E-ISSN: 2636-7688

january 2026
volume 33
issue 1

ANNALS OF MEDICAL RESEARCH

The Official Journal of the Inonu University Faculty of Medicine

Original Articles

İNÖNÜ UNIVERSITY PRESS

ANN MED RES

Ann Med Res

Annals of Medical Research

The Official Journal of Inonu University Faculty of Medicine

Editorial Board | Volume: 33 | Issue: 1 | January 2026

Ann Med Res

Annals of Medical Research

The Official Journal of Inonu University Faculty of Medicine

Editorial Advisory Board | Volume: 33 | Issue: 1 | January 2026

Ann Med Res

Annals of Medical Research

The Official Journal of Inonu University Faculty of Medicine

Table of Contents | Volume: 33 | Issue: 1 | January 2026

Original Articles

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

The effect of tartrazine and thymoquinone on the antioxidant system and lipid peroxidation in brain tissue

Zeynep Erdemli [a], Nursena Demircigil [a], Mehmet Erman Erdemli [a,*]

[a] *İnönü University, Faculty of Medicine, Department of Medical Biochemistry, Malatya, Türkiye*
**Corresponding author: ermanerdemli@hotmail.com (Mehmet Erman Erdemli)*

MAIN POINTS

- The relationship between Tartrazine and Thymoquinone in brain tissue was investigated.
- We revealed that Tartrazine induces oxidative stress in brain tissue.
- Thymoquinone has a very strong antioxidant effect and is recommended to be consumed regularly for brain health.

Cite this article as: Erdemli Z, Demircigil N, Erdemli ME. The effect of tartrazine and thymoquinone on the antioxidant system and lipid peroxidation in brain tissue. *Ann Med Res*. 2026;33(1):1--5. doi: 10.5455/annalsmedres.2025.07.181.

ABSTRACT

Aim: To investigate the possible effects of harmful Tartrazine and protective Thymoquinone on brain tissue in Wistar albino rats.

Materials and Methods: Thirty-two rats were divided into 4 groups: Control (n=8), Tartrazine (n=8), Thymoquinone (n=8), and a combination of Tartrazine and Thymoquinone (n=8). Tartrazine and Thymoquinone were administered orally via gavage for 3 weeks, and the brain tissues were collected after the 3-week period. Oxidant-antioxidant parameters such as malondialdehyde (MDA), glutathione (GSH), glutathione peroxidase (GSH-Px), superoxide dismutase (SOD), catalase (CAT) were examined.

Results: There were differences between the Tartrazine and all other groups. Tartrazine administration led to a significant increase in MDA and SOD enzymes activity levels in brain tissue. It also caused a significant decrease in reduced GSH, GSH-Px and CAT activities. Thymoquinone administration caused an increase in GSH, GSH-Px enzymes activity, and CAT enzymes activity levels compared to all other groups.

Conclusion: This study examined the effects of tartrazine and thymoquinone on brain tissue found that tartrazine has neurotoxic effects. We believe that the neurotoxicity caused by tartrazine results from oxidative stress, increased oxidant capacity, and decreased antioxidant capacity. Thymoquinone may serve as a neuroprotective agent because it significantly increased antioxidant capacity.

Keywords: Tartrazine, Thymoquinone, Brain, Oxidative stress, Rats

Received: Aug 14, 2025 **Accepted:** Oct 15, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

Tartrazine is an azo dye with the chemical formula of $C_{16}H_9N_4Na_3O_9S_2$. It is commonly used as a yellow dye [1,2]. It enhances the visual appeal of products while maintaining consistent color. Tartrazine (E 102) is a yellow powder employed as a colorant in the food industry. This compound belongs to the triarylmethane group, a class of organic compounds characterized by azo bonds. It exhibits numerous chemical properties and applications in the cosmetics industry, including makeup products, perfumes, and skincare items. It is also utilized in the pharmaceutical and textile industries. As the global population and food demand grow, various physical tools and chemical products have been developed to support food supply. These social changes have increased the demand for processed foods that can be transported from farms to cities while preserving nutritional value and organoleptic properties [3]. The availability of food additives has also expanded due to advancements in the food industry and the rise in processed food production [4,5]. Tartrazine is present in various food items, such as cakes, candies, canned vegetables, cheeses, gums, sausages, ice cream, orange drinks, salad dressings, seasonal salads, desserts, jams, baked goods, snacks, canned fish, ready- to- serve soups, soft drinks, and ketchup [6]. Tartrazine can negatively affect intestinal microbiota, and it is believed to significantly reduce mucus thickness and contribute to colon cancer, obesity, and diabetes [7,8]. It may have carcinogenic effects and trigger allergic reactions. These reactions include angioedema or chronic urticaria, and symptoms such as atopic dermatitis, flushing, abdominal pain, diarrhea, hypotension, and severe anaphylactoid or anaphylactic reactions similar to asthma. It can also induce hyperactivity in children [9]. Studies on experimen-

tal animals have reported that Tartrazine causes hepatotoxicity, nephrotoxicity, reproductive and developmental toxicity, neurotoxicity, and tissue damage [10-18].

Although endemic to the Eastern Mediterranean, Nigella sativa, a member of the Ranunculaceae family, is also commonly found in Turkey and is an annual plant. It contains nearly forty bioactive compounds, the most active of which is thymoquinone ($C_{10}H_{12}O_2$, 2-isopropyl -5-methyl -1,4-benzoquinone). Studies have shown that thymoquinone has antioxidant, anti-inflammatory, immunoregulatory, anticancer, antimicrobial, neuroprotective, hepatoprotective, nephroprotective, gastroprotective, hypoglycemic, antidiabetic, hypolipidemic, and antihistamine properties, and is effective in treating heart and respiratory system diseases, as well as in inhibiting apoptosis [19-25].

In this research, we aimed to investigate biochemical methods the effects of possible neurotoxicity of tartrazine and protective role of thymoquinone administration on rat brain tissues.

MATERIALS AND METHODS

Rats and experimental groups

The study involved 250 ± 25 g male Thirty-two Wistar albino rats bred at the Inonu University Faculty of Medicine Experimental Animal Breeding and Research Center. After approval by the ethics committee (2020/17-4), the rats were randomly divided into four cages, with 8 rats in each group.

Corn oil was given to the control group. Tartrazine 100 mg/kg/day of was given to the Tartrazine group [26]. Thymoquinone 50 mg/kg/day of was given to the Thymoquinone group [27]. A dose of 100 mg/kg/day Tartrazine and 50 mg/kg/day Thymoquinone was given to the Tartrazine + Thymoquinone group. Administrations were given orally by gavage. All doses were one mL and administered at the same time every day for 21 days. Tartrazine was dissolved in physiological serum, Thymoquinone was dissolved in corn oil.

Tissue collection

After the applications treatments, rats were anesthetized with 10 mg/kg Xylazine (Rompun®, Bayer, Topkapı, Turkey) and 40 mg/kg Ketamine (Ketalar, Pfizer, Istanbul, Turkey) were administered intraperitoneally. Brain tissues from the rats were removed under anesthesia. Samples were quickly stored in a freezer at -75°C.

Preparation of the tissues

Brain tissues were quickly weighed on ice packs, and phosphate buffer was added in a volume nine times their weight. This prepared the tissues for homogenization (IKA ultra turrax T 25 basic). The tissues were homogenized for 1 minute, and a sample was taken from this homogenate for malondialdehyde (MDA) measurement. The homogenate was then centrifuged. The supernatant was obtained by centrifuging for 20 minutes at 650 × g. A sample was taken from this supernatant for the analysis of reduced glutathione (GSH), glutathione peroxidase (GSH-Px), superoxide dismutase (SOD), catalase (CAT), and protein levels.

Measurements

MDA

0.25 mL of brain tissue homogenates was pipetted into glass tubes. To each tube, 1.5 mL of 1% H_3PO_4 and 0.5 mL of 0.6% TBA were added. The tubes were then stored in a water bath at 100 °C for 45 minutes. After cooling under tap water, 2 mL of n-butanol was added. The tubes were vortexed and centrifuged at 4400 rpm at 25 °C for 24 minutes. Readings were taken at 535 nm from the upper section of the glass tube [28].

GSH

125 µl of brain tissue supernatant was pipetted into polypropylene tubes, and 125 µl of tricarboxylic acid was added. The tubes were vortexed and centrifuged at 4400 rpm for 15 minutes at +4 °C. Then, 30 µl of the mixture was transferred to another polypropylene tube. 30 µl of DTNB and 235 µl of Na_2HPO_4 were added, and the tubes were gently vortexed. The samples were then removed, and readings were taken at 410 nm [29].

SOD

500 µl of the brain tissue supernatant was pipetted into polypropylene tubes, and 2450 µl of a mixture containing xanthine, Na_2EDTA, NBT, BSA, and Na_2CO_3, along with 50 µl of XO dissolved in $(NH_4)_2SO_4$, was added. The tubes were vortexed thoroughly, then incubated for 25 minutes. One milliliter of $CuCl_2$ was added to this mixture, and the samples were read at 560 nm [30].

CAT

10 µl of the brain tissue supernatant was pipetted into polypropylene tubes. 2990 µl of a mixture of KH_2PO_4, Na_2HPO_4, and H_2O_2 was added, and the tubes were capped and inverted. Readings were taken at 240 nm. The readings were monitored for one minute [31].

GSH-Px

20 µl of brain tissue supernatant was pipetted into polypropylene tubes. 2650 µl phosphate buffer with EDTA, 10 µl reduced glutathione, 10 µl NADPH, 10 µl GSH reductase, and 10 µl NaN_3 were added, and the tubes were mixed thoroughly. Readings were taken at 340 nm. The readings were monitored for three minutes [32].

Protein

12.5 µl of brain tissue supernatant was pipetted into polypropylene tubes. 237.5 µl of distilled water and 250 µl

https://doi.org/10.5455/annalsmedres.2025.07.181

Table 1. Oxidant -- antioxidant parameters.

	Control (n=8)	Thymoquinone (n=8)	Tartrazine (n=8)	Tartrazine +Thymoquinone (n=8)	p
MDA nmol/gwt tissue	325 (318-365)[a]	327 (307-347)[a]	577 (498-623)[b]	405 (367-472)[c]	≤ 0.05
GSH nmol/gwt tissue	270 (198-315)[a]	341 (294-450)[b]	207 (177-282)[c]	280 (203-312)[a]	≤ 0.05
SOD U/g protein	15 (10-19)[a]	17 (11-22)[a]	21 (15-37)[b]	14 (11-23)[a]	≤ 0.05
CAT K/g protein	9 (3-11)[a]	17 (9-23)[b]	4 (2-7)[c]	12 (5-19)[d]	≤ 0.05
GSH-Px U/g protein	18 (7-24)[a]	21 (16-26)[b]	13 (5-17)[c]	24 (17-32)[d]	≤ 0.05

MDA, malondialdehyde; GSH, reduced glutathione; SOD, superoxide dismutase; CAT, catalase; GSH-Px, glutathione peroxidase. Data are expressed as median (min-max) (n = 8). gwt; gram wet tissue. Groups with different letters in columns are significantly different from each other ($p \leq 0.05$).

of Alkaline copper were added, and the tubes were vortexed thoroughly. 1 ml of phenol was added, and samples were read at 700 nm [33].

Statistical analysis

Quantitative data are presented as medians; minimums and maximums. Intra-group comparisons were conducted using the Kruskal-Wallis test, with the Conover test applied for pairwise comparisons. The significance level was set at $p \leq 0.05$ for all analyses. To compare the four groups at %95 conf idence level (α=0.05) and %80 power (β=0.20) when the effect size is considered to be 0.69, the required minimum sample size for each group is calculated as eight. Each group had eight samples, so a non-parametric hypothesis test was used to compare the groups. The descriptive statistics of the quantitative data are presented as medians, minimum values, maximum values. Independent group comparisons were performed using the Kruskal-Wallis test, and the Conover test was used for pairwise comparisons. The significance level was accepted as 0.05 in all tests. IBM Statistical Software Package for Social Sciences (SPSS) for Windows version 26.0 (Armonk, NY) and https://biostatapps.inonu.edu.tr were employed in statistical analysis.

RESULTS

The biochemical examination of brain tissues revealed significant alterations between the control, Tartrazine, and Thymoquinone groups. Administration of Tartrazine induced a state of oxidative stress, characterized by a significant elevation in MDA levels and SOD activity compared to the control and Thymoquinone groups. This was accompanied by a marked depletion of endogenous antioxidants, including glutathione (GSH) levels and the enzymatic activities of GSH-Px and CAT. In contrast, Thymoquinone demonstrated a potent antioxidant effect, significantly increasing the levels of GSH, GSH-Px, and CAT, while reducing MDA and SOD activity relative to the other treatment groups. Crucially, the co-administration of Thymoquinone with Tartrazine appeared to ameliorate Tartrazine's detrimental effects. This combination group exhibited significantly reduced MDA and SOD activity and elevated levels of GSH, GSH-Px, and CAT when compared to the group receiving Tartrazine alone (Table 1).

DISCUSSION

Oxidative stress plays a crucial role in brain aging. The brain is the most vulnerable organ to oxidative stress because it has a higher oxygen demand than all other organs. This stress causes significant damage to lipids, proteins, and DNA within brain tissue. Such damage contributes to brain aging, nerve cell death, and declining cognitive functions. Furthermore, oxidative stress is a primary cause of neurodegenerative diseases like Alzheimer's, Parkinson's, and Huntington's disease, all involving damage to brain tissue. Essawy et al. developed a model of 7.5 mg/kg/day Tartrazine-induced neurotoxicity over 4 weeks and administered 10 mg/kg/day Melatonin as a protective agent. When examining the rats' brain tissues at the end of the study, they reported that Tartrazine exposure reduced GSH, GSH-Px enzyme activity, and CAT enzyme activity levels compared to other groups [17]. Woopara et al. administered 10 mg/kg of Tartrazine for six weeks and studied the rats' brain tissues. After this period, they observed increased MDA levels and a significant decrease in GSH levels in the Tartrazine group compared to the control [34]. Albasher et al. gave pregnant rats 2.5 mg/kg, 5 mg/kg of Tartrazine through drinking water during pregnancy and for 15 days after birth. Male rats were then selected for brain tissue analysis. They found that Tartrazine increased MDA levels and decreased GSH levels in a dose-dependent manner compared to the control group [35]. Bhatt et al. observed increased MDA levels and decreased CAT enzyme activity levels in brain tissue after administering 7.5 mg/kg of Tartrazine for 21 days [36]. They concluded that Tartrazine raises the oxidant capacity in brain tissue, inducing oxidative stress a key factor in causing nerve cell death and damage in various neurodegenerative and nervous system diseases. They emphasized that this damages brain tissues. Our findings align with these studies, showing that Tartrazine elevates the oxidant capacity and reduces antioxidant capacity in the brain, leading to oxidative stress.

Brain tissue is highly vulnerable to oxidative stress because of its need high oxygen demand, lipid content, limited antioxidant capacity. In the brain, the primary defense against oxidative stress is provided by the antioxidant enzyme system [superoxide dismutase (SOD), glyoxalase, glutathione reductase, glutathione peroxidase, and catalase], along with low molecular weight antioxidants such as glutathione, uric acid, ascorbic acid, and melatonin [37]. Recently, research has focused on

https://doi.org/10.5455/annalsmedres.2025.07.181

antioxidants capable of resisting the development of oxidative stress during neurodegenerative processes in the brain. One promising natural antioxidant from plants is thymoquinone [38]. Shrief et al. gave 20 mg/kg of thymoquinone to rats for 4 weeks as a prophylactic in an $AlCl_3$-induced neurotoxicity model. After 4 weeks, they observed a decrease in brain MDA levels and a significant increase in GSH levels compared to the $AlCl_3$ group [39]. Imam et al. administered 10 mg/kg of thymoquinone as a prophylactic in a cypermethrin-induced neurotoxicity model for 14 days. They noted an increase in GSH levels and a decrease in MDA levels in the thymoquinone group compared to controls [40]. Our findings align with these results. Thymoquinone significantly enhances the brain's antioxidant capacity, mainly by increasing GSH, GSH-Px enzyme activity, CAT enzyme activity levels. Additionally, thymoquinone reduces MDA levels, an indicator of oxidative capacity. Therefore, it helps prevent oxidative stress by decreasing oxidant capacity and boosting antioxidant defenses.

CONCLUSION

Tartrazine caused oxidative stress in brain tissue which led to damage in the brain. Thymoquinone administration resulted in a significant increase in antioxidant capacity. It could serve as a protective agent against Tartrazine-induced brain damage.

