compounded cardiovascular burden through parallel dysfunctions in both systemic and pulmonary vascular systems.[8]

The coexistence of HT and COPD may serve as a dual risk factor for ascending aortic dilatation. In a retrospective study on individuals with COPD, the prevalence of aortic aneurysm was reported as 14.3%, and 56% of those affected also had HT.[7] Nevertheless, studies specifically focusing on the impact of COPD on ascending aortic diameter in hypertensive individuals remain limited. Given the high likelihood of coexistence between COPD and cardiovascular disease—and the potential for more severe clinical outcomes—exploring this relationship in greater detail is essential for effective cardiovascular risk management.[9]

Recent radiological and prospective studies have demonstrated that COPD can lead to systemic vascular effects beyond the pulmonary system, such as increased aortic inflammation and arterial sclerosis.[10] Moreover, systemic inflammation, hyperinflation, and intrathoracic pressure changes associated with COPD may increase wall stress in large vessels, including the ascending aorta, thereby promoting dilatation.[5] Population-based imaging studies have further shown that advanced age, male sex, smoking, dyslipidemia, and obesity are significantly associated with age-related vascular remodeling processes and with ascending aortic enlargement.[11,12]

In this context, radiological evaluation of the effect of COPD on ascending aortic diameter in individuals diagnosed with HT may enhance our understanding of the combined impact of HT and respiratory disease on aortic morphology. Furthermore, it may contribute to the development of more personalized and targeted cardiovascular risk stratification strategies by incorporating individual risk factors.[13,14]

METHODS

Ethics

The study was conducted with the approval of the Non-interventional Clinical Researches Ethics Committee of Yozgat Bozok University Faculty of Medicine (Date: 02.07.2025, Decision No: 2025-GOKAEK-2513_2025.07.02_540). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki. As a retrospective study, no diagnostic or therapeutic intervention was made in the course of patient care, and all data were obtained from existing clinical and radiological records.

Study Design and Participants

This study aimed to radiologically evaluate the impact of COPD on ascending aortic diameter in hypertensive individuals who presented to the Radiology, Pulmonology, and Cardiology departments of Yozgat Bozok University Research and Training Hospital. The study was based on clinical and radiological evaluations performed between January 1, 2023, and June 1, 2025.

The study sample consisted of individuals diagnosed with HT who had undergone thoracic CT. Participants were categorized into two groups based on the presence or absence of a clinical diagnosis of COPD. Only individuals aged 18 years and older were included. Patients with active infections, malignancies, systemic connective tissue diseases, a history of immunosuppressive therapy, or current use of pharmacological agents that could affect immune function were excluded.

Imaging Protocol

Measurements of the AsAo diameter were retrospectively evaluated using thoracic CT images obtained for clinical indications. CT scans were performed with a Philips Brilliance 64 Slice CT scanner using a thoracic protocol with 1-mm slice thickness, 120–80 kV, and 180 mAs dose modulation, and were acquired without contrast. All scans were performed with ECG gating. Individuals with available imaging were included in the study. Measurements were performed using standard techniques at the level of the pulmonary artery bifurcation, with the maximum anteroposterior diameter on the axial plane taken as the reference. All measurements were independently conducted by a cardiovascular surgeon and a radiologist, both with clinical and radiological experience, and no significant interobserver variability was observed between the measurements.

Clinical and Demographic Data

Data were retrieved digitally through the hospital information management system (HIMS). Systematic records were collected for all participants, including age, sex, smoking status, and cardiovascular risk factors such as HT, diabetes mellitus (DM), dyslipidemia, obesity, and family history. In addition to clinical data, laboratory parameters were reviewed; however, the primary focus of the study was to examine the relationship between radiological measurements and accompanying clinical variables.

Statistical Analysis

The data analyses were performed using IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, NY, USA). Continuous variables were presented as mean±standard deviation (SD), and categorical variables as numbers and percentages (%). The distribution of variables was assessed using the Kolmogorov-Smirnov test. Independent samples t-test was used to compare normally distributed groups, while the Mann-Whitney U test was applied for non-normally distributed data. Categorical variables were analyzed using the Chi-square test. Pearson or Spearman correlation tests were employed to assess the relationship between the presence of COPD, ascending aortic diameter, and cardiometabolic risk factors. A p-value <0.05 was considered statistically significant.

RESULTS

Of the 60 individuals included in the study, 51.7% (n=31) were male and 48.3% (n=29) were female, indicating a relatively balanced gender distribution. No statistically significant difference was observed between the groups in terms of gender (p=1.000). The participants' ages ranged from 38 to 92 years. The mean age in the COPD group was 73.8±7.92 years, whereas it was 67.93±12.59 years in the non-COPD group. The independent samples t-test revealed a statistically significant age difference between the groups (p=0.036, p<0.05).

When evaluating comorbidities among participants, DM was the most commonly observed condition. DM was present in 43.3% of the cohort, while heart failure (HF) was identified in 13.3% of individuals.

Lipid parameters were compared between individuals with and without COPD. The mean HDL cholesterol level was 45.96±11.55 mg/dl in the COPD group and 42.31±11.67 mg/dl in the control group. LDL levels were 97.76±32.45 mg/dl and 110.09±40.12 mg/dl, respectively. The mean triglyceride levels were 128.83±60.03 mg/dl in the COPD group and 159.63±68.50 mg/dl in the non-COPD group. Total cholesterol was measured at 167.65±37.24 mg/dl and 184.35±49.42 mg/dl, respectively. According to independent samples t-tests, there were no statistically significant differences in HDL (p=0.229), LDL (p=0.195), triglyceride (p=0.069), or total cholesterol (p=0.145) levels between the groups (p>0.05) (**Table 1**).

Table 1. Comparison of lipid profiles

Parameter	COPD (+) mean±SD	COPD (–) Mean±SD	p-value
HDL (mg/dl)	45.96±11.55	42.31±11.67	0.229
LDL (mg/dl)	97.76±32.45	110.09±40.12	0.195
Triglycerides (mg/dl)	128.83±60.03	159.63±68.50	0.069
Total cholesterol (mg/dl)	167.65±37.24	184.35±49.42	0.145

HDL: High-density lipoprotein, LDL: Low-density lipoprotein, SD: Standard deviation, COPD: Chronic obstructive pulmonary disease

The AsAo diameter was compared between individuals with and without a diagnosis of COPD. In the COPD group, the mean AsAo diameter was measured as 40.40±4.90 mm, whereas in the non-COPD group it was 36.01±3.79 mm. This difference was evaluated using the independent Samples t-test and was found to be statistically significant (p=0.00028) (**Figure, Table 2**).

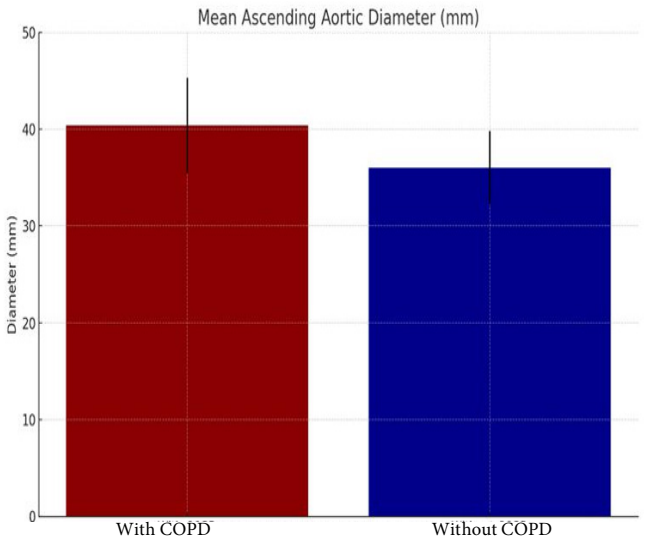


Figure. Comparison of ascending aortic diameter based on the presence of COPD
COPD: Chronic obstructive pulmonary disease

Table 2. Comparison of ascending aortic diameter (AsAo)

Group	AsAo diameter (mm)±SD	p-value
COPD present	40.40±4.90	0.00028
COPD absent	36.01±3.79	0.00028

COPD: Chronic obstructive pulmonary disease, AsAo: Ascending aorta, SD: Standard deviation

In this study, the diameter of the Arcus Ao was compared between individuals with and without a diagnosis of COPD. The mean Arcus Ao diameter in the COPD group was calculated as 32.32±3.48 mm, whereas in the non-COPD group it was 30.89±3.20 mm. The statistical significance of this difference was evaluated using an independent Samples t-test, yielding a p-value of 0.103 (**Table 3**).