Ethics Committee Approval: The study was approved by the İnönü University Faculty of Medicine Animal Experimentation Local Ethics Committee (HADYEK) (no: 2020/17-4).

Informed Consent: Not required for this study.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors declare that they have no conflict of interest.

Author Contributions: Concept: ZE; Design: ND; Supervision: MEE; Fundings: MEE; Materials: ZE; Data Collection and/or Processing: ND; Analysis and/or Interpretation: ND; Literature Review: ZE; Writing: ZE, MEE; Critical Review: MEE.

Financial Disclosure: The research received no financial support.

Artificial Intelligence Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Kumar N, Singh A, Sharma D, et al. Kishore, K. Toxicity of Food Additives. In Food Safety and Human Health, Edited by R.L. Singh, S. Mondal. London, United Kingdom. *Academic Press.* 2019;402. doi: 10.1016/B978-0-12-816333-7.00003-5.
2. Li J, Liu M, Jiang J, et al. Morphology-controlled electrochemical sensing properties of CuS crystals for tartrazine and sunset yellow. *Sens Actuators B.* 2019;288:552-563. doi: 10.1016/j.snb.2019.03.028.
3. Soylak M, Uzcan F. A novel ultrasonication-assisted deep eutectic solvent microextraction procedure for tartrazine at trace levels from environmental samples. *J Iran Chem Soc.* 2020;17(2):461-467. doi: 10.1007/s13738-019-01781-5.
4. de Lima Barizão AC, Silva MF, Andrade M, et al. Green synthesis of iron oxide nanoparticles for tartrazine and bordeaux red dye removal. *J Environ Chem Eng.* 2020;8(1):103618. doi: 10.1016/j.jece.2019.103618.
5. Wu S, Yin ZZ, Chen X, et al. Electropolymerized melamine for simultaneous determination of nitrite and tartrazine. *Food Chem.* 2020;333:127532. doi: 10.1016/j.foodchem.2020.127532.
6. Elhkim MO, Héraud F, Bemrah N, et al. New considerations regarding the risk assessment on Tartrazine An update toxicological assessment, intolerance reactions and maximum theoretical daily intake in France. *Regul Toxicol Pharmacol.* 2007;47(3):308-16. doi: 10.1016/j.yrtph.2006.11.004.
7. Chassaing B, Koren O, Goodrich JK, et al. Dietary emulsifiers impact the Mouse gut microbiota promoting colitis and metabolic syndrome. *Nature.* 2015;519(7541):92-96. doi: 10.1038/nature14232.
8. Cowan TE, Palmnas M, Reimer R, et al. Artifical sweetener consumpption differentially affects the gut microbiota-host metabolic interactions. *The FASEB journal.* 2013;27(1):224-227. doi: 10.1096/fasebj.27.1_supplement.224.7.
9. Skypala IJ, Williams M, Reeves L, et al. Sensitivity to food additives, vasoactive amines and salicylates: a review of the evidence. *Clin Transl Allergy.* 2015;5(34):2-11. doi: 10.1186/s13601-015-0078-3.
10. Demircigil N, Gul M, Gokturk N, et al. Thymoquinone played a protective role against tartrazine-induced hepatotoxicity. *Iran J Basic Med Sci.* 2023;26(1):99-106. doi: 10.22038/IJBMS.2022.67341.14763.
11. Velioglu C, Erdemli ME, Gul M, et al. Protective effect of crocin on food azo dye tartrazine-induced hepatic damage by improving biochemical parameters and oxidative stress biomarkers in rats. *Gen Physiol Biophys.* 2019;38(1):73-82. doi: 10.4149/gpb_2018039.
12. Erdemli Z, Gul M, Gokturk N, et al. Ameliorative effects of thymoquinone on the caspase 3, kidney function and oxidative stress tartrazine-induced nephrotoxicity. *Toxicon.* 2024;241:107660. doi: 10.1016/j.toxicon.2024.107660.
13. Erdemli ME, Gul M, Altinoz E, et al. The protective role of crocin in tartrazine induced nephrotoxicity in Wistar rats. *Biomed Pharmacother.* 2017;96:930-935. doi: 10.1016/j.biopha.2017.11.150.
14. Ezz El-Arab RF, Waly HSA, Al-Salahy MB, et al. Role of gallic acid against hepatic functional and histological deteriorations in tartrazine-intoxicated rats. *Food Chem Toxicol.* 2025;197:115303. doi: 10.1016/j.fct.2025.115303.
15. Öztürk O, Uçar S, Doğanyiğit Z, et al. Evaluation of bone development and organs in rat fetuses exposed to tartrazine. *Heliyon.* 2024;30(1):e41456. doi: 10.1016/j.heliyon.2024.e41456.
16. Mahmoud ES, Abd Elwahab AH, Yousuf AF. The alleviative effect of thymoquinone on cerebellar injury induced by food azo dye in rats. *Al-Azhar Med J.* 2018;47(2):403-421. doi: 10.12816/0052264.
17. Essawy AE, Mohamed AI, Ali RG, et al. Analysis of Melatonin-Modulating Effects Against Tartrazine-Induced Neurotoxicity in Male Rats: Biochemical, Pathological and Immunohistochemical Markers. *Neurochem Res.* 2023;48(1):131-141. doi: 10.1007/s11064-022-03723-9.
18. Erdemli Z, Altinoz E, Erdemli ME, et al. Ameliorative effects of crocin on tartrazine dye-induced pancreatic adverse effects: a biochemical and histological study. *Environ Sci Pollut Res Int.* 2021;28(2):2209-2218. doi: 10.1007/s11356-020-10578-6.
19. Hafez MH, Ez Elarab SM, Tohamy HG, et al. Thymoquinone attenuates diabetes-induced hepatic damage in rat via regulation of oxidative/nitrosative stress, apoptosis, and inflammatory cascade with molecular docking approach. *Sci Rep.* 2024;6:(1):13016. doi: 10.1038/s41598-024-62780-y.
20. Yalçin T, Kaya S. Effect of thymoquinone on hippocampal spexin levels in cisplatin-induced rats. *Neurol Res.* 2025;47(9):817-825. doi: 10.1080/01616412.2025.2504158.
21. Hamaamin KS, Marouf BH. Chemopreventive Efficacy of Thymoquinone in Chemically Induced Urinary Bladder Carcinogenesis in Rat. *Biomed Res Int.* 2022:6276768. doi: 10.1155/2022/6276768.

22. Erdemli ME, Yigitcan B, Erdemli Z, et al. Thymoquinone protection against 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin induced nephrotoxicity in rats. *Biotech Histochem.* 2020;95(8):567-574. doi: 10.1080/10520295.2020.1735520.
23. El-Benhawy SA, Morsi MI, El-Tahan RA, et al. Radioprotective Effect of Thymoquinone in X-irradiated Rats. *Asian Pac J Cancer Prev.* 2021;22(9):3005-3015. doi: 10.31557/APJCP.2021.22.9.3005.
24. Erdemli ME, Yigitcan B, Gul M, et al. Thymoquinone is protective against 2,3,7,8-tetrachlorodibenzo-p-dioxin induced hepatotoxicity. *Biotech Histochem.* 2018;93(6):453-462. doi: 10.1080/10520295.2018.1453549.
25. Kapan M, Tekin R, Onder A, et al. Thymoquinone ameliorates bacterial translocation and inflammatory response in rats with intestinal obstruction. *Int J Surg.* 2012;10(9):484-488. doi: 10.1016/j.ijsu.2012.06.006.
26. Balta I, Sevastre B, Mireşan V, et al. Protective effect of blackthorn fruits (Prunus spinosa) against tartrazine toxicity development in albino Wistar rats. *BMC Chem.* 2019;13(1):104. doi: 10.1186/s13065-019-0610-y.
27. Kong LY, Li GP, Yang P, et al. Protective effect of thymoquinone on cholestatic rats with liver injury. *Genet Mol Res.* 2015;14(4):12247–12253. PMID: 26505373.
28. Ohkawa H, Ohishi N, Yagi K. Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. *Anal Biochem.* 1979;95(2):351–358. doi: 10.1016/0003-2697(79)90738-3.
29. Ellman GL. Tissue sulphydryl groups. *Arch Biochem Biophys.* 1959;82(1):70–77. doi: 10.1016/0003-9861(59)90090-6.
30. Paglia D, Valentine WN. Studies on the quantitative and qualitative charecterization of erythrocyte glutathione peroxidase. *J Lab Clin Med.* 1967;70:158-169. PMID: 6066618.
31. Sun Y, Oberley LW, Li Y. A simple method for clinical assay of superoxide dismutase. *Clin Chem.* 1988;34(3):497–500. doi: 10.1093/clinchem/34.3.497.
32. Aebi H, Bergmeyer HU. Methods of Enzymatic Analysis. Weinheim/NewYork. *Academic Press.* 1974;2:673-680. doi: 10.1016/b978-0-12-091302-2.50032-3.
33. Lowry OH, Rosebrough NJ, Farr AL, et al. Protein measurement with the Folin phenol reagent. *J Biol Chem.* 1951;193(1):265–275. doi: 10.1016/S0021-9258(19)52451-6.
34. Wopara I, Modo EU, Adebayo OG, et al. Anxiogenic and memory impairment effect of food color exposure: upregulation of oxido-neuroinflammatory markers and acetyl-cholinestrase activity in the prefrontal cortex and hippocampus. *Heliyon.* 2021;17(3):e06378. doi: 10.1016/j.heliyon.2021.e06378.
35. Albasher G, Maashi N, Alfarraj S, et al. Perinatal Exposure to Tartrazine Triggers Oxidative Stress and Neurobehavioral Alterations in Mice Offspring. *Antioxidants (Basel).* 2020;9(1):53. doi: 10.3390/antiox9010053.
36. Bhatt D, Vyas K, Singh S, et al. Tartrazine induced neurobiochemical alterations in rat brain sub-regions. *Food Chem Toxicol.* 2018;113:322-327. doi: 10.1016/j.fct.2018.02.011.
37. Hulbert AJ, Pamplona R, Buffenstein R, et al. Life and death: metabolic rate, membrane composition, and life span of animals. *Physiol Rev.* 2007;87(4):1175-213. doi: 10.1152/physrev.00047.2006.
38. Ahmad N, Ahmad R, Alam MA, et al. Quantification and evaluation of thymoquinone loaded mucoadhesive nanoemulsion for treatment of cerebral ischemia. *Int J Biol Macromol.* 2016;88:320-332. doi: 10.1016/j.ijbiomac.2016.03.019.
39. Shrief AI, Elshenawy DS, Elsukary AE, et al. Behavioral and histological study on the neuroprotective effect of thymoquinone on the cerebellum in AlCl3-induced neurotoxicity in rats through modulation of oxidative stress, apoptosis, and autophagy. *J Mol Histol.* 2025;56(2):81. doi: 10.1007/s10735-025-10361-2.
40. Imam AL, Okesina AA, Sulaimon FA, et al. Thymoquinone ameliorate oxidative stress, GABAergic neuronal depletion and memory impairment through Nrf2/ARE signaling pathway in the dentate gyrus following cypermethrin administration. *BMC Neurosci.* 2024;25(1):45. doi: 10.1186/s12868-024-00896-7.

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Comparison of neuroendocrine responses between caudal block and local infiltration techniques in pediatric inguinal hernia surgery: A prospective comparative clinical study

Olcay Dilken [a], Mustafa Altinay [a], Batuhan Bostan [a,*], Feyzullah Orkut Duran [a], Betul Duran [a], Sevde Sumeyye Iraz Soyalp [a], Merve Ipekoglu [a], Ayse Surhan Cinar [a]

[a] *University of Health Sciences, Şişli Hamidiye Etfal Training and Research Hospital, Department of Anesthesiology and Reanimation, Istanbul, Türkiye*

**Corresponding author: batubstn@gmail.com (Batuhan Bostan)*

MAIN POINTS

- Caudal block and local infiltration produced comparable postoperative cortisol and prolactin responses, as well as similar CHEOPS pain scores.
- Anesthesia and surgical durations were significantly longer in the caudal block group.
- Local wound infiltration may serve as a practical alternative due to its shorter application time and ease of use.
- A significant positive correlation was observed between preoperative and postoperative prolactin levels, independent of the analgesic technique.

Cite this article as: Dilken O, Altinay M, Bostan B, Duran FO, Duran B, Iraz Soyalp SS, Ipekoglu M, Surhan Cinar A. Comparison of neuroendocrine responses between caudal block and local infiltration techniques in pediatric inguinal hernia surgery: A prospective comparative clinical study. *Ann Med Res*. 2026;33(1):6–12. doi: 10.5455/annalsmedres.2025.07.202.

ABSTRACT

Aim: This study aimed to compare the effects of caudal anesthesia and local wound infiltration on postoperative stress hormone responses (cortisol and prolactin) and pain scores in pediatric patients undergoing inguinal hernia repair.

Materials and Methods: A total of 64 children aged 1–4 years, classified as ASA I, scheduled for elective unilateral inguinal hernia surgery were randomized into two groups: caudal anesthesia (Group C) and local infiltration (Group I). All patients received general anesthesia. Group C received 1 mL/kg of 0.25% bupivacaine via the caudal route, while Group I received 0.25 mL/kg of 5 mg/mL bupivacaine via 23-gauge needle infiltration. Postoperative pain was assessed using the CHEOPS scale at 1st, 2nd, 3rd, and 4th hours by a blinded anesthesiologist. Blood samples for cortisol and prolactin levels were collected before induction and 40 minutes after surgery and analyzed by ECLIA.

Results: No significant differences were found between groups in postoperative cortisol and prolactin levels. Pain scores were also similar at all time points ($p>0.05$). A significant positive correlation was observed between preoperative and postoperative prolactin levels ($p<0.05$). Anesthesia and surgical durations were longer in the caudal group ($p<0.05$), while fentanyl requirements were similar.

Conclusion: Both caudal block and wound infiltration provide comparable postoperative analgesia and hormonal stress response control in pediatric inguinal hernia surgeries. Given its shorter application time, local infiltration may offer a practical alternative in suitable clinical settings.

Keywords: Postoperative pain, Inguinal hernia, Wound infiltration, Caudal block, Stress hormone

Received: Jul 25, 2025 **Accepted:** Sep 30, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

Pediatric inguinal hernia is among the common congenital anomalies requiring surgical treatment, particularly during the neonatal and infantile periods. It is observed 4–10 times more frequently in boys compared to girls, with an incidence ranging from 0.8% to 5% in the general pediatric population [1]. In premature infants, this rate may rise to as high as 30% [2]. These cases are typically indirect hernias, developing due to delayed obliteration of the processus vaginalis. Diagnosis is usually established through physical examination, and surgical intervention is necessary to prevent potential complications such as incarcerated hernia, bowel ischemia, or perforation [1]. Early surgical repair is critical in reducing morbidity and preventing such complications.

Due to the physiological and psychological differences in children, postoperative pain management in these surgical inter-

ventions is of great importance. Although pain assessment in pediatric patients is inherently challenging, inadequate analgesia can prolong recovery, extend hospital stay, and has been linked to the development of anxiety, behavioral disorders, and chronic pain in the long term [3]. Currently, multimodal approaches are preferred for postoperative pain management, including opioids, paracetamol, nonsteroidal anti-inflammatory drugs (NSAIDs), regional nerve blocks, and local anesthetic infiltration techniques [3]. However, due to the adverse effects associated with opioids—such as nausea, vomiting, respiratory depression, and sedation—safer and minimally invasive techniques are often favored in children [4].

In this context, caudal anesthesia and local wound infiltration are two commonly used techniques in pediatric surgical procedures. Caudal block is a regional anesthesia technique performed by administering local anesthetic into the epidural space via the sacral hiatus and is widely used in surgeries under the umbilicus [5]. When administered with long-acting local anesthetics such as ropivacaine or levobupivacaine, it provides effective and safe analgesia [5]. Moreover, caudal block has been shown to significantly reduce stress hormone levels such as plasma cortisol and prolactin in the postoperative period [6]. On the other hand, local wound infiltration involves the administration of local anesthetics around the surgical incision and is a non-invasive and easily applicable method. This technique has also been reported to reduce cortisol levels and to be effective in postoperative pain control in several studies [7]. However, it remains unclear which of these two methods is more effective in modulating the stress response, and the literature presents conflicting results on this topic [8].

This study aims to compare the effects of caudal anesthesia and local wound infiltration on postoperative stress hormone levels in pediatric patients undergoing inguinal hernia surgery. The findings are expected to contribute to the identification of the most appropriate analgesic method in this patient population.