Table 3. Comparison of AsAo and aortic arch diameters by COPD group

Aortic segment	COPD present (mm)±SD	COPD absent (mm)±SD	p-value	Statistical significance
AsAo	40.40±4.90	36.01±3.79	0.00028	Significant (p<0.05)
Aortic arch (Arcus Ao)	32.32±3.48	30.89±3.20	0.103	Not significant (p>0.05)

COPD: Chronic obstructive pulmonary disease, AsAo: Ascending aorta, Arcus Ao: Aortic arch, SD: Standard deviation

DISCUSSION

Among the 60 individuals included in this study, 51.7% (n=31) were male and 48.3% (n=29) were female. No statistically significant difference was found in gender distribution between those with and without COPD (p=1.000), indicating that gender was not a significant determinant of COPD status in this sample. However, previous literature widely recognizes male sex as a risk factor for the development of COPD.[14] Furthermore, studies have shown that ascending aortic diameters are significantly greater in men than in women.[12] In our study, due to the relatively balanced gender distribution, the effect of gender on outcomes was not prominently observed.[13]

Participants' ages ranged from 38 to 92 years. The mean age was 73.8±7.92 years in the hypertensive group with COPD and 67.93±12.59 years in the group without COPD. The age

difference between the groups was statistically significant (p=0.036), suggesting that individuals diagnosed with COPD were older, and that age may play a key role in the development of COPD.[6] Piccari et al.[8] reported that age is not only associated with COPD but also with cardiovascular complications. Similarly, Guo et al.[13] highlighted a significant relationship between age and AsAo diameter, underscoring age-related vascular changes.

In the assessment of comorbidities among hypertensive individuals, DM was the most frequently reported condition, with a notable proportion of COPD patients also having DM. Additional comorbidities included congestive HF, asthma, coronary artery disease, hypothyroidism, benign prostatic hyperplasia, and chronic kidney disease. The higher number of comorbidities observed in the COPD group likely reflects the systemic inflammatory burden of COPD, which can affect multiple organ systems. These findings are in line with those of Ando et al.,[7] who reported frequent co-occurrence of aortic aneurysm, HT, and HF in COPD patients. Polman et al.[9] also described a strong association between COPD and cardiometabolic disorders. Piccari et al.[8] further emphasized that vascular dysfunction in COPD may contribute to the overall comorbidity burden.

Regarding lipid profiles, no statistically significant differences were observed between groups in HDL, LDL, or triglyceride levels. In the COPD group, mean HDL was 45.96±11.55 mg/dl, LDL was 97.76±32.45 mg/dl, and triglycerides were 128.83±60.03 mg/dl. In the non-COPD group, these values were 42.31±11.67 mg/dl, 110.09±40.12 mg/dl, and 159.63±68.50 mg/dl, respectively. A borderline significance was observed for triglyceride levels (p=0.069), which may indicate a clinical tendency toward dyslipidemia in COPD. Polman et al.[9] pointed to the high prevalence of metabolic syndrome among individuals with COPD. Fisk et al.[10] proposed that vascular inflammation and atherosclerotic mechanisms contribute to dyslipidemia in COPD. Labaki et al.[24] also demonstrated that CT imaging can help identify dyslipidemia and other chronic diseases in COPD. Nonetheless, some studies have reported inconclusive effects of COPD on lipid profiles.[25]

When examining the AsAo, the diameter was significantly greater in individuals with COPD (40.40±4.90 mm) compared to those without COPD (36.01±3.79 mm) (p=0.00028). This suggests that COPD may contribute to ascending aortic dilatation. Known mechanisms in COPD, such as chronic hypoxemia, oxidative stress, endothelial dysfunction, and elastic fiber degeneration, may promote vascular remodeling.[5,8,14] Fujikura et al.[5] observed that thoracic aortic diameters are notably larger in individuals with COPD, potentially due to emphysematous lung changes. Xiong et al.[15] reported similar findings for abdominal aortic enlargement in COPD. Ng et al.[16] proposed that the pulmonary artery-to-aorta ratio may serve as an indicator of vascular alterations in COPD.