MATERIALS AND METHODS

This study was approved by the Ethics Committee of Şişli Etfal Training and Research Hospital on 23.06.2009, with protocol number 73, and was conducted between July 2009 and July 2010 in the Departments of Pediatric Surgery and Anesthesiology at Şişli Etfal Training and Research Hospital Hospital. All procedures adhered to the ethical standards outlined in the Declaration of Helsinki (2008) and were reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines.

Patient selection

Children aged 1–4 years scheduled for elective unilateral inguinal hernia repair and classified as ASA I according to the American Society of Anesthesiologists Physical Status Classification were included. Exclusion criteria were: history of prematurity, neurological disorders, chronic medication use, known allergy to local anesthetics, previous abdominal surgery, bleeding disorders, bilateral hernia, or recurrent hernia.

Sample size

Sample size was calculated based on expected variations in cortisol levels and effect sizes reported in similar studies [9]. Assuming $\alpha = 0.05$, $\beta = 0.20$ (power = 80%), and a between-group difference of 1 standard deviation (effect size $d \approx 1.0$), a minimum of 32 patients per group was required. Accordingly, 64 patients were enrolled (Group C = 32, Group I = 32), achieving the targeted statistical power.

Anesthesia and monitoring

No premedication was administered. Age, sex, and body weight were recorded preoperatively. Standard monitoring included noninvasive arterial blood pressure, ECG, peripheral oxygen saturation (SpO_2), and end-tidal CO_2. Anesthesia was delivered using a Julian® anesthesia machine (Dräger, Lübeck, Germany). Induction was achieved with intravenous propofol (1 mg/kg), followed by insertion of a laryngeal mask airway. Fentanyl (1 mcg/kg IV) was administered, and anesthesia was maintained with 2% sevoflurane in 50% air/oxygen.

Fluid management involved 5% dextrose in 0.2% sodium chloride solution (Isolyte® P). If heart rate or blood pressure increased more than 20% above baseline, an additional dose of fentanyl (0.5 mcg/kg IV) was administered and recorded as part of total fentanyl consumption. All surgical procedures were performed by the same pediatric surgical team, with hemodynamic stability maintained throughout.

Randomization and analgesia protocol

Patients were randomized using a computer-generated block randomization method, ensuring equal allocation between groups. Group C (caudal group) received caudal anesthesia with 0.25% bupivacaine (Chirocaine®, 5 mg/mL; Abbott S.p.A, Latina, Italy) at a dose of 1 mL/kg. The required volume was prepared by diluting the commercially available 0.5% (5 mg/mL) solution 1:1 with sterile 0.9% sodium chloride to achieve a final concentration of 0.25%. This solution was then drawn into a sterile syringe and administered via a 22-gauge short-beveled needle through the sacral hiatus, with the patient placed in the lateral decubitus position, under strict sterile conditions and after confirming negative aspiration.

Group I (infiltration group) received local wound infiltration using the same 0.25% bupivacaine solution. A volume of 0.25 mL/kg was infiltrated into the planned surgical site by the pediatric surgeon prior to skin incision, using a 23-gauge needle. The preparation of the infiltration solution followed the same dilution protocol as in Group C to ensure consistency in drug concentration across groups.

All procedures were performed by the same pediatric surgical team. Analgesia was administered after randomization and

https://doi.org/10.5455/annalsmedres.2025.07.202

prior to surgical incision by either an anesthesiologist (Group C) or the surgeon (Group I). No crossover occurred between groups, and the integrity of the intervention allocation was maintained throughout the study.

Postoperative management

After the procedure, anesthesia was discontinued, and patients were extubated once spontaneous breathing resumed, and adequate respiratory parameters were achieved. Patients were then transferred to the Post-Anesthesia Care Unit (PACU) in stable condition. Anesthesia duration was defined as the time from induction to extubation, and surgical duration was defined as the time from induction to final suture placement—both recorded by the attending anesthesiologist.

In the PACU, noninvasive blood pressure, heart rate, respiratory rate, SpO_2, and consciousness were monitored. Postoperative pain was assessed using the Children's Hospital of Eastern Ontario Pain Scale (CHEOPS) by a blinded anesthesiologist at the 1st, 2nd, 3rd, and 4th postoperative hours. Patients were transferred to the ward after stabilization.

Hormone measurement

Blood samples for cortisol and prolactin measurements were obtained twice: immediately before induction and 40 minutes after the end of surgery. Samples were collected in EDTA tubes, centrifuged, and stored at –20°C until analysis. Plasma cortisol (normal ranges: 1–39 µg/dL in the morning, 3–18 µg/dL in the afternoon) and prolactin levels (girls: 3.2–25.3 ng/mL; boys: 2.9–17.1 ng/mL) were measured using electrochemiluminescence immunoassay (ECLIA) on an Elecsys E-170 analyzer (Roche Diagnostics, Basel, Switzerland).

Statistical analysis

Descriptive statistics were presented as mean ± standard deviation (SD) or frequency (%). Categorical variables were compared using the chi-square test. Continuous variables were compared using independent samples t-test under equal variance assumptions. Distribution of the data was assessed with the Shapiro-Wilk test. Postoperative parameters were analyzed with Repeated Measures ANOVA with Bonferroni correction. Repeated measures were analyzed using linear mixed-effects models with random intercepts per patient. Postoperative biomarkers were analyzed using linear regression, with baseline values and group interaction terms as covariates. A two-tailed p-value < 0.05 was considered statistically significant. All analyses were performed using R software version 4.4.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

A total of 64 pediatric patients were included in the study (Figure 1). There were no statistically significant differences between the two groups in terms of sex distribution, age, or body weight (p>0.05). Baseline demographic characteristics of the participants are summarized in Table 1.

In the caudal block group, both mean anesthesia duration and surgical time were significantly longer compared to the infil-

Table 1. Baseline characteristics by treatment.

Variable	Caudal N = 32[1]	Infiltration N = 32[1]	p-value[2]
Sex			0.4
F	7 (22%)	11 (34%)	
M	25 (78%)	21 (66%)	
Age (months)	30 (19)	33 (26)	0.5
Body Weight (kg)	12.7 (3.4)	13.7 (5.3)	0.4

[1] n (%); Mean (SD), [2] Pearson's Chi-squared test; Two Sample t-test

Table 2. Fentanyl dose, anesthesia and operation durations by group.

Variable	Caudal[1]	Infiltration[1]	p-value[2]
Mean Anesthesia Duration (mins)	71 (36)	50 (23)	0.007
Mean Operation Duration (mins)	60 (33)	46 (18)	0.047
Mean Fentanyl Dose (mcg)	13.2 (4.1)	13.3 (7.7)	>0.9

[1] Mean (SD); [2] Two Sample t-test

Table 3. CHEOPS by hour and groups.

Hours After Operation	Caudal[1]	Infiltration[1]	p-value[2]
1	8.16 (2.50)	7.81 (1.99)	0.5
2	7.66 (2.24)	7.59 (2.00)	>0.9
3	7.78 (2.15)	7.84 (2.32)	>0.9
4	7.66 (2.06)	7.78 (2.04)	0.8

[1] Cheops: Mean (SD); [2] Two-way ANOVA with Bonferroni correction.

Table 4. Results of linear regression analysis of variables affecting postoperative cortisol levels.

Characteristic	Beta	95% CI[1]	p-value
Preoperative Cortisol Level	0.31	-0.23, 0.84	0.3
Group			
Caudal	---	---	
Infiltration	-4.6	-15, 6.0	0.4
Preoperative Cortisol Level * Group			
Preoperative Cortisol Level * Infiltration	0.28	-0.44, 1.0	0.4

[1] CI = Confidence Interval.

Table 5. Results of linear regression analysis of variables affecting postoperative prolactin levels.

Characteristic	Beta	95% CI[1]	p-value
Preoperative Prolactin Level	0.42	0.22, 0.62	<0.001
Group			
Caudal	---	---	
Infiltration	9.6	-4.6, 24	0.2
Preoperative Prolactin Level * Group			
Preoperative Prolactin Level * Infiltration	-0.35	-0.93, 0.24	0.2

[1] CI = Confidence Interval.

https://doi.org/10.5455/annalsmedres.2025.07.202

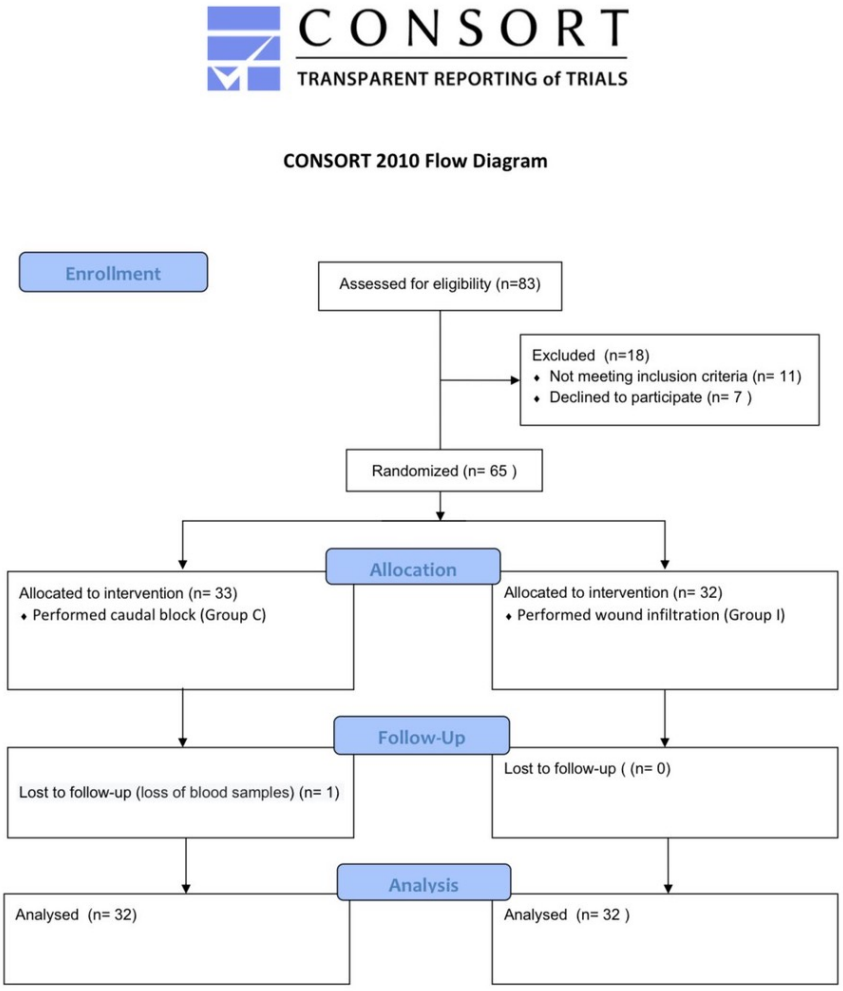

Figure 1. CONSORT flow diagram of the study.

tration group. The difference in anesthesia duration was –21 minutes (95% CI: –36 to –6, p = 0.007), while the difference in surgical duration was –14 minutes (95% CI: –26 to 0, p = 0.047) (Table 2). The mean fentanyl dose was similar between the groups (difference: 0.1 mcg, 95% CI: –3 to +3, p>0.9).

Postoperative CHEOPS pain scores did not differ significantly between the two groups. The mean difference in scores was –0.4, which was not statistically significant (95% CI: –1.7 to 0.8, p = 0.5). Additionally, time had no significant effect on CHEOPS scores (hour effect: –0.14, 95% CI: –0.5 to +0.2, p = 0.4) (Table 3).

Neither the type of analgesia nor preoperative cortisol levels had a significant effect on postoperative cortisol concentrations (Figure 2, Table 4). Similarly, the type of analgesia did not influence postoperative prolactin levels. However, preoperative prolactin levels showed a positive correlation with postoperative levels (Figure 3, Table 5).

DISCUSSION

This study aimed to compare the effects of caudal block and local wound infiltration on postoperative stress hormone response (cortisol and prolactin) and analgesic efficacy in children aged 1–4 years undergoing inguinal hernia repair. Our findings demonstrate that both techniques provided similar postoperative pain control and did not produce significant differences in cortisol and prolactin levels. Notably, anesthesia and surgical durations were longer in the caudal group.

Caudal anesthesia is a widely used technique in pediatric patients, particularly in subumbilical surgeries, and has been shown in several studies to improve postoperative analgesia and reduce pain scores. For instance, Ingelmo et al. reported that caudal administration of ropivacaine significantly reduced plasma prolactin and cortisol levels [6]. However, that study did not emphasize pain scores and was not conducted under general anesthesia, limiting direct comparison

https://doi.org/10.5455/annalsmedres.2025.07.202

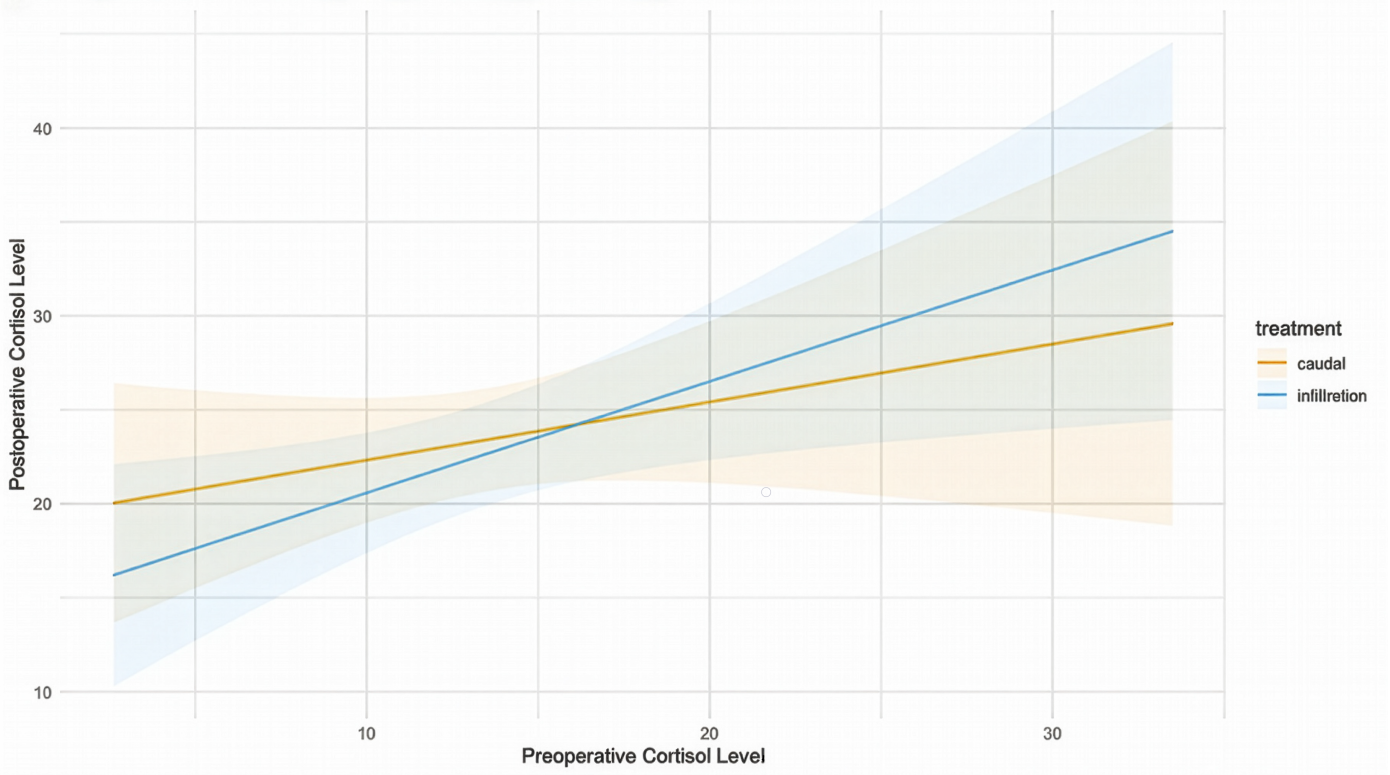

Figure 2. Interaction between preoperative and postoperative cortisol levels according to type of analgesia (caudal vs. infiltration). Shaded areas represent 95% confidence intervals. (No statistically significant interaction was observed between preoperative cortisol levels and type of analgesia (p = 0.4)).