Regarding the Arcus Ao, the average diameter was 32.32±3.48 mm in the COPD group and 30.89±3.20 mm in the non-COPD group. However, the difference was not statistically significant (p=0.103). This suggests that the Arcus Ao may be less affected by COPD-related changes. Anatomically, the aortic arch is more stable and may carry less hemodynamic burden compared to the AsAo.[12,14,17] Hatemi et al.[14] found generalized aortic dilatation in individuals with COPD, but the changes were most prominent in the ascending segment. Similarly, Mensel et al.[12] noted that Arcus Ao diameter is relatively stable despite age and other risk factors.

The overall findings of this study support the concept that COPD is not only a pulmonary disease but also a systemic and cardiovascular condition. In particular, the observed increase in AsAo diameter may reflect the vascular burden imposed by COPD. Additionally, studies have shown that the pulmonary artery-to-aorta ratio is associated with right heart function, underscoring the value of vascular imaging in this population.[19,23] Labaki[24] and Rueda-Ochoa et al.[25] have clearly demonstrated associations between thoracic aortic diameter and both cardiovascular events and mortality.

Limitations

This study has several limitations. First, its retrospective design and relatively small sample size limit the generalizability of the results. In addition, the absence of longitudinal follow-up data and the inability to account for potential confounding factors such as smoking pack-years and medication use should be considered when interpreting the findings. Future prospective studies with larger cohorts are warranted to validate these results and to determine whether COPD should be incorporated into cardiovascular risk stratification models for hypertensive patients. In conclusion, the coexistence of COPD and HT should be carefully evaluated in terms of their effects on cardiovascular structures, and multidisciplinary follow-up strategies should be developed for these individuals.