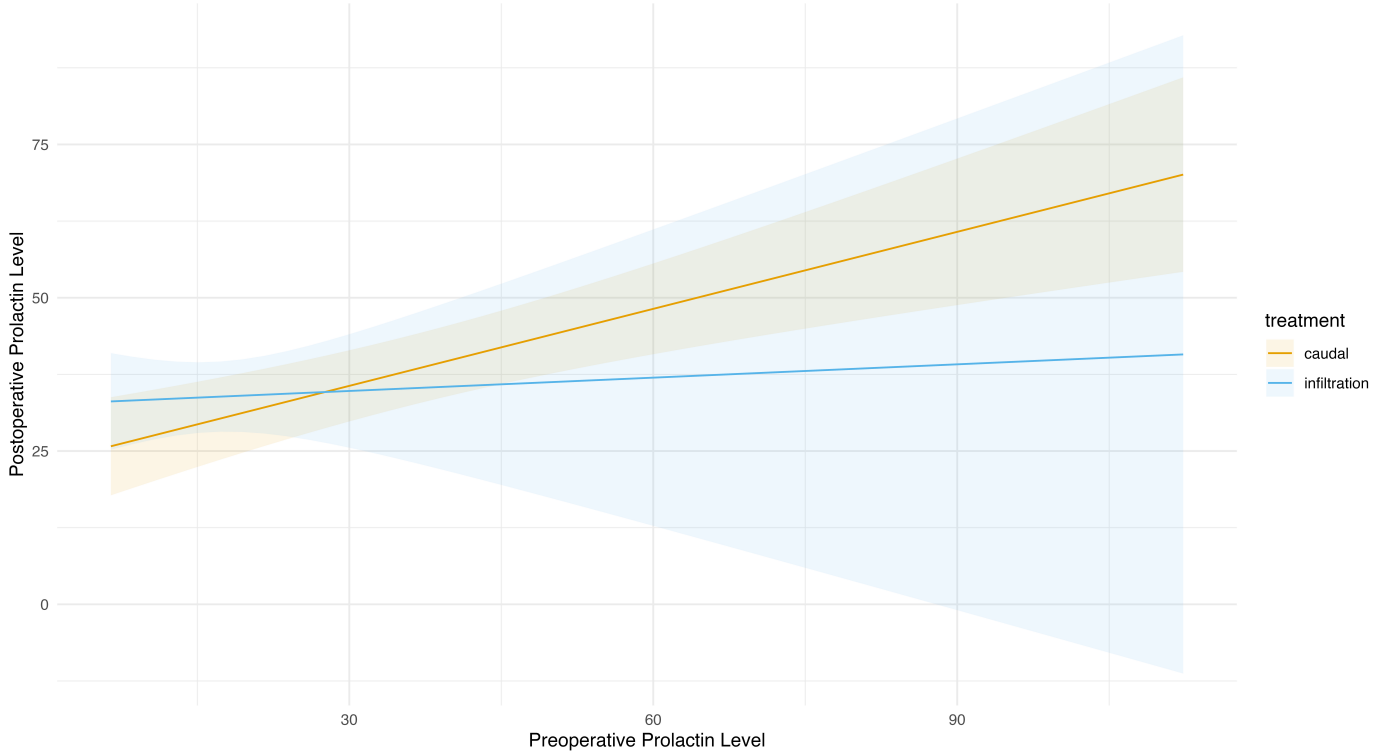

Figure 3. Relationship between preoperative and postoperative prolactin levels by analgesia type (caudal vs. infiltration). Shaded areas represent 95% confidence intervals. (A significant positive association was observed between preoperative and postoperative prolactin levels, regardless of analgesia type (p<0.05)).

https://doi.org/10.5455/annalsmedres.2025.07.202

with our results.

In our study, both groups received general anesthesia, allowing for a more isolated assessment of the analgesic technique's effect. This likely explains the similar pain scores between groups. Çınar et al. reported a significant decrease in cortisol levels in children receiving levobupivacaine via wound infiltration [7]. While a similar hormonal suppression was observed in our study, no statistically significant difference was found between the two techniques.

Another study by Sakellaris et al. reported a significant reduction in postoperative prolactin levels in children who received ropivacaine wound infiltration [8]. Our findings align with this in the infiltration group; however, we also observed a positive correlation between preoperative and postoperative prolactin levels, suggesting that individual hormonal responses may be partially predictable.

In a randomized controlled trial by Gavrilovska-Brzanov et al. caudal block and local infiltration were compared, revealing that although caudal block provided longer analgesia, both techniques similarly reduced cortisol levels [9]. These findings support our results, indicating that both methods, when appropriately applied, can effectively modulate postoperative stress responses.

Park and Lee also found that wound infiltration shortened overall procedure time and was easier to administer [10]. In our study, the longer time required for positioning and performing the caudal block led to prolonged anesthesia and surgical durations, highlighting an important practical consideration.

There are several limitations to our study. It was not blinded, and the sample size was relatively small. However, the use of a consistent surgical team and blinded pain assessment by an independent anesthesiologist mitigate some of these limitations. Additionally, hormonal levels were measured only at two time points (preoperative and 40 minutes postoperative); longer-term follow-up could provide further insight into hormonal adaptation dynamics.

In conclusion, our findings show that both caudal block and local wound infiltration are similarly effective in managing postoperative pain and stress hormone responses in pediatric patients undergoing inguinal hernia repair. In clinical practice, the choice of analgesic technique should be guided by patient characteristics, ease of application, and procedural duration. Although caudal block may offer longer-lasting analgesia, wound infiltration remains a technically simpler and time-efficient alternative. Both techniques can be safely utilized in appropriate clinical settings.

CONCLUSION

This study demonstrated that caudal anesthesia and local wound infiltration provide comparable efficacy in postoperative pain control and hormonal stress response in pediatric inguinal hernia surgeries. While no significant differences were observed in cortisol or prolactin levels, or in pain scores and fentanyl consumption, our findings offer several clinically relevant insights.

Notably, caudal anesthesia required significantly longer anesthesia and surgical time, likely due to positioning and procedural complexity, which may affect workflow and resource allocation in busy clinical settings. In contrast, local wound infiltration is technically simpler, more time-efficient, and can be performed directly by the surgeon—factors that may contribute to lower costs and increased procedural autonomy.

Moreover, caudal block necessitates lateral positioning of the child, potentially introducing a confounding variable in hormonal stress response due to physical manipulation or procedural anxiety. Additionally, the circadian rhythm of stress hormones—particularly cortisol—raises the possibility that timing of sample collection may influence results. Future studies should account for these factors, including economic analysis, timing-related hormonal fluctuations, and role differentiation between anesthesia and surgical teams, to optimize analgesic strategies in pediatric practice.

Ethics Committee Approval: This study was approved by the Ethics Committee of Şişli Etfal Training and Research Hospital (Protocol No: 73, Date: 23.06.2009).

Informed Consent: Written and verbal consent was obtained from the parents of the children.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors declare that they have no conflicts of interest.

Author Contributions: Concept: O.D, A.S.Ç.; Design: M.A., S.S.IS.Ç., A.S.Ç.; Supervision: O.D., M.A., B.B., F.O.D., B.D., S.S.I.Ç., M.İ., A.S.Ç.; Materials: B.D., M.İ.; Data Collection and/or Processing: M.A., F.O.D., M.İ.; Analysis and/or Interpretation: O.D., B.B., F.O.D., B.D.; Literature Review: S.S.I.Ç.; Writing: O.D., B.B., F.O.D., S.S.I.Ç., M.İ.; Critical Review: M.A., B.B., B.D., A.S.Ç.

Financial Disclosure: The authors received no financial support for this study.

Artificial Intelligence Disclosure: During the preparation of this manuscript, the authors used ChatGPT (OpenAI; GPT-4) solely for language editing and improving readability (grammar, sentence structure, and clarity) in selected parts of the manuscript (Abstract, Introduction, Discussion). The AI tool was not used for data analysis, statistical processing, interpretation of results, or generation of scientific conclusions. All AI-assisted content was reviewed, verified, and edited by the authors. The authors retain full responsibility for the integrity, accuracy, and originality of the work.

REFERENCES

1. Yeap E, Pacilli M, Nataraja RM. Inguinal hernias in children. *Aust J Gen Pract.* 2020;49(1–2):38–43. doi: 10.31128/AJGP-08-19-5037.

2. Sharma S, Curry J. Inguinal hernia in premature infants. *J Indian Assoc Pediatr Surg.* 2015;20(1):25-26. PMID: 25552827.
3. Chou R, Gordon DB, de Leon-Casasola OA, et al. Management of Postoperative Pain: A Clinical Practice Guideline From the American Pain Society, the American Society of Regional Anesthesia and Pain Medicine, and the American Society of Anesthesiologists' Committee on Regional Anesthesia, Executive Committee, and Administrative Council. *J Pain.* 2016;17(2):131-157. doi: 10.1016/j.jpain.2015.12.008.
4. Anderson BJ, Palmer GM. Recent developments in the pharmacological management of pain in children. *Curr Opin Anaesthesiol.* 2006;19(3):285–292. doi: 10.1097/01.aco.0000192802.33291.6f.
5. Ecoffey C. Pediatric regional anesthesia – update. *Curr Opin Anaesthesiol.* 2007;20(3):232–235. doi: 10.1097/ACO.0b013e328105da3f.
6. Solak M, Ulusoy H, Sarihan H. Effects of caudal block on cortisol and prolactin responses to postoperative pain in children. *Eur J Pediatr Surg.* 2000;10(4):219-223. doi: 10.1055/s-2008-1072362.
7. Çınar SÖ, Kum Ü, Cevizci N, et al. Effects of levobupivacaine infiltration on postoperative analgesia and stress response in children following inguinal hernia repair. *Eur J Anaesthesiol.* 2009;26(5):430–434. doi: 10.1097/EJA.0b013e32832974fc.
8. Sakellaris G, Petrakis I, Makatounaki K, Arbiros I, Karkavitsas N, Charissis G. Effects of ropivacaine infiltration on cortisol and prolactin responses to postoperative pain after inguinal hernioraphy in children. *J Pediatr Surg.* 2004;39(9):1400-1403. doi: 10.1016/j.jpedsurg.2004.05.002.
9. Gavrilovska-Brzanov A, Kuzmanovska B, Kartalov A, et al. Evaluation of Anesthesia Profile in Pediatric Patients after Inguinal Hernia Repair with Caudal Block or Local Wound Infiltration. *Open Access Maced J Med Sci.* 2016;4(1):89-93. doi: 10.3889/oamjms.2016.023.
10. Markham SJ, Tomlinson J, Hain WR. Ilioinguinal nerve block in children. A comparison with caudal block for intra and postoperative analgesia. *Anaesthesia.* 1986;41(11):1098-1103. doi: 10.1111/j.1365-2044.1986.tb12956.x.

https://doi.org/10.5455/annalsmedres.2025.07.202

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Prevalence of asthma and allergic diseases and associated risk factors among primary school children in the Malatya region

Murat Caglar Sahin [a,*], Ercan Yilmaz [b], Erdem Topal [b]

[a] İnönü University, Faculty of Medicine, Department of Pediatrics, Malatya, Türkiye
[b] İnönü University, Faculty of Medicine, Department of Pediatrics, Division of Allergy and Immunology, Malatya, Türkiye
*Corresponding author: muratcaglarsahin@gmail.com (Murat Caglar Sahin)

MAIN POINTS

- Among 6–7-year-old children in Malatya, the prevalence of physician-diagnosed asthma was 4.2%, while 18.6% experienced wheezing in the past year.
- Allergic rhinitis symptoms affected 29.9% of participants and showed a strong association with a positive family history.
- Atopic dermatitis was reported in 7.9% of children, significantly linked to parental history of atopic dermatitis.
- Prematurity, together with family history of allergic diseases, emerged as the most consistent risk factors.
- This study provides updated regional epidemiological data that may guide preventive strategies and healthcare planning.

Cite this article as: Sahin MC, Yilmaz E, Topal E. Prevalence of asthma and allergic diseases and associated risk factors among primary school children in the Malatya region. *Ann Med Res*. 2026;33(1):13–18. doi: 10.5455/annalsmedres.2025.08.229.

ABSTRACT

Aim: In recent years, a significant increase has been observed in the prevalence of allergic diseases during childhood. This study aims to determine the prevalence of asthma, allergic rhinitis, and atopic dermatitis among children aged 6--7 living in the city center of Malatya, as well as to analyze potential environmental and familial risk factors associated with these conditions.

Materials and Methods: The study was conducted between December 2023 and January 2024 in primary schools located in central Malatya, using the International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire protocol. A total of 431 children were included. Participants were evaluated based on demographic characteristics, environmental exposures, family history, and symptoms related to asthma, allergic rhinitis, and atopic dermatitis. The collected data were analyzed statistically.

Results: Lifetime wheezing was reported in 20% of the children, while 18.6% reported wheezing within the past 12 months. The prevalence of physician-diagnosed asthma was 4.2%. Allergic rhinitis symptoms were reported in 32.7% of the children throughout their lifetime and in 29.9% during the last year, while the rate of physician-diagnosed allergic rhinitis was 6.7%. The lifetime prevalence of atopic dermatitis was 8.8%, and the physician-diagnosed rate was 2.6%. Family history of asthma, allergic rhinitis, atopic dermatitis, and premature birth were identified as significant risk factors ($p<0.05$). No statistically significant association was found with gender, exposure to tobacco smoke, pet ownership, mode of delivery, parental consanguinity, or the presence of mold in the home.

Conclusion: These findings confirm that allergic diseases represent a significant public health burden among children in Malatya. The results specifically highlight family history and perinatal factors as major determinants of disease prevalence in this population.

Keywords: Hypersensitivity, Asthma, Prevalence, Risk factors, Population based study

Received: Sep 01, 2025 **Accepted:** Oct 15, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

Asthma and other allergic diseases are among the most common chronic conditions in childhood, with an increasing prevalence observed both in Türkiye and worldwide [1,2]. The frequency of these diseases can vary significantly between countries and even across different regions within the same country [3-7]. In addition to genetic predisposition, environmental factors are also known to play a significant role in the development of allergic diseases [8].

Understanding the prevalence of allergic diseases in a given region is of great importance for planning healthcare services, implementing preventive measures, and assessing the potential economic burden on society. With this aim, the International Study of Asthma and Allergies in Childhood (ISAAC) group was established in the early 1990s, which has led to large-scale epidemiological studies in various countries and regions. These studies employed standardized questionnaires and objective assessment tools with global validity [9].

Following ISAAC Phase I, which primarily utilized questionnaires, Phase II incorporated advanced diagnostic methods such as skin prick testing, pulmonary function testing, bronchial provocation testing, complete blood counts, and serum total IgE measurements [10]. Approximately five years later, ISAAC Phase III reassessed changes in allergic disease prevalence through renewed, questionnaire-based surveys [11].

The present study aimed to determine the prevalence of allergic diseases, including asthma, allergic rhinitis, and atopic dermatitis, among children aged 6-7 years living in the central district of Malatya, and to comprehensively evaluate the environmental and familial risk factors associated with these conditions.

MATERIALS AND METHODS

This cross-sectional study included children aged between 6 and 7 years who were residing in the city center of Malatya between December 2023 and January 2024. According to data from the Turkish Statistical Institute, approximately 24,700 children are in this age group in the Malatya city center.

The study protocol was approved by the Local Ethics Committee of İnönü University Faculty of Medicine (Approval No: 2023/4503). Written informed consent was obtained from the parents of all participating children.

Power analysis

The minimum required sample size for the study was calculated using G*Power version 3.1 software. Based on previous regional studies reporting a prevalence of respiratory allergies ranging between 15% and 25%, a 95% confidence level, a statistical power of 80%, and a 5% margin of error were assumed. According to these parameters, the minimum required number of participants was determined as 354. Ultimately, 431 children were included in the study, ensuring adequate statistical power for all primary comparisons.

To ensure representation of different socioeconomic levels, the city center was divided into four regions, and one primary school was randomly selected from each region. The Turkish-adapted version of the International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire was used for data collection. Questionnaires were distributed by teachers and completed by parents. To improve comprehension, medical terms were explained in the classroom before distribution.

A total of 500 students were reached, and 431 questionnaires (86.2%) were fully completed and included in the final analysis. The definitions of asthma, allergic rhinitis, and atopic dermatitis were based on the ISAAC criteria, incorporating both symptom-based history and physician diagnosis.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics, version 26.0 (IBM Corp., Armonk, NY, USA). Categorical variables were presented as frequencies and percentages. Depending on the test assumptions, Pearson's chi-square test or continuity-corrected chi-square test was applied for statistical comparisons. A two-tailed p-value of less than 0.05 was considered statistically significant.

RESULTS

The mean age of the 431 children included in the study was 6.22 ± 0.41 years. Of these, 232 (53.8%) were girls and 199 (46.2%) were boys. Passive exposure to tobacco smoke at home was reported in 175 children (40.6%), while 30 children (7%) had mothers who smoked during pregnancy, and 27 (6.3%) were born prematurely. Regarding the mode of delivery, 231 children (53.6%) were delivered by cesarean section and 200 (46.4%) by vaginal birth. Parental consanguinity was present in 88 cases (20.4%). Mold growth in the household was reported by 50 participants (11.6%), and 72 (16.7%) reported having pets at home. A family history of asthma was noted in 83 children (19.3%), allergy in 95 (22%), allergic rhinitis in 64 (14.8%), and atopic dermatitis in 62 (14.4%) (Table 1).

Detailed analyses of risk factors for each allergic condition are presented separately in Table 3A (asthma), Table 3B (allergic rhinitis), and Table 3C (atopic dermatitis).

Table 1. Demographic characteristics.

	n (%)
Gender (Female)	232 (53.8)
Passive Exposure to Cigarette Smoke	175 (40.6)
Maternal Smoking During Pregnancy	30 (7.0)
Preterm Birth	27 (6.3)
Mode of Delivery, Cesarean Section	231 (53.6)
Consanguineous Marriage	88 (20.4)
Presence of Mold in the Home	50 (11.6)
Keeping Pets at Home	72 (16.7)
Family History of Asthma	83 (19.3)
Family History of Allergy	95 (22.0)
Family History of Allergic Rhinitis	64 (14.8)

Descriptive data are presented as frequencies and percentages.

Wheezing episodes / Asthma and associated risk factors

Regarding symptoms associated with asthma, 20% of children reported experiencing wheezing at least once in their lifetime, and 18.6% reported having this symptom within the past 12 months. In the same period, 3.7% had more than three episodes, 13.2% experienced wheezing severe enough to disturb sleep at night, and 4.4% reported symptoms severe enough to impair speech. Exercise-induced wheezing was observed in 5.3%, while 15.8% experienced night awakenings due to cough. The prevalence of physician-diagnosed asthma was 4.2% (Table 2).

Analysis of risk factors revealed a significant association between a history of wheezing at any time and a family history of

https://doi.org/10.5455/annalsmedres.2025.08.229

Table 2. Prevalence of asthma-related symptoms.

Wheezing / Coughing Symptoms	n (%)
Wheezing at any time in life	86 (20)
Wheezing in the past 12 months	80 (18.6)
Number of wheezing attacks in the past 12 months	
1-3 times	66 (15.3)
4-12 times	15 (3.5)
≥12 times	1 (0.2)
Sleep disturbance due to wheezing in the past 12 months	
Never awakened	374 (86.8)
< 1 night/week	42 (9.7)
> 1 night/week	15 (3.5)
Difficulty speaking due to severe wheezing in the past 12 months	19 (4.4)
Wheezing during or after exercise in the past 12 months	23 (5.3)
Night waking due to coughing in the past 12 months	68 (15.8)
Physician-diagnosed asthma	18 (4.2)

Descriptive data are presented as frequencies and percentages.

asthma, allergic disease, atopic dermatitis, and allergic rhinitis ($p<0.05$). Premature birth was also found to be associated with the development of wheezing ($p=0.022$). In contrast, sex, maternal smoking during pregnancy, household smoking exposure, pet ownership, mode of delivery, parental consanguinity, and presence of household mold showed no statistically significant association with wheezing symptoms (Table 3A).

Allergic rhinitis and associated risk factors

For allergic rhinitis, lifetime symptoms were reported by 32.7% of participants, and symptoms within the past 12 months were reported by 29.9%. During the same period, conjunctivitis symptoms were reported in 13% of children, and 56.5% experienced symptoms that affected daily activities more than moderately. The prevalence of physician-diagnosed allergic rhinitis was 6.7% (Table 4).

Allergic rhinitis symptoms were significantly associated with a family history of asthma, allergy, atopic dermatitis, and allergic rhinitis ($p<0.05$). However, sex, maternal smoking during pregnancy, household smoking exposure, mode of delivery, birth timing, pet ownership, parental consanguinity, and the presence of household mold were not significantly associated (Table 3B).

Atopic dermatitis and associated risk factors

With respect to atopic dermatitis, 8.8% of children reported having itchy rashes at some point in their lifetime, and 7.9% reported experiencing them within the past 12 months. Night awakenings due to itching during this period were observed in 26.3% of cases, while the prevalence of physician-diagnosed atopic dermatitis was 2.6% (Table 5).

For itchy rashes lasting six months or more, significant associations were observed with a family history of asthma, allergy, atopic dermatitis, and allergic rhinitis ($p<0.05$). No statistically significant associations were found with other factors (Table 3C).

DISCUSSION

This study aimed to determine the prevalence of asthma, allergic rhinitis, and atopic dermatitis among primary school children aged six to seven years living in the city center of Malatya, and to evaluate potential risk factors associated with these conditions. According to our findings, the prevalence of asthma in Malatya was 18.6%, allergic rhinitis was 29.9%, and atopic dermatitis was 7.9%. Among the risk factors examined, a history of premature birth was identified as a risk factor for asthma, while a family history of asthma, allergy, atopic dermatitis, or allergic rhinitis was found to be a risk factor for all three conditions.

In our study, lifetime wheezing prevalence was 20%, wheezing within the past year was 18.6%, and physician-diagnosed asthma was 4.2%. In a previous study conducted in Malatya by Topal et al. in 2015 involving 413 children aged six to seven years, lifetime wheezing prevalence was reported as 20.3%, past-year wheezing prevalence as 12.3%, and physician-diagnosed asthma as 9% [12]. Similarly, studies conducted in different regions by Akçay et al. in Denizli in 2003 among six- to seven-year-olds [5], Çetemen et al. in Aydın in 2009 among six- to seven-year-olds [13], Gürkan et al. in Diyarbakır in 2000 among six- to fifteen-year-olds [14], Arslan et al. in Sivas in 2008 among seven- to fifteen-year-olds [15], Talay et al. in Bolu in 2006 among seven- to fourteen-year-olds [16], and Tomaç et al. in Zonguldak among six- to sixteen-year-olds [17] reported lifetime wheezing prevalence rates of 22.3%, 30.8%, 22.4%, 26.8%, 15.5%, and 15.5%, respectively. Past-year wheezing prevalence rates in these studies were 9.9%, 19.3%, 14.7%, 6.5%, 9.9%, and 11.2%, while physician-diagnosed asthma prevalence was 17.3%, 17.1%, 14.1%, 10.1%, 5.6%, and 4.9%, respectively.

When comparing our findings with those from previous

https://doi.org/10.5455/annalsmedres.2025.08.229

Table 3. Risk Factors for Asthma, Allergic Rhinitis, and Atopic Dermatitis.

(a) Risk factors associated with wheezing (asthma).

Variable	Wheezing present n (%)	p value
Sex, Male	33 (16.6)	0.328
Family history of asthma	37 (44.6)	**<0.001***
Family history of allergy	36 (37.9)	**<0.001***
Family history of atopic dermatitis	20 (32.3)	**0.005***
Family history of allergic rhinitis	27 (42.2)	**<0.001***
Maternal smoking during pregnancy	8 (26.7)	0.347
Passive smoking at home	40 (22.9)	0.058
Keeping pets at home	4 (19.4)	0.964
Mode of delivery, Cesarean section	46 (19.9)	0.438
Preterm birth	10 (37)	**0.022***
Parental consanguinity	13 (14.8)	0.384
Presence mold in the home	11 (22)	0.637

* p-values were calculated using continuity-corrected Chi-square test.

(b) Risk factors associated with allergic rhinitis.

Variable	Rhinitis symptoms present n (%)	p value
Sex, Male	56 (28.1)	0.452
Family history of asthma	48 (57.8)	**<0.001***
Family history of allergy	54 (56.8)	**<0.001#**
Family history of atopic dermatitis	35 (56.5)	**<0.001***
Family history of allergic rhinitis	38 (59.4)	**<0.001***
Maternal smoking during pregnancy	14 (46.7)	0.062
Passive smoking at home	52 (29.7)	0.935
Keeping pets at home	22 (30.6)	1
Mode of delivery, Cesarean section	77 (33.3)	0.097
Preterm birth	12 (44.4)	0.138
Parental consanguinity	23 (26.1)	0.384
Presence mold in the home	11 (22)	0.255

* p-values were calculated using continuity-corrected Chi-square test. # p-values were calculated using Pearson's Chi-square test.

(c) Risk factors associated with atopic dermatitis.

Variable	Itchy rash present n (%)	p value
Sex, Male	15 (7.5)	0.486
Family history of asthma	13 (15.7)	**0.026***
Family history of allergy	15 (15.8)	**0.012***
Family history of atopic dermatitis	13 (21)	**<0.001***
Family history of allergic rhinitis	11 (17.2)	**0.020***
Maternal smoking during pregnancy	5 (16.7)	0.169
Passive smoking at home	15 (8.6)	1.000
Keeping pets at home	7 (9.7)	0.945
Mode of delivery, Cesarean section	21 (9.1)	0.964
Preterm birth	0 (0)	0.154
Parental consanguinity	9 (10.2)	0.755
Presence mold in the home	4 (8)	1

* p-values were calculated using continuity-corrected Chi-square test.

studies conducted in various regions of Türkiye and in our province, we observed a noticeable increase in past-year wheezing prevalence, except for the study by Çetemen et al., alongside a marked decrease in physician-diagnosed asthma prevalence. This may be explained by the increased air pollution in the region following the major earthquake on February 6, 2023, which could have contributed to the higher symptom frequency. Additionally, the reduced accessibility of healthcare services in the post-earthquake period may have contributed to the decline in physician-diagnosed asthma rates.

In our study, the lifetime prevalence of allergic rhinitis symptoms was 32.7%, past-year prevalence 29.9%, and physician-diagnosed allergic rhinitis 6.7%. In the 2015 study conducted in Malatya by Topal et al., lifetime rhinitis prevalence among six- to seven-year-olds was reported as 37%, past-year rhinoconjunctivitis prevalence as 12.1%, and physician-diagnosed rhinitis prevalence as 3.4% [12]. Similarly, studies by Akçay et al. in Denizli in 2003 [5], Çetemen et al. in Aydın

https://doi.org/10.5455/annalsmedres.2025.08.229

Table 4. Prevalence of allergic rhinitis-related symptoms.

Symptom or Condition	n (%)
Allergic rhinitis (AR) symptoms at any time in life	141 (32.7)
AR symptoms in the past 12 months	129 (29.9)
Allergic conjunctivitis symptoms in the past 12 months	56 (13)
Impact of AR symptoms on child's activities in the past 12 months	
None	11 (8.5)
Very little	45 (34.8)
Moderate	69 (53.4)
Severe	4 (3.1)
Physician-diagnosed AR	29 (6.7)

Descriptive data are presented as frequencies and percentages.

Table 5. Prevalence of atopic eczema-related symptoms.

Symptom or Condition	n (%)
Atopic dermatitis (AD) symptoms at any time in life	38 (8.8)
AD symptoms in the past 12 months	34 (7.9)
Night waking due to itching in the past 12 months	
None	25 (73.5)
Less than once per week	7 (20.5)
Once or more per week	2 (5.8)
Physician-diagnosed AR	11 (2.6)

Descriptive data are presented as frequencies and percentages.

in 2009 [13], and Arslan et al. in Sivas in 2008 [15] reported lifetime allergic rhinitis prevalence rates of 33.5%, 30.4%, and 25.2%, respectively. Past-year prevalence rates were 8%, 11.1%, and 17.7%, and physician-diagnosed allergic rhinitis rates were 6.1% and 8.3% in the latter two studies.

When compared with earlier studies from our region and other parts of the country, the observed increase in past-year allergic rhinitis symptoms in our study is likely linked to earthquake-related air pollution.

With respect to atopic dermatitis, our study found lifetime symptom prevalence of 8.8%, past-year prevalence of 7.9%, and physician-diagnosed prevalence of 2.6%. In the study by Topal et al. involving six- to seven-year-olds, these rates were 7.5%, 6.5%, and 7.3%, respectively [12]. Similarly, studies by Akçay et al. in Denizli in 2003 [5], Çetemen et al. in Aydın in 2009 [13], and Arslan et al. in Sivas in 2008 [15] reported lifetime atopic dermatitis prevalence rates of 11.3%, 9.6%, and 28.3%, respectively; past-year prevalence rates were 8.2%, 7.8%, and 20.5%; respectively and physician-diagnosed prevalence rates were 2.8% and 2.9% in the latter two studies.

The lower physician-diagnosed atopic dermatitis prevalence in our study compared with others may be partly due to reduced access to healthcare services following the February 2023 earthquake. Additionally, variations in genetic predisposition, lifestyle, dietary habits, and the methodology of questionnaire administration may also have influenced prevalence rates.

Regarding risk factors, a significant association was found between premature birth and wheezing, supporting the hypothesis that incomplete lung development in preterm infants increases susceptibility to respiratory diseases [18]. Furthermore, children with a family history of asthma, allergy, atopic dermatitis, or allergic rhinitis were significantly more likely to experience symptoms of all allergic diseases, reinforcing the role of genetic predisposition in their development, as supported by previous literature [19].

In contrast, no significant associations were observed between asthma or other allergic disease symptoms and environmental risk factors such as maternal smoking during pregnancy, exposure to household tobacco smoke, pet ownership, or the presence of household mold.

Limitations

Limitations of this study include the use of a questionnaire-based data collection method, which may be subject to recall bias, parental misreporting, and social desirability bias. In addition, the study relied exclusively on questionnaire-based data without objective clinical measurements such as spirometry, skin prick testing, or serum IgE levels. Future studies incorporating such quantitative assessments would provide a more comprehensive understanding of allergic disease prevalence and risk factors. Additionally, the potential indirect effects of post-earthquake changes in living conditions on the results should be considered.

CONCLUSION

In conclusion, allergic diseases appear to be an increasingly significant public health concern in the Malatya region. While

https://doi.org/10.5455/annalsmedres.2025.08.229

nonmodifiable risk factors such as family history and birth conditions play an important role, targeted interventions addressing environmental factors remain essential for regional healthcare planning.

Ethics Committee Approval: The study protocol was approved by the Local Ethics Committee of İnönü University Faculty of Medicine (Approval No: 2023/4503).

Informed Consent: Informed consent was obtained from the parents of all children who participated in the study.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors declare that they have no conflict of interest.

Author Contributions: M.Ç.Ş.: Study conception and design, data collection, statistical analysis, manuscript drafting. E.Y.: Literature review, data interpretation, manuscript drafting. E.T.: Study design, supervision, critical revision of the manuscript.

Financial Disclosure: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Artificial Intelligence Disclosure: The authors declare that an artificial intelligence–based tool (ChatGPT, OpenAI, GPT-4) was used solely for language editing and improving the clarity and readability of the manuscript. The AI tool did not contribute to the study design, data collection, statistical analysis, interpretation of results, or the formulation of scientific conclusions. All content was critically reviewed and approved by the authors, who take full responsibility for the accuracy, originality, and scientific integrity of the manuscript.