CONCLUSION

In hypertensive individuals, the presence of COPD appears to negatively affect cardiovascular structures, particularly through the significant enlargement observed in the ascending aorta. This finding underscores that COPD is not solely a respiratory disease, but rather a systemic condition with notable vascular implications. The absence of a significant difference Arcus Ao diameter may be attributed to the anatomically more stable nature of this segment. Continued radiological monitoring and a multidisciplinary approach are essential in individuals with COPD for the early detection of potential cardiovascular complications and for improving clinical outcomes.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was conducted with the approval of the Non-interventional Clinical Researches Ethics Committee of Yozgat Bozok University Faculty of Medicine (Date: 02.07.2025, Decision No: 2025-GOKAEK-2513_2025.07.02_540).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Williams B, Mancia G, Spiering W, et al. 2018 ESC/ESH guidelines for the management of arterial hypertension: the Task Force of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). *Eur Heart J*. 2018;39(33):3021-3104. doi:10.1093/eurheartj/ehy339
2. Covella M, Milan A, Totaro S, et al. Echocardiographic aortic root dilatation in hypertensive patients: a systematic review and meta-analysis. *J Hypertens*. 2014;32(10):1928-1935. doi:10.1097/HJH.0000000000000286
3. DePaolo J, Levin MG, Tcheandjieu C, et al. Relationship between ascending thoracic aortic diameter and blood pressure: a Mendelian randomization study. *Arterioscler Thromb Vasc Biol*. 2023;43(2):359-366. doi:10.1161/ATVBAHA.122.318149
4. Mancusi C, Basile C, Fucile I, et al. Aortic remodeling in patients with arterial hypertension: pathophysiological mechanisms, therapeutic interventions and preventive strategies-a position paper from the heart and hypertension working group of the Italian society of hypertension. *High Blood Press Cardiovasc Prev*. 2025;32(3):255-273. doi:10.1007/s40292-025-00710-3
5. Fujikura K, Albini A, Barr RG, et al. Aortic enlargement in chronic obstructive pulmonary disease (COPD) and emphysema: the Multi-Ethnic Study of Atherosclerosis (MESA) COPD study. *Int J Cardiol*. 2021;331:214-220. doi:10.1016/j.ijcard.2021.02.017
6. Liang X, Chou OHI, Cheung BM. The association between systemic arterial hypertension and chronic obstructive pulmonary disease. Results from the U.S. National Health and Nutrition Examination Survey 1999-2018: a cross-sectional study. *Chronic Obstr Pulm Dis*. 2023;10(2):190-198. doi:10.15326/jcopdf.2022.0306
7. Ando K, Kaneko N, Doi T, Aoshima M, Takahashi K. Prevalence and risk factors of aortic aneurysm in patients with chronic obstructive pulmonary disease. *J Thorac Dis*. 2014;6(10):1388-1395. doi:10.3978/j.issn.2072-1439.2014.10.01
8. Piccari L, Del Pozo R, Blanco I, et al. Association between systemic and pulmonary vascular dysfunction in COPD. *Int J Chron Obstruct Pulmon Dis*. 2020;15:2037-2047. doi:10.2147/COPD.S257679
9. Polman R, Hurst JR, Uysal OF, Mandal S, Linz D, Simons S. Cardiovascular disease and risk in COPD: a state of the art review.*Expert Rev Cardiovasc Ther*. 2024;22(4-5):177-191. doi:10.1080/14779072.2024.2333786
10. Fisk M, Cheriyan J, Mohan D, et al. Vascular inflammation and aortic stiffness: potential mechanisms of increased vascular risk in chronic obstructive pulmonary disease. *Respir Res*. 2018;19(1):100. doi:10.1186/s12931-018-0792-1
11. Chen T, Yang X, Fang X, et al. Potential influencing factors of aortic diameter at specific segments in population with cardiovascular risk. *BMC Cardiovasc Disord*. 2022;22(1):32. doi:10.1186/s12872-022-02479-y
12. Mensel B, Heßelbarth L, Wenzel M, et al. Thoracic and abdominal aortic diameters in a general population: MRI-based reference values and association with age and cardiovascular risk factors. *Eur Radiol*. 2016;26(4):969-978. doi:10.1007/s00330-015-3926-6
13. Guo X, Wang Y, Li H, et al. Relationship between ascending thoracic aortic diameter and blood pressure: a cross-sectional CT study. *Arteriosclerosis, Thrombosis, Vasc Biol*. 2022;42(7):1345-1353. doi:10.1161/ATVBAHA.122.318149
14. Hatemi A, Yao Y, Xu M, et al. Risk factors for dilatation of ascending, arch, and descending aorta: a population-based cohort. *J Am College Cardiol*. 2022;78(7):678-690.
15. Xiong J, Wu Z, Chen C, Guo W. Chronic obstructive pulmonary disease effect on the prevalence and postoperative outcome of abdominal aortic aneurysms: a meta-analysis. *Sci Rep*. 2016;6:25003. doi:10.1038/srep25003
16. Ng CS, Wells AU, Padley SP. A CT sign of chronic pulmonary arterial hypertension: the ratio of main pulmonary artery to aortic diameter. *J Thorac Imaging*. 1999;14(4):270-278. doi:10.1097/00005382-199910000-00007
17. Chettrit D, Amitai OB, Tamir I, et al. PHT-bot: a deep learning system for pulmonary hypertension risk stratification in COPD. *Proceedings of SPIE*. 2019;10950:171-178. doi:10.1117/12.2512469
18. Mohamed Hoesein FA, Besselink T, Pompe E, et al. Accuracy of CT pulmonary artery diameter for pulmonary hypertension in end-stage COPD. *Lung*. 2016;194(5):813-819. doi:10.1007/s00408-016-9926-8
19. Cuttica MJ, Bhatt SP, Rosenberg SR, et al. Pulmonary artery to aorta ratio is associated with cardiac structure and functional changes in mild-to-moderate COPD. *Int J Chron Obstruct Pulmon Dis*. 2017;12:1439-1446. doi:10.2147/COPD.S131413
20. Barr RG. The epidemiology of vascular dysfunction relating to chronic obstructive pulmonary disease and emphysema. *Proc Am Thorac Soc*. 2011;8(6):522-527. doi:10.1513/pats.201101-008MW
21. Andersen KH, Iversen M, Kjaergaard J, et al. Prevalence, predictors, and survival in pulmonary hypertension related to end-stage chronic obstructive pulmonary disease. *J Heart Lung Transplant*. 2012;31(4):373-380. doi:10.1016/j.healun.2011.11.020
22. Terzikhan N, Bos D, Lahousse L, et al. Pulmonary artery to aorta ratio and risk of all-cause mortality in the general population: the Rotterdam study. *Eur Respir J*. 2017;49(6):1602168. doi:10.1183/13993003.02168-2016
23. Janda S, Shahidi N, Gin K, Swiston J. Diagnostic accuracy of echocardiography for pulmonary hypertension: a systematic review and meta-analysis. *Heart*. 2011;97(8):612-622. doi:10.1136/hrt.2010.212084
24. Labaki WW, Agusti A, Bhatt SP, et al. Using CT to detect COPD and comorbidities. *Am J Respiratory Crit Care Med*.2024;210(3):281-287. doi:10.1164/rccm.202402-0407PP
25. Rueda-Ochoa OL, Bons LR, Zhu F, et al. Thoracic aortic diameter and cardiovascular events and mortality among women and men. *Radiology*. 2022;304(1):208-215. doi:10.1148/radiol.210861