REFERENCES

1. Masoli M, Fabian D, Holt S, Beasley R. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy.* 2004;59(5):469–478. doi: 10.1111/j.1398-9995.2004.00526.x.
2. Duksal F, Tamay Z, Guler N, et al. The prevalence of asthma diagnosis and symptoms is still increasing in early adolescents in Turkey. *Allergol Int.* 2014;63(2):189–197. doi: 10.2332/allergolint.13-OA-0612.
3. ISAAC Steering Committee. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. *Lancet.* 1998;351(9111):1225–1232. PMID: 9643741.
4. Ece A, Gürkan F, Haspolat K, et al. Prevalence of asthma and other allergic disorders among schoolchildren in Diyarbakir, Turkey. *Turk J Pediatr.* 2001;43(4):286–292. PMID: 11765156.
5. Akçay A, Bavbek N, Guler N, Altınel N. Prevalence of allergic diseases in 6–7-year-old schoolchildren in Denizli. *Ege Med J.* 2007;46(2):87–91.
6. Kalyoncu AF, Selçuk ZT, Enünlü T, et al. Prevalence of childhood asthma and allergic diseases in Ankara, Turkey. *Allergy.* 1994;49(6):485–488. doi: 10.1111/j.1398-9995.1994.tb00846.x.
7. Tamay Z, Akçay A, Ferik F, Guler N. Dietary habits and prevalence of allergic rhinitis in 6 to 7-year-schoolchildren in Turkey. *Allergol Int.* 2014;63(4):553–562. doi: 10.2332/allergolint.13-OA-0661.
8. Farid-Hossaini R, Ghaffari J, Ranjbar A, Haghshenas MR, Rafatpanah H. Infections in children with asthma. *J Pediatr Rev.* 2013;1(1):34-45.
9. Asher MI, Keil U, Anderson HR, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J.* 1995;8(3):483–491. doi: 10.1183/09031936.95.08030483.
10. Weiland SK, Björkstén B, Brunekreef B, et al. Phase II of the International Study of Asthma and Allergies in Childhood (ISAAC II): rationale and methods. *Eur Respir J.* 2004;24(3):406–412. doi: 10.1183/09031936.04.00090303.
11. Ellwood P, Asher MI, Beasley R, et al. The International Study of Asthma and Allergies in Childhood (ISAAC): phase three rationale and methods. *Int J Tuberc Lung Dis.* 2005;9(1):10–16. PMID: 15675544.
12. Topal E, Kaplan F, Türker K, Kutlutürk K, Bağ HG. The prevalence of allergic diseases and related risk factors in 6–7 age children living in Malatya, Turkey. *Asthma Allergy Immunol.* 2017;15(3):129–134. doi: 10.21911/aai.348.
13. Çetemen A, Yenigün A. Prevalences of asthma and allergic diseases in primary school children in Aydın. *Asthma Allergy Immunol.* 2022;10(2):84–92.
14. Gürkan F, Ece A, Haspolat K, et al. Asthmatic children and risk factors in a province in southeast Turkey. *Allergol Immunopathol (Madr).* 2002;30(1):25–29. doi: 10.1016/s0301-0546(02)79083-6.
15. Arslan S, Uğurlu S, Demirel Y, et al. Sivas Yöresinde İlköğretim Çocuklarında Astım Ve Alerjik Hastalıkların İlişkisi Ve Prevalansı. Nobel Med. 2012;8(1):30–34.
16. Talay F, Kurt B, Tug T, el al. Prevalence and risk factors of asthma and allergic diseases among schoolchildren in Bolu, Turkey. *Acta Paediatr.* 2008;97(4):459-62. doi: 10.1111/j.1651-2227.2008.00726.x.
17. Tomac N, Demirel F, Acun C, et al. Prevalence and risk factors for childhood asthma in Zonguldak, Turkey. *Allergy Asthma Proc.* 2005;26(5):397-402. PMID: 16450575.
18. Trønnes H, Wilcox AJ, Markestad T, et al. The association of preterm birth with severe asthma and atopic dermatitis: a national cohort study. *Pediatr Allergy Immunol.* 2013;24(8):782–787. doi: 10.1111/pai.12170.
19. Holberg CJ, Elston RC, Halonen M, et al. Segregation analysis of physician diagnosed asthma in Hispanic and non Hispanic white families: a recessive component? *Am J Respir Crit Care Med.* 1996;154(1):144–150. doi: 10.1164/ajrccm.154.1.8680670.

https://doi.org/10.5455/annalsmedres.2025.08.229

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Eating behaviors in individuals over the age of 18 years who applied to a family health center and the associated factors

Erhan Atici [a,*], Aliye Bulut [b]

[a]Elazığ Provincial Health Directorate, Elazığ, Türkiye
[b]Gaziantep Islamic University of Science and Technology, Faculty of Medicine, Department of Public Health, Gaziantep, Türkiye
*Corresponding author: erhanatici8123@gmail.com (Erhan Atici)

MAIN POINTS

- The present study revealed elevated levels of emotional eating among participants, whereas other eating behaviors were moderate.
- The EBQ scores were notably higher in older adults, married individuals, those without children, high-income groups, and participants with chronic illnesses or physical disabilities.
- Significant positive correlations were identified between AEBQ scores and factors such as age, number of children, duration of illness and disability, and BMI.
- Furthermore, both the illiterate and highly educated groups exhibited significant variations in the AEBQ scores.
- These results underscore the multifaceted influence of demographic variables on eating behaviors among adults.

Cite this article as: Atici E, Bulut A. Eating behaviors in individuals over the age of 18 years who applied to a family health center and the associated factors. *Ann Med Res*. 2026;33(1):19--26. doi: 10.5455/annalsmedres.2025.05.135.

ABSTRACT

Aim: This study aimed to determine the eating behaviors of individuals over the age of 18 years who applied to a family health center as well as the associated factors.

Materials and Methods: This cross-sectional study included 1,684 adults aged 18 years who attended two Family Health Centers in Elazığ, Turkey. Participants voluntarily completed a questionnaire, including sociodemographic items and the AEBQ, through face-to-face interviews. Ethical approval was obtained from the Ethics Committee of the Non-Interventional Clinical Trials. Statistical analyses, including t-tests, analysis of variance, Pearson correlation, and multiple linear regression, were used to examine associations between eating behaviors and sociodemographic and health-related factors.

Results: The participants' emotional eating (EE) scores were high, whereas the other AEBQ subscale and total scores were moderate. Higher AEBQ scores were observed in individuals aged ≥55 years, married, childless, with higher income, with chronic diseases, or with physical disabilities. Positive correlations were found between age, number of children, chronic disease duration, physical disability duration, BMI, and AEBQ scores.

Conclusion: The study demonstrates significant associations between sociodemographic and health-related factors and EEBs. These findings support the need for further research using the AEBQ and may guide interventions to promote healthier eating habits and improve the health of the population.

Keywords: Eating behaviors, Eating behavior factors, Family health center

Received: May 27, 2025 **Accepted:** Oct 23, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

As one of the most basic human needs, nutrition is important for protecting and promoting health and enhancing the quality of life [1]. Individuals' nutritional needs vary according to age, gender, physical activity intensity, disease condition, and genetic structure [2].

An energy deficit emerges if the daily energy intake is less than the total energy consumed. The lower daily energy intake may be due to reasons such as consciously or unconsciously unbalanced and inadequate nutrition, abnormal eating behaviors, and the inability to meet the energy requirement due to energy intake restriction [3].

Behaviors related to eating habits are established over time and are affected by both psychosocial and social factors [4]. Many factors affect food choice, including age, habits, educational level, accessibility, environment and ethics, body weight control, perceived stress level, genetics, nutritional knowledge, physical activity level, and concern about good health [5]. These factors, which we refer to as personal characteristics and environmental factors, may cause eating behavior disturbances [6].

The gradual increase in obesity all over the world is an important issue for us to question the eating habits of adults [7]. Obesity is a health problem that negatively affects psycholog-

ical and physical health and leads to an elevated mortality rate [8]. Overeating and reduced physical activity are causes of weight gain [9].

Eating behavior may also show different trends when affected by individuals' sociopsychological state. These include eating behavior disorders such as emotional eating, external eating, and restrictive eating [10]. Emotional Eating can be described as a response to emotions such as joy, stress, and sadness by referring to food.External Eating refers to eating in response to external stimuli such as the smell and presentation of food, even though one is not hungry. Restrictive eating is defined as excessive restriction of food intake due to anxiety about weight gain [11].

This study aimed to determine the eating behaviors of individuals over the age of 18, identify the factors affecting eating behaviors, examine these factors, and make recommendations for regulating the eating behaviors of individuals.

MATERIALS AND METHODS

Data analysis

This study was conducted using a cross-sectional method with quantitative data. Quantitative research is conducted within the positivist paradigm framework and emphasizes the use of probabilistic sampling methods in large populations and samples [12]. This study employed an observational model, a type of quantitative method, and collected data through face-to-face interviews. The researchers informed the participants through the informed consent form attached to the questionnaire, which included the principles of the Declaration of Helsinki. The inclusion criteria of the study were voluntary participation and registration in the FHC. The exclusion criteria were withdrawal from the study during or after the completion of the questionnaire. The study was completed with 1684 participants.

The research was conducted between May 31 and August 31, 2023, at Family Health Units No. 91 and No. 105 in the city center of Elâzığ. During the specified dates, 1292 and 958 patients presented to Family Health Unit No. 91 and 958 to Family Health Unit No. 105, making a total of 2250 applications. The study included all individuals aged 18 and over who volunteered to participate, and the research was completed with a total of 1,684 individuals. This number corresponds to approximately 75% of adult applicants during the study period. In descriptive cross-sectional studies, when a very large portion of the universe is reached, calculating a minimum sample size is unnecessary [13].

The questionnaire form used in the study consists of two sections: The first part of the form contains the Individual Information Form, which includes questions about sociodemographic characteristics such as age, gender, marital status, educational level, health conditions such as diseases and allergies, and questions that investigate the attitudes and behaviors of the participant toward eating habits. The second part of the form is the "Adult Eating Behavior Questionnaire (AEBQ)."

The dependent variable of the study is the total score and subscale scores of the AEBQ. The independent variables were sociodemographic and health-related characteristics, such as age, gender, marital status, education level, parental status, occupation, perceived economic status, difficulty in accessing the health center, presence of a chronic disease, presence of a physical disability, height, weight, body mass index, number of main meals/snacks, dietary habits, and the number of healthcare service applications. The main hypothesis of the research is that these independent variables have significant effects on the AEBQ scores.

The original form of the AEBQ developed by Hunot et al. [14] in 2016 consists of 35 questions. Participants are expected to self-assess using a 5-point Likert-type response ranging from "Strongly agree" to "Strongly disagree." Accordingly, the minimum and maximum possible scores are 35 and 175, respectively. Higher scores indicate that individuals exhibit specific eating behaviors more prominently. Items in the AEBQ are classified on eight scales, four of which are food approach and four food avoidance scales. The scale has eight subscales:Enjoyment of Eating (EE), Emotional Over-eating (EO), Emotional Under-eating (EU), Food Fussiness (FF), Food Responsiveness (FR), Slowness in Eating (SE), Hunger (H), and Satiety Responsiveness (SR). The validity and reliability study conducted by Şengül et al. [15] reported Cronbach's alpha values as EE = 0.70, EO = 0.71, EU = 0.86, FF = 0.80, FR = 0.81, SE = 0.93, H = 0.77, and SR = 0.70 [15]. However, the Cronbach's alpha values in this study were found to be EE=0.78,EO=0.97,EU=0.97,FF=0.73,FR=0.89,SE=0.65, H=0.86,and SR=0.81. There is no cutoff value for the scale. Scale evaluation is based on eight subheadings and a total average score. In the validity and reliability study conducted by Yardımcı et al. in a similar manner, the descriptive statistics (mean ± SD), internal consistency (Cronbach's alpha), and test–retest reliability of the AEBQ-TR were also reported. Internal consistency demonstrated that the questionnaire had adequate reliability, with Cronbach's alpha coefficients exceeding 0.70 for all subscales except satiety responsiveness. Test–retest reliability was also above 0.70 (ranging between 0.95 and 0.98) for all subscales of the AEBQ-TR [16].

The Non-Interventional Clinical Trials Ethics Committee of the Gaziantep Islam Science and Technology University granted written approval before the study (approval no: 2023/214).

Statistical analysis

Analyses were conducted using a statistical software package and subsequently verified using R, an open-source and free software. Categorical data are expressed in numbers and percentages. For numerical data,a normality analysis was performed. To assess the applicability of parametric tests, the assumption of normality was examined using the Kolmogorov-Smirnov test, and the skewness (0.372) and kurtosis (1.032)

https://doi.org/10.5455/annalsmedres.2025.05.135

values were found to be within acceptable limits. The Levene test was used to evaluate the homogeneity of variances, and the assumption of homogeneity was met. Therefore, following the one-way analysis of variance, the Tukey HSD post hoc test was used to determine the source of the difference between the groups. Pearson correlation analysis, parametric analyses (independent sample t-test and one-way analysis of variance test), and multiple linear regression analysis were performed. The independent variables to be included in the analysis were selected from variables found to be significant in the univariate analyses in our study and from previous literature findings. Additionally, a check for multicollinearity was performed as the variables were entered into the model, and the VIF values were found to be <10. A statistical significance level of $p<0.05$ was set.

RESULTS

The average age of the participants was 39.45±15.58 (min: 18, max: 84 years), and 44.7% were male. Table 1 shows the descriptive characteristics of the participants.

This study revealed that 6.1% of the participants reported their marital status as "other." Among respondents, 33.8% identified themselves as single, while 60.1% reported being married. The mean number of children was 2.64±1.40 (min: 1, max: 8). 42.3% of the participants indicated their profession as "other," while 281 (39.1%) of them stated that they were employed in the private sector, 100 (13.9%) of them stated that they were retired, 82 (11.4%) stated that they were teachers; 72 (10.0%) stated that they were health workers; 66 (9.2%) stated that they were students; 53 (7.4%) stated that they were academicians, 23 (3.2%) stated that they were engineers, 23 (3.2%) stated that they were business owners, 13 (1.8%) stated that they were police officers, and 5 (0.7%) stated that they were pharmacists. The rate of those who reported having difficulty in accessing health services was 71.7%, 671 people (55.2%) stated that they could not make an appointment, 295 people (24.3%) stated that the number of queues at the hospital was too long, 142 people (11.7%) stated that transportation conditions were difficult, 100 people (8.2%) stated financial constraints, and 8 people (0.7%) stated physical disability as one of the reasons why they could not access services. The rate of participants with chronic diseases was 26.2%, and the mean duration of the disease was 10.93±6.67 (min: 2, max: 40 years). The mean number of family physician visits in the last year was 2.69±1.52 (min: 1, max: 10 times), while the mean number of hospitalizations in the last year was 3.02±2.11 (min: 0, max: 15 times).

The mean height of the participants was 168.36±9.53 cm (min: 150, max: 192 cm), their mean weight was 71.08±9.78 (min: 48, max: 99 kg), and their mean body mass index was 25.13±3.76 (min: 17.65, max: 39.55).

The mean number of main meals consumed by the participants per day was 2.32±0.50 (min: 1, max: 4), and the mean number of snacks was 1.02±0.82 (min: 0, max: 3). Almost half of the participants skipped the main meal. Table 2 shows the nutritional characteristics of the participants.

As shown in Table 3, the participants' mean EE score was high. The mean scores of the other subscales and the AEBQ mean score were medium.

In the evaluation conducted to identify the general eating behaviors of adults, the mean total AEBQ score among participants aged 18 years and older was found to be 101.95 ± 11.66. Among the subscales, the highest mean score was observed in the EE subscale (10.81 ± 2.08), while the lowest mean score was observed in the FR subscale (7.77 ± 2.91). These findings provide a descriptive overview of the study population's general eating behaviors (Table 3).

As shown in Table 4, the age range of the participants caused a difference in the distribution of AEBQ scores. The mean scores of those who were 55 years old and older were higher than those in the other two age ranges. This was highly significant ($p<0.001$). The variable of sex did not cause a difference in the distribution of AEBQ scores. At the educational level, it was found that both illiterates and those who held a higher degree of education were statistically highly significant groups causing the difference ($p<0.001$). Similarly, those who had no children ($p<0.001$), those who earned more income ($p<0.001$), those who suffered from chronic diseases ($p<0.001$), and those who were physically disabled had higher AEBQ scores.

Table 5 shows the correlation between the characteristics of the participants and AEBQ. There was a statistically significant positive correlation between the participant's age, number of children, duration of chronic disease, duration of physical disability, and BMI values and their AEBQ scores, but these variables were very weakly correlated with a one-unit change in the AEBQ.

In this study, the model created to establish the variables affecting the AEBQ score was found to be significant ($p<0.05$). The AEBQ score was predicted in terms of the variables of educational level, perception of income level, and having a physical disability. The level of education explained 1.14% of the 1 (one) standard deviation change in the AEBQ score, 0.49% by the perception of income level, and 2.07% by the presence of a physical disability. explained (Table 6).

Age, educational level, number of children, perceived income level, presence of chronic disease, and physical disability were identified as factors influencing eating behaviors. These variables were significantly associated with AEBQ scores, thereby indicating the factors affecting eating behaviors.

DISCUSSION

In this study, which was conducted to determine the eating behaviors of individuals over the age of 18 who applied to the family health center, the mean age of the participants was 39.45±15.58 (min: 18, max: 84 years), and 44.7% were male. In a similar study by Şengül, it was reported that the mean age of the participants was 19.36±1.70 [15].

https://doi.org/10.5455/annalsmedres.2025.05.135

Table 1. Some descriptive characteristics of the participants (N= 1684).