Journal of Radiology Medicine in

Review

Doi: 10.51271/JRM-0040

Radiology in Turkiye's healthcare system

Abdullah Enes Berksoy[1], Süleyman Ekiz[2]

[1]Department of Radiology, Göksun State Hospital, Kahramanmaraş, Turkiye
[2]Department of Radiology, Niğde Training and Research Hospital, Niğde, Turkiye

Received: 20.05.2025 ◆ **Accepted**: 30.07.2025 ◆ **Published**: 15.10.2025

Cite this article: Berksoy AE, Ekiz S. Radiology in Turkiye's healthcare system. *J Radiol Med.* 2025;2(4):81-83.

Corresponding Author: Abdullah Enes Berksoy, enesberksoy513@gmail.com

ABSTRACT

Radiology is one of the key departments in healthcare services. To provide effective healthcare to patients, solutions must be developed for the problems in the radiology system. In Turkiye, the demand for radiology is increasing at a rate above the global average. The Turkish healthcare system, which lacks an adequate number of radiology specialists, faces a significant increase in workload. In Turkiye, there are 5 radiologists per 100.000 people, which is below the average of developed countries. According to a 2022 report by the Republic of Turkey's Ministry of Health, there are 295 computed tomography (CT) scans and 207 magnetic resonance imaging (MRI) scans performed per 1.000 people. When compared to organisation for economic co-operation and development (OECD) countries, Turkiye ranks first in the number of radiological examinations per capita. The increasing number of examinations means a higher workload for radiologists. Health policies aim to address these issues, but sufficient progress has not yet been achieved. In this review, we will attempt to identify the causes of this situation and propose solutions.

Keywords: Turkiye (D014421), radiologist (D000072177), health policy (D006291), teleradiology (D019112)

INTRODUCTION

Turkiye's growing population, increasing hospital admissions, efforts to provide healthcare services to more people, and rising healthcare investments are among the main reasons for the increase in radiological examinations. Additionally, clinicians' concerns about malpractice have increased the need and number of imaging studies. Recently, the time allocated per patient has significantly decreased, which is seen as the biggest obstacle to detailed patient evaluation. For all these reasons, the number of radiological examinations is progressively increasing under current conditions.

METHODS

Statistical data obtained from the databases of official institutions in Turkiye and worldwide were retrospectively analyzed. A literature review was conducted on education, healthcare services, and current operations in Turkiye and other countries.

FINDINGS

Healthcare services in Turkiye are provided by both public and private sectors. Until January 25, 2025, healthcare services obtained from public institutions were free of charge except for a $0.20 copayment. In 2022, there were 1.555 hospitals nationwide. Of these, 915 belonged to the Ministry of Health, 68 to universities, and 572 to the private sector. Healthcare services are primarily provided by institutions affiliated with the Ministry of Health. The facilitation of access to healthcare services by the government led to a significant increase in hospital admissions. The number of hospital admissions per capita was recorded as 4.5 in Ministry of Health institutions, 0.5 in universities, and 0.9 in the private sector.