Characteristics		n	%
Age range	35 years and under	820	48.7
	Between 36-55 years	591	35.1
	55 years and over	273	16.2
Gender	Male	752	44.7
	Female	932	55.3
Marital Status	Married	1012	60.1
	Single	570	33.8
	Other (Widow/Divorced/Estranged)	102	6.1
Educational Level	Illiterate	38	2.3
	Literate	31	1.8
	Primary school	99	5.9
	Secondary school	135	8.0
	High school	601	35.7
	University	780	46.3
Having a child	Yes	652	38.7
	No	1032	61.3
Profession n (1679)	Civil servant	302	18.0
	Worker	139	8.3
	Farmer	23	1.4
	Self-employed	115	6.8
	Housewife	390	23.2
	Other	710	42.3
Perceptions of income status	Income more than expenses	334	19.8
	Income equal to expenses	697	41.4
	Income less than expenses	653	38.8
The status of having difficulty in accessing the health centre	Never	476	28.3
	Rarely	463	27.5
	Sometimes	523	31.1
	Usually	161	9.6
	Always	61	3.5
The status of having chronic disease	No	1243	73.8
	Yes	441	26.2
The status of having a physical disability	No	1671	99.2
	Yes	13	0.8

Table 2. Some nutritional characteristics of the participants (N= 1684).

Characteristics		n	%
Skipping main meals	No	904	53.7
	Yes	144	8.6
	Sometimes	628	37.3
	Often	8	0.5
Which meal do they skip the most?	Breakfast	337	20.0
	Lunch	409	24.3
	Dinner	49	2.9
	Don't skip any at all	889	52.8
Diet	Vegetarian	5	0.3
	Vegan	88	5.2
	Pescatarian	48	2.9
	Omnivorous	1543	91.6
Having a food allergy	Yes	41	2.4
	No	1643	97.6

In the present study, 60.1% of the individuals were married, 26.2% had a chronic disease, and 53.7% did not skip the main meal. In their study, Arslantaş et al. reported that 1.7% of the participants were married, 14.9% had a chronic disease, and 75.1% had an eating habit of 3 or more meals per day [17]. While the present study shows similar characteristics to

https://doi.org/10.5455/annalsmedres.2025.05.135

Table 3. Scores of AEBQ and its subscales used in the study.

	Mean ± SD	Median	Min.	Max.	95% CI
EE	10.81±2.08	11.00	3.00	15.00	10.71-10.90
EO	8.25±3.42	9.00	3.00	15.00	8.08-8.41
EU	12.68±4.38	12.00	4.00	20.00	12.47-12.75
FF	13.35±4.12	14.00	5.00	25.00	13.16-13.55
FR	7.77±2.91	7.50	3.00	15.00	7.63-7.91
SE	12.01±4.64	12.00	4.00	52.00	11.79-12.24
H	12.73±4.66	12.00	5.00	25.00	12.50-12.95
SR	12.15±3.61	12.00	4.00	20.00	11.98-12.33
AEBQ	101.95±11.66	102.00	73.00	145.00	101.39-102.50

REMARKS: EE; Enjoyment of Eating, EO; Emotional Over-eating, EU; Emotional Under-eating, FF; Food Fussiness, FR; Food Responsiveness, SE; Slowness in Eating, H; Hunger, SR; Satiety Responsiveness, AEBQ; Adult Eating Behaviour Questionnaire.

Table 4. The variables causing a difference on the distribution of AEBQ mean scores of the participants (N= 1684).

Characteristics	Variable	n	AEBQ Mean±SD	Test values	Effect Size η^2
Age range	35 years and under	820	101.45±12.41[a]	F= 21.11	0.012-0.040
	Between 36-55 years	591	100.74±11.44[b]	**p= <0.001**	
	55 years and over	273	106.03±8.52[a,b]		
Gender	Male	752	102.26±12.08	t= 0.990	
	Female	932	101.69±11.30	p= 0.322	
Marital status	Married	1012	102.78±11.26[a,b]	F= 18.603	0.010-0.037
	Single	570	99.76±12.06[a,c]	**p= <0.001**	
	Other	102	105.82±11.23[b,c]		
Educational level	Illiterate	38	111.55±10.65[a]	F= 17.621	0.030-0.069
	Literate	31	110.16±5.17[b]	**p= <0.001**	
	Primary school	99	107.53±6.51[c]		
	Secondary school	135	98.37±10.40[a,b,c]		
	High school	601	102.23±12.54[a,b,c]		
	University	780	100.84±11.29[a,b,c]		
Having a child	Having a child	652	100.46±12.18	t= -4.341	-0.315-0.119
	Having no child	1032	102.92±11.21	**p= <0.001**	
Profession n (1679)	Civil servant	302	104.46±13.46[a]	F= 30.234	0.057-0.106
	Worker	139	107.26±9.89[b,e]	**p= <0.001**	
	Farmer	23	106.60±8.92[c]		
	Self-employed	115	103.56±11.13[d]		
	Housewife	390	103.92±10.00[e,f]		
	Other	710	98.12±10.78[a,b,c,d,f]		
Income status?	Income more than expenses	334	103.22±12.61[a]	F= 4.373	0.000-0.013
	Income equal to expenses	697	101.04±11.96[a]	**p= 0.013**	
	Income less than expenses	653	102.27±11.66		
Having difficulty in accessing the health centre?	Never	476	98.70±11.23[a]	F= 39.544	0.061-0.111
	Rarely	463	100.99±10.77[a,b]	**p= <0.001**	
	Sometimes	523	102.39±12.38[a]		
	Usually	161	109.13±7.84[a,b]		
	Always	61	111.73±9.58[a,b]		
Chronic disease?	No	1243	101.16±12.25	t= -4.678	-0.368-0.150
	Yes	441	104.16±9.44	**p= <0.001**	
Physical disability?	No	1671	101.77±11.52	t= -7.289	-2.579-1.479
	Yes	13	125.07±2.53	**p= <0.001**	

REMARKS: [a,b,c,d,e,f] show the groups causing the difference and Tukey test was run. SD: Standard Deviation, AEBQ: Adult Eating Behaviour Questionnaire.

Arslantaş's study in some aspects, it also shows differences in some aspects; this difference is due to the population in which the studies have been conducted and the different population groups.

In the present study, the AEBQ mean score was 101.95±11.66 (min: 73, max: 145 points), which was at aa moderate level. The highest score was 13.35±4.12, obtained in the FF subscale (95% CI: 13.16-13.55) and the lowest score was 7.77±2.91, obtained in the FR subscale (95% CI: 7.63-7.91). In a similar study conducted by Yardımcı et al., it was reported that the

https://doi.org/10.5455/annalsmedres.2025.05.135

Table 5. The correlation of some characteristics of the participants and AEBQ (N= 1684).

		Age	Number of Children	Duration of chronic disease	Duration of physical disability	BMI
AEBQ	r	0.113**	0.111**	0.159	1.000**	0.263**
	p	**<0.001**	<0.001	**<0.001**	**0.001**	**0.001**

**Significant level at the level of 0.01, * at the level of 0.05.

Table 6. Some of effective characteristics on the AEBQ score.

	Stan.Coef. β	t	p	Subscale/Component	95% CI	VIF
Constant	91.243	19.102	<0.001		81.87-100.61	
Age	-0.197	-0.364	0.716	-0.009	-1.261-0.866	2.09
Marital status	-0.791	-1.461	0.144	-0.036	-1.852-0.271	1.41
Educational level	-1.420	-4.419	**<0.001**	-0.107	-2.050-0.790	1.74
Having a child	0.486	0.605	0.545	0.015	-1.090-2.063	2.00
Income	-1.336	-3.140	**0.002**	-0.076	-2.171-0.502	1.30
Chronic disease	1.365	1.570	0.117	0.038	-0.341-3.071	1.91
Physical disability	19.729	5.969	**<0.001**	0.144	13.246-26.212	1.09

AEBQ mean score was at a moderate level, the highest score was observed in the EE subscale, and the lowest score was observed in the EO subscale [16]. Mostafazadeh et al reported the highest score in the EE subscale and the lowest score in the EO subscale [18]. The results of the present study are similar to the study byYardımcı et al., in terms of the AEBQ score, and the differences in scores in the subscales are thought to be caused by the different cultural structures and eating habits of the people of the different countries where the study was conducted.

In the present study, the AEBQ score of individuals aged 55 and over was significantly higher than that of individuals aged between 36 and 55 years and those aged 35 and under. In their study, Şengür et al found that the AEBQ score increased with increasing age. The results of this study are similar to those reported in the literature [19]. The AEBQ score is expected to increase with increasing age, as lifestyle and cultural changes acquired during adulthood are important factors on food preferences and nutrition.

In this study, no statistically significant difference was found in the AEBQ score according to gender. In their study, Yurt and Özdemir reported no significant difference [20]. Zulu et al. also found that although there was no significant difference between eating behaviors and gender, the scores of women were higher than those of men [21]. Kahraman et al found that the eating behavior score was statistically significantly higher in female students than in male students [22]. The difference between the present study and other studies supporting the present study and Kahraman's study is because Kahraman's study was conducted only among the young student population.

In the study, it was determined that the AEBQ scores of secondary school, high school, and university graduates were significantly higher than those of illiterate, literate, and primary school graduates. Mutlu et al observed that there was no statistically significant difference in terms of adult eating behaviors in terms of education [21]. This difference is thought to be due to the increase in the level of education and the fact that individuals' body perception and eating behaviors include restrictive and regulatory eating behaviors.

When the AEBQ score was evaluated from an economic perspective, the AEBQ score of those who stated that their income was more than expenses was found to be significantly higher than those who stated that their income was equal to and less than expenses. In the study by Tanriverdi et al., the eating attitude behavior scores of those who described their family's income status as very high were significantly higher than those who described their family's income status as middle or low [23]. It is expected that a high income level and the perception of having a high income level will increase the AEBQ score. The results of the present study are compatible with the literature results.

In the present study, the AEBQ score of married people was higher than that of single people, and the AEBQ score of other marital statuses (divorced and widowed) was higher than that of single and married people. Ceyhan's study found no significant correlation between marital status and eating behaviors [24]. This difference between the results of the present study and Ceyhan's study is thought to be due to the fact that Ceyhan's study was conducted among students and the number of married individuals was very low compared to single individuals. Single individuals are also thought to be at a higher risk of eating disorders, and due to their irregular lifestyle, they cannot pay much attention to their eating habits and have worse eating attitudes than married individuals.

The AEBQ score of those without children was higher than those who had children. This is one of the most important reasons why parents with children have a negative impact on their eating behaviors, as they have the concerns and worries of all family members as a whole due to the concept of family, in addition to their own individual responsibilities.

https://doi.org/10.5455/annalsmedres.2025.05.135

The AEBQ score of those with chronic diseases was higher than those without chronic diseases. In Ayaz's study, individuals with chronic diseases had higher eating attitude and behavior scores than those without chronic diseases [25]. Individuals with chronic diseases show more care and attention to their eating behavior because of the need to prevent further disease progression and to live a better quality of life. Therefore, it is important for individuals with chronic diseases to establish a more regular and careful eating behavior discipline for their own health stability.

The AEBQ score of those with physical disabilities was higher than those without. Disabled individuals behave in a more multidisciplinary manner on this issue because they believe that regular consumption of food and beverages, and therefore their eating attitudes and behaviors, are an indicator of being fully healthy.

When the correlation between some characteristics of the participants and AEBQ scores was examined, a positive correlation was observed between AEBQ scores and age, chronic disease duration, physical disability duration, and BMI. The AEBQ score is expected to increase as the BMI increases, as obesity is not only a problem in developed countries but also one of the most important public health problems in developing countries due to globalization. Similar results were obtained by Shinde et al. [26].

CONCLUSION

In this study, the participants' EE mean scores were high, and the mean scores of the other subscales and AEBQ were moderate. However, the AEBQ score was higher in individuals aged 55 and over than in the other age groups, married individuals compared to single individuals, and individuals without children than those with children. Those who had a high income, suffered from chronic diseases, and had physical disabilities also had high AEBQ scores. A statistically significant positive correlation was found between the participants' age, number of children, duration of chronic disease, duration of physical disability, and BMI values and AEBQ scores. The illiterate and those with advanced education were the groups that caused the difference. As the scale used in this study is new in the literature and has not been widely applied, increasing the number of studies using the AEBQ for the assessment of adult eating behaviors and expanding solution proposals based on the findings of these studies are recommended. In line with the first objective of the study, the eating behaviors of individuals over the age of 18 were examined, with participants showing high mean scores in EE and moderate scores in the other subscales. These findings provide an overview of the participants' general eating behavior profile and are expected to contribute to the development of healthier eating habits and healthier societies.

Ethics Committee Approval: Approval of the Ethics Committee of the Gaziantep Islam Science and Technology University (protocol 2023/214, approval date: 02.05.2023) was taken for the study.

Informed Consent: Informed consent was obtained from all the subjects participating in the study.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors declare no conflict of interest.

Author Contributions: Conception, E.A.; Design E.A.; Supervision, E.A.; and A.B; Fundings, E.A.; and A.B; Materials; E.A.; and A.B; Data Collection and/or Processing, E.A.; and A.B; Analysis and/or İnterpretation, E.A.; and A.B.; Literature Review, E.A; and A.B.; Writing, E.A; and A.B.; Critical Review, E.A; and A.B. All authors read and agreed to the published version of the manuscript.

Financial Disclosure: This research received no specific grant from any funding agency, commercial or not-forprofit sectors.

Artificial Intelligence Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

REFERENCES

1. Hacettepe University Department of Nutrition and Dietetics. Nutrition Guide Specific to Turkey 2004. Available at www.bdb.hacettepe.edu.tr/torehberi.pdf. Accessed July 21, 2023.
2. Baysal A. Impact of social inequalities on nutrition. *CMJ.* 2003;25(4):66-72.
3. Mountjoy ML, Burke LM, Stellingwerff T, Sundgot-Borgen J. Relative energy deficiency in sport: the tip of an iceberg. *Int J Sport Nutr Exerc Metab.* 2018;28(4):313-315. doi:10.1123/ijsnem.2018-0149.
4. Ganasegeran K, Al-Dubai SA, Qureshi AM, Al-abed AA, Am R, Aljunid SM. Social and psychological factors affecting eating habits among university students in a Malaysian medical school: a cross-sectional study. *Nutr J.* 2012;11:48. doi: 10.1186/1475-2891-11-48.
5. Phan UT, Chambers E. Motivations for choosing various food groups based on individual foods. *Appetite.* 2016;105:204-211. doi: 10.1016/j.appet.2016.05.031.
6. Altınok M. The evaluation of eating attitudes, emotional eating, mindful eating and orthorexia with the metacognitions. Master's thesis, University of Fatih Sultan Mehmet Department of Psychology, Istanbul 2020.
7. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 2014;6736(14):1-16. doi: 10.1016/S0140-6736(14)60460-8.
8. Flegal KM, Kit BK, Orpana H. Association of all-cause mortality with overweight and obesity using standard Body Mass index categories. A systematic review and meta-analysis. *JAMA.* 2013;309(1):71-82. doi: 10.1001/jama.2012.113905.
9. Llewellyn CH, Wardle J. Behavioral susceptibility to obesity: gene-environment interplay in the development of weight. *Physiol Behav.* 2015;152(Pt B):494-501. doi: 10.1016/j.physbeh.2015.07.006.
10. Johnson F, Pratt M, Wardle J. Dietary restraint and self-regulation in eating behaviour. *Int J Obes.* 2012;36(5):665–674. doi: 10.1038/ijo.2011.156.
11. Bayrak E, Serter R, Isikli S. The effect of emotional, external and restraint eating disorders on the development of metabolic syndrome. *Med J Ankara Train Res Hospital.* 2020;53(1):36-43. doi: 10.20492/aeahtd.669570.

https://doi.org/10.5455/annalsmedres.2025.05.135

12. Garip S. A Theoretical Study on the Quantitative Research Tradition in Social Sciences. *IJSSR.* 2023;12(1):1-19.
13. Aktürk, Z., Acemoğlu, H., Akçay, F., & Işık, M. Sağlık çalışanları için araştırma ve pratik istatistik: Örnek problemler ve SPSS çözümleri. Ankara: Akademisyen Kitabevi. 2016.
14. Hunot C, Fildes A, Croker H, Llewellyn C, Wardle J, Beeken R. Appetitive Traits and relationships with BMI in adults: Development of the adult eating behaviour questionnaire. *Appetite.* 2016;105:356-363. doi: 10.1016/j.appet.2016.05.024.
15. Sengül H. University students eating behavior scale: validity and reliability study. *Ejons J.* 2020;4(15):708-721. doi: 10.38063/ejons.316.
16. Yardimci H, Ersoy N, Aslan Cin NN. Validity and reliability of Turkish version of the Adult Eating Behaviour Questionnaire for adult participants. *Public Health Nutrition.* 2023;26(11):2218–2225. doi: 10.1017/S1368980023001751.
17. Arslantas H, Adana F, Ogut S, Ayakdas D, Korkmaz A. Relationship Between Eating Behaviors of Nursing Students and Orthorexia Nervosa (Obsession with Healthy Eating): A Cross-Sectional Study. *The Journal of Psychiatric Nursing.* 2017;8(3):137–144. doi: 10.14744/phd.2016.36854.
18. Mostafazadeh P, Jafari MJ, Mojebi MR, Nemati-Vakilabad R, Mirzaei A. Assessing the relationship between nutrition literacy and eating behaviors among nursing students: a cross-sectional study. *BMC Public Health*. 2024;24(1):18. doi: 10.1186/s12889-023-17468-9.
19. Sengür E, Yilmaz G, Turasan I. Examination of Eating Behaviors of Students Studying in Sports Sciences Faculties of Universities. *Journal of Sports Education.* 2023;7(3):257-265. doi: 10.55238/seder.1373290.
20. Yurt S, Ozdemir T. Determination of Nutrition Exercise Behaviors of University Students. *JERN.* 2020;17(3):238-243. doi: 10.5222/HEAD.2020.58234.
21. Gökensel Okta P, ince N, Akyol Mutlu A. Determination of Eating Behaviors of Volleyball Players. *Research in Sport Education and Sciences.* 2022;24(3):87-92. doi: 10.5152/JPESS.2022.22714.
22. Kahraman S, Çiftçi EK, Yılmaz D. Comparison of nutritional habits of the nursing and faculty of economics students. *IAAOJ Health Sciences.* 2020;6(3):302-314.
23. Tanrıverdi D, Savaş E, Gönüllüoğlu N, Kurdal E, Balık G. Determination of High School Students' Eating Attitudes, Eating Behavior and Self-Esteem. *Eur J Therapeutics.* 2011;17(1):33-39. doi: 10.58600/eurjther.2011-17-1-749-arch.
24. Ceyhan H. The Relationship Between Eating Behaviors of Health Sciences Students with Anthropometric and Sociodemographic Features. Master's thesis, University of Istanbul, Istanbul 2021.
25. Ayaz A. Researching the eating attitudes and mindful eating of individuals aged 18-65. Master's thesis, University of Istanbul Medipol, Istanbul 2021.
26. Shinde VV. Relationship of body mass index to job stress and eating behaviour in health care professionals-an observational study. *Obes Med.* 2019;14:100070. doi: 10.1016/j.obmed.2018.12.002.

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Evaluation of clinical and neonatal outcomes of cancer in pregnancy and patients in remission

Burcu Bozkurt Ozdal [a,*], Atakan Tanacan [b], Ayse Altindis Bal [a], Ece Atalay [a], Esra Karatas [a], Sukru Bakirci [a], Fatma Doga Ocal [b], Dilek Sahin [b]

[a]Republic of Turkey Ministry of Health, Ankara City Hospital, Clinic of Perinatology, Ankara, Türkiye
[b]Republic of Turkey Ministry of Health, Ankara City Hospital, Clinic of Gynecology Obstetrics, Division of Perinatology, Ankara, Türkiye
*Corresponding author: burcu_bzkrt@hotmail.com (Burcu Bozkurt Ozdal)

MAIN POINTS

- It was conducted for the first time in Türkiye to examine the clinical problems and treatment treatments in patients receiving cancer treatment in detail and to create permanence in this field.
- It is compatible with clinical practices in Türkiye and adds new and local data to the literature by showing that maternal and neonatal pregnancies can be managed safely with treatment.
- The management of patients diagnosed with cancer during pregnancy by expert multidisciplinary teams, early diagnosis and adoption of appropriate treatment strategies are of great importance for both maternal and infant health

Cite this article as: Bozkurt Ozdal B, Tanacan A, Altindis Bal A, Atalay E, Karatas E, Bakirci S, Ocal FD, Sahin D. Evaluation of clinical and neonatal outcomes of cancer in pregnancy and patients in remission. *Ann Med Res*. 2026;33(1):27–32. doi: 10.5455/annalsmedres.2025.08.219.

ABSTRACT

Aim: To retrospectively examine patients diagnosed with cancer during pregnancy and compare the maternal and natal outcomes of these patients with those in cancer remission and healthy pregnant women.

Materials and Methods: This study was designed at a single tertiary care center. A total of 99 patients were included, of whom 18 were diagnosed with cancer during pregnancy, 21 had a pregestational history of cancer and were in remission, and 60 were healthy controls.

Results: Breast cancer was the most frequently detected malignancy during pregnancy. The average cancer during pregnancy was seen at the 21st week. Diagnostic methods included various biopsy methods and magnetic resonance imaging. Some patients underwent cancer-related surgery and chemotherapy during pregnancy. Two patients diagnosed with cancer during pregnancy died. Neonatal and maternal intensive care requirements were found to be higher in patients diagnosed with cancer during pregnancy compared to the other two groups. Additionally, neonatal birth weight was lower than that observed in healthy pregnancies.

Conclusion: We consider the monitoring and management of patients diagnosed with cancer during pregnancy to be a critical issue for ensuring optimal maternal and fetal health outcomes, and we emphasize the need for further research in this field.

Keywords: Active cancer during pregnancy, Maternal mortality, Neonatal intensive care requirements

Received: Aug 07, 2025 **Accepted:** Oct 27, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

A portion of cancers observed in women of reproductive age manifests during pregnancy [1], and cancers in this age group are among the leading causes of maternal mortality [1,2]. However, cancers diagnosed during pregnancy are exceedingly rare, with an incidence of approximately 0.38–0.50 per 1,000 patients [3]. The most common malignant neoplasms during pregnancy include breast cancer, thyroid cancer, cervical cancer, melanoma, Hodgkin's lymphoma, ovarian cancer, and leukemia [4].

The incidence of cancer during pregnancy is increasing due to the tendency of women to delay childbearing [5]. The care of pregnant women diagnosed with cancer is both medically and ethically challenging [6] and necessitates a multidisciplinary approach. The coexistence of maternal and fetal life raises concerns among physicians regarding the use of diagnostic and therapeutic modalities such as magnetic resonance imaging (MRI), computed tomography, positron emission tomography, surgery, chemotherapy, and radiotherapy. In recent years, surgical interventions during pregnancy have been more safely performed, as they do not pose significant risks to the fetus [5]. While earlier data suggested that chemotherapy and radiotherapy could have long-term adverse effects on the fetus, recent studies indicate that when administered after the first trimester, these treatments improve short-term pregnancy outcomes [7,8]. Consequently, patients diagnosed

with cancer during pregnancy undergo a risk-benefit assessment to determine the appropriateness of chemoradiotherapy [9].

This study aimed to evaluate patients diagnosed with cancer during pregnancy and compare their fetal and maternal outcomes with those of patients who had a history of cancer but achieved remission before pregnancy, as well as with healthy pregnancies.

MATERIALS AND METHODS

Study population

This retrospective study was conducted in a tertiary care center between January 2021 and March 2025. The study included patients diagnosed with cancer during pregnancy and those with a history of cancer who had achieved remission before conception. The names and diagnoses of newly admitted patients were recorded in a Microsoft Excel database maintained by our clinic. Additionally, patients with cancer were identified through the hospital records of pregnant women hospitalized in various departments and ICUs. Patient information was obtained from the hospital's database.

This retrospective study was approved by the Ethics Committee of Ankara City Hospital (TABED 1-25-868). Data were collected from electronic medical records in strict adherence to patient confidentiality, and all identifying information was anonymized. Informed consent was obtained from patients when necessary. All study procedures adhered to the principles of the Declaration of Helsinki.

The recorded data included the patients' age, gravida, parity, gestational week at cancer diagnosis, mode of cancer diagnosis, cancer type, cancer stage, cancer treatment received during pregnancy, maternal ICU requirement, maternal survival status, fetal outcomes, neonatal outcomes (gestational age at birth, birth weight, first- and fifth-minute Apgar scores, neonatal ICU [NICU] admission requirements), and remission duration for pregnant patients with a history of cancer.

The case group included patients diagnosed with cancer during pregnancy and those diagnosed with cancer before pregnancy that were in remission. The control group consisted of randomly selected pregnant women.

Statistical analysis

Statistical analysis was performed using the Statistical Package for the Social Sciences version 22.0 (SPSS Inc., Chicago, IL, USA). The normality of continuous variables was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Variables showing a normal distribution were compared using Tukey's post hoc test with one-way ANOVA. Categorical variables were compared using the chi-square exact test as appropriate. A p-value of 0.05 was considered statistically significant.

RESULTS

A total of 99 patients were included in this study, comprising 18 patients diagnosed with cancer during pregnancy, 21 patients with a pregestational history of cancer who were in remission, and 60 healthy controls.

The patients diagnosed with cancer during pregnancy were examined first. The distribution of cancer types diagnosed during pregnancy is summarized in. Breast cancer was the most common (27.7%), followed by thyroid cancer (22.2%).

Clinical data related to cancer types and maternal outcomes are summarized in Table 1. The mean gestational age at cancer diagnosis was 21 weeks (range: 8–40 weeks).

Cancer diagnoses were established via various methods: breast and thyroid cancers through fine-needle aspiration biopsy, ovarian cancer via ultrasound-guided biopsy, gastric cancer through endoscopic biopsy, and acute myeloid leukemia M3 (AML-M3) through frozen section analysis following splenic rupture.

Among these patients, metastasis was detected in nine cases, while cancer remained localized in nine cases. Eleven patients underwent cancer-related surgery during pregnancy, six received chemotherapy.

Maternal intensive care unit (ICU) admission was necessary for 11 patients, and seven did not require ICU care. There were two maternal deaths in this group.

Of the patients who received surgery and chemotherapy, only one did not require neonatal intensive care unit (NICU) admission, whereas nine required NICU support.

Among pregnant patients diagnosed with cancer, three had preterm deliveries, two underwent medical termination, and 11 delivered at term.

The distribution of cancer types among patients with a pregestational history of cancer in remission is provided in Table 2. In this group, thyroid and breast cancers were the most common. Metastatic relapse was observed in only one patient, who was previously diagnosed with borderline mucinous tumor. Gastric metastasis was observed.

The remission duration ranged from 3 to 180 months. Two patients in remission opted for elective pregnancy termination, while 18 pregnancies resulted in live births.

Table 3 compares the clinical and demographic differences between those diagnosed with cancer during pregnancy, neonatal outcomes, and maternal outcomes those in remission, and those with healthy pregnancies. Intra-group and inter-group differences for the variables of interest are presented in Table 4.

According to the post-hoc Tukey test, a difference was observed between the active cancer group and the control group in terms of maternal age ($p=0.041$), while no difference was observed between the cancer and remission groups ($p>0.05$). A difference was observed between the remission and control groups ($p=0.006$). Age was higher in the active cancer and remission groups.

https://doi.org/10.5455/annalsmedres.2025.08.219

Table 1. Clinical data and maternal outcomes by cancer type.

Variable	Breast cancer	Thyroid cancer	Lymphoma	Thymoma	Cervical cancer	Leukemia	Ovarian cancer	Bladder cancer	Gastric cancer
	n = 5	n = 4	n = 1	n = 1	n = 2	n = 1	n = 2	n = 1	n = 1
Gestational week at diagnosis (median)	14 w	17 w	30 w	12 w	12w	Postpartum first day	15 w	8 w	20 w
Diagnosis method	FNAB	FNAB	FNAB	MRI	Cervical biopsy	Postpartum splenic rupture frozen section	USG intraoperative frozen section	TUR biopsy	Endoscopic biopsy
Type of invasion									
Metastasis	2	0	1	1	1	2	1		1
Localized	3	4					1	1	
Surgery	3	3	0	0	2	1	2	0	0
Chemotherapy	3	0	1	0	0	0	2	0	0
Maternal ICU requirement	3	2	1	0	1	1	2		1
Maternal death	1	0	0	0	0	0	0	0	1
Medical termination	0	0	0	1 (23 w)	2 (18 w, 19w)	0	0	1 (14 w)	0
Preterm birth	1 (25 w)		1 (30 w)	0	0	0	0	0	1 (24 w)

w: week, FNAB: fine-needle aspiration biopsy, MRI: magnetic resonance imaging, USG: ultrasonography, TUR: transurethral resection.

Table 2. Cancer diagnoses of patients in remission*.

Tumor type	n	%
Thyroid cancer	9	42.8
Breast cancer	5	23.8
Lymphoma	3	14.2
Leukemia	2	9,5
Brain tumor	1	4.7
Metastatıc relaps(musınos borderline over carcinoma)	1	4.7
Total	21	100.0

*Remission durations ranged from 3 to 180 months.

Table 3. Clinical and demographic characteristics of patients and maternal and neonatal outcomes.

Variable	Pregnant women with active cancer (n = 18)	Pregnant women in cancer remission (n = 21)	Healthy pregnant women (controls) (n = 60)	p-value
Age	32.0±6.9[a]	33.0±6.9[a]	28.0±4.5[b]	**0.005**[1]
Gravida	2.8±1.0	3.5±2.4	2.0±1.3	0.650[1]
Parity	0.8±0.9[a]	1.6±1.3[b]	0.9±0.9[a]	**0.019**[1]
Gestational age at birth (week)	36.3±2.8[a]	37.6±1.1[b]	37.7±2.1[b]	**0.049**[1]
Birth weight (g)	2613.0±626.0[a]	3028.0±407.0[a]	3266.0±560.0[b]	**<0.001**[1]
First-minute Apgar score	8.4±0.7	8.7±0.7	8.9±0.5	0.585[1]
Fifth-minute Apgar score	7.1±0.9	7.2±0.9	7.4±1.3	0.535[1]
NICU requirement	8 (44.0%)	2 (9.5%)	0 (0.0%)	**<0.001**[2]
Medical Termination	4 (22.0%)	3 (14.2%)	0 (0.0%)	**<0.001**[2]
Spontaneous preterm labor	4 (22.0%)	0 (0.0%)	0 (0.0%)	**<0.001**[2]
Maternal ICU requirement	11 (61.0%)	0 (0.0%)	0 (0.0%)	**<0.001**[2]
Maternal outcome	16 (88.8%) survived 2 (11.2%) died	21 (100.0%) survived	60 (100.0%) survived	**<0.001**[2]

NICU: neonatal intensive care unit, ICU: intensive care unit. Quantitative variables are summarized as mean±sd, qualitative data as frequency (percentage).
1: One Way ANOVA test. 2: Pearson Chi-Square test. Note: Means not sharing subscripts differ significantly at a=0.05 as indicated by Tukey's HSD.

According to the post-hoc Tukey test, a difference was observed between the active cancer group and the remission group in parity (p= 0.042), while no difference was observed between cancer patients and the control group. A difference was observed between the remission group and the control group (p=0.041).

https://doi.org/10.5455/annalsmedres.2025.08.219

Table 4. ANOVA test results for meaningful variables

Variable	Source	df	SS	MS	F	p	$\eta 2$
Maternal Age	Between Groups	2	431.3	215.8	5.82	0.005	0.183
	Within Groups	52	1926.4	37.47			
	Total	54	2358.1				
Parity	Between Groups	2	9.969	4.985	4.309	0.019	0.102
	Within Groups	50	57.842	1.157			
	Total	52	67.811				
Gestational age at birth	Between Groups	2	65.841	32.920	6.510	0.002	0.138
	Within Groups	81	409.615	5.057			
	Total	83	475.456				
Birth weight	Between Groups	2	5181397.01	2590598.5	8.436	<0.001	0.172
	Within Groups	81	24875835.1	307109.0			
	Total	83	30057232.1				

P-value <0.05 is significant.

According to the post-hoc Tukey test, no difference was observed in birth weight between active cancer patients and patients in remission (p>0.05). A difference was observed between cancer patients and control patients (p=<0.001). A difference was observed between patients in remission and control patients (p=0.008). Birth weight was lower in the active cancer group and the remission group.

According to the post-hoc Tukey test, a difference was observed between patients in the cancer and remission groups in terms of gestational age (p= 0.029). A difference was observed between cancer patients and control patients (p=0.002). No difference was observed between remission and control patients (p>0.05). The active cancer group was found to have an earlier gestational age (Table 3).

A higher rate of premature birth and NICU requirement was also detected in this group. The need for NICU was determined to be 44.4%, while this rate was 0.0% in the control group (p<0.001). Similarly, maternal ICU requirement was significantly increased in the active cancer group, and the maternal mortality rate was also determined to be 11.2%; this was found to be significantly higher compared to the other groups (p<0.001).

DISCUSSION

In this study, we examined patients diagnosed with cancer during pregnancy and compared the fetal and maternal outcomes among pregnant women with a pregestational history of cancer in remission and healthy pregnant women. Breast cancer was the most frequently detected malignancy during pregnancy. Cancer diagnoses were made as early as eight weeks of gestation and as late as 40 weeks, with a mean