According to the 2022 report, of the total 973 magnetic resonance imaging (MRI) devices, 480 were in the private sector, 379 in the Ministry of Health, and 114 in university hospitals. Similarly, of the total 1.331 computed tomography (CT) devices, 545 were in the private sector, 636 in the Ministry of Health, and 150 in university hospitals. Additionally, of the total 6.255 ultrasonography (USG) devices, 2.431 were in the private sector, 2.715 in the Ministry of Health, and 1.109 in university hospitals.[1]

Basic medical education in Turkiye lasts 6 years. Basic radiology training generally varies and is typically in the form of 1-2 week seminars. In some educational institutions, radiology rotations are not included in the curriculum. In

such cases, newly graduated physicians request unnecessary and incorrect examinations without awareness of radiological tests.

Radiology specialty training is generally provided by different institutions worldwide. As shown in a study conducted in 17 countries, it is provided by local hospitals (23%), universities (36%), national boards (36%), and ministries of health (6%).[2] In Turkiye, radiology specialty training is provided through state universities, foundation universities, and city hospitals.

In Turkiye, radiology specialty training is conducted within a systematic training program framework that defines general and discipline-specific components. The Turkish Radiology Association Competency Board has prepared a comprehensive guide aimed at regulating the quality of radiology specialty training in accordance with national and international standards. This guide specifies the objectives and goals of radiology specialty training.

In Turkiye, radiology specialty training has been implemented as a 5-year program since 2017.[3] This duration varies among countries. For example, in South Korea, this period is 4 years, while in the United Kingdom, it is 5 years.[4,5]

Examinations are a common element in the training process worldwide. Theoretical and practical exams are administered. Pre-specialty structured examinations are conducted, and competency is granted by the academic board.

In Turkiye, residency training includes daily shifts and on-call duties. On average, a resident works 5-8 on-call shifts per month in addition to weekday shifts. Post-call leave is utilized. In other countries, shift durations range from 5 to 26 hours, while the number of monthly on-call shifts ranges from 3 to 6.[2]

Significant differences in radiologist salaries have been observed among countries. Annual salaries range between $8.000 and $75.000 US dollars.

In Europe, there are 13 radiologists per 100.000 people, while in the United Kingdom, this number is 8.5. In Turkiye, there are only 5 radiologists per 100.000 people.[6]

Population growth is one of the reasons for the increase in imaging numbers. According to the address-based population registration system, the country's population was 76.667.864 in 2013, while by the end of 2022, this number was recorded as 85.279.553.[7]

In 2013, the number of USG examinations was 24.819.662; Doppler USG , 8.098.999; CT, 11.039.694; and MRI, 9.073.582. In 2022, these numbers were 20.224.312 for USG, 21.163.138 for Doppler USG, 25.178.827 for CT, and 17.634.202 for MRI (**Table**).[1,8]

The negative effects of increasing examination intensity include rising radiation exposure, increased malpractice risks for radiologists, and higher frequency of contrast agent side effects.

Table. Changes in radiological examination numbers between 2013–2022

Imaging modality	2013	2022	Change
USG	24.819.662	20.224.312	-18.52%
Doppler USG	8.098.999	21.163.138	+161.31%
CT	11.039.694	25.178.827	+128.08%
MRI	9.073.582	17.634.202	+94.35%

USG: Ultrasonography, CT: Computed tomography, MRI: Magnetic resonance imaging

Due to this high demand increase, the Ministry of Health has felt the need to implement various methods. The first of these was to increase the number of radiologists. While the number of research assistants starting radiology specialty training was 1.191 between 2016 and 2021, this number increased to 2.770 during the 2022–2024 period. Another method is to provide radiology services in regions without radiologists through teleradiology. The Teleradiology System enables access to patients' radiological examinations via the internet, remote interpretation of images, exchange of ideas among radiologists, evaluation of image and report quality, transfer to the e-Nabız personal health system, and sharing with citizens and physicians via web and mobile platforms 24/7.[9]

Artificial intelligence (AI) applications, seen as a solution source worldwide and in our country, have gained momentum. Organizations such as the Radiological Society of North America (RSNA), medical image computing and computer-assisted intervention society (MICCAI), and aviation, space and technology festival (TEKNOFEST), or online platforms like grand-challenge, are undertaking various initiatives in this regard.[10] Countries with aging populations, such as Singapore, are integrating AI and deep learning models into their healthcare systems.[11] Our Ministry of Health has recently implemented mammography computer-aided detection (CAD) applications for calcification and early breast cancer detection in mammography.[12]

CONCLUSION

The role of radiology in healthcare services is of undeniable importance. To provide effective healthcare services to patients, solutions must be developed for the problems in the radiology system. The radiology crisis in Turkiye reflects broader healthcare system issues, such as increasing demand, inequalities in resource distribution, and delays in workforce policies.

Although recent steps such as teleradiology and AI pilot applications are promising, a sustainable and multi-stakeholder approach is essential. Access to previously performed radiological examinations should be facilitated through digital platforms. Future studies should evaluate the cost-effectiveness of AI and teleradiology in Turkiye. Teleradiology use should be user-friendly at the national level.

Increasing health literacy in society may reduce the number of hospital visits. Enhancing clinicians' knowledge about examinations and implementing malpractice reforms to create suitable working conditions for clinicians could decrease workload. Reducing the number of visits would

increase the time clinicians allocate per patient. A freer working environment should be provided where clinicians do not feel the need to request radiological examinations for defensive purposes.

Radiology-focused education should be increased in medical curricula, and awareness of radiological tests should be raised. Studies targeting physicians receiving basic medical education suggest that educational activities in radiology can improve the clinical use of imaging.[13]

The number of radiologists per capita should be increased. However, the number of scientific meetings and studies should also be increased to enhance radiologists' qualifications. Existing AI applications should be encouraged, and their integration into the system should be accelerated.

ETHICAL DECLARATIONS

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Republic of Turkey Ministry of Health (RTMH). Accessed date: February 3, 2025. https://dosyasb.saglik.gov.tr/Eklenti/48055/0/siy2022eng050420241pdf.pdf
2. Sosna J, Pyatigorskaya N, Krestin G, et al. International survey on residency programs in radiology: similarities and differences among 17 countries. *Clin Imaging*. 2021;79:230-234. doi:10.1016/j.clinimag.2021.05.011
3. Turkish Medical Specialization Board (TMSB). Accessed date: February 4, 2025. https://tuk.saglik.gov.tr/TR,30486/radyoloji-uzmanlik-egitimi-suresinin-5-yil-olmasi-hakkinda-karar-no--729.html
4. Kassamali RH, Hoey ET. Radiology training in United Kingdom: current status. *Quant Imaging Med Surg*. 2014;4(6):447-448. doi:10.3978/j.issn.2223-4292.2014.10.10
5. Lee JH, Park JS, Lee AL, Lim YJ, Jung SE. Training of radiology residents in Korea. *Korean J Radiol*. 2025;26(4):291-293. doi:10.3348/kjr.2024.1151
6. Radiological Society of North America (RSNA). Accessed date: February 4, 2025. https://www.rsna.org/news/2022/may/global-radiologist-shortage
7. Turkish Statistical Institute (TUIK). Accessed date: February 9, 2025. https://data.tuik.gov.tr/Bulten/Index?p=Adrese-Dayali-Nufus-Kayit-Sistemi-Sonuclari
8. Republic of Turkey Ministry of Health (RTMH). Accessed date: February 3, 2025. https://dosyasb.saglik.gov.tr/Eklenti/45952/0/ekutuphanekitaplarhealthstatisticsyearbook2013pdf.pdf
9. Turkish Radiology Association (TRA). Accessed date: February 9, 2025. https://www.turkrad.org.tr/assets/slider-photos/Radyolojik-Tetkik-Yogunlugu-Raporu.pdf
10. Beşler MS, Koç U. Systematic review of artificial intelligence competitions in radiology: a focus on design, evaluation, and trends. *Diagn Interv Radiol*. 2025. doi:10.4274/dir.2025.243152
11. Goh CX, Ho FC. The growing problem of radiologist shortage: perspectives from Singapore. *Korean J Radiol*. 2023;24(12):1176-1178. doi:10.3348/kjr.2023.0966
12. Yorgancioglu Tarcan G, Yalçın Balçık P, Sebik NB. Artificial intelligence in health services in Turkey and the World. *Mersin Univ Med J*. 2024;14(1):50-60. doi:10.31020/mutftd.1278529
13. Altinmakas E, Dogru OF, Yucel U, et al. Knowing the ABCs: teaching the principles of radiology to medical students in Turkey. *BMC Med Educ*. 2022;22:857. doi:10.1186/s12909-022-03885-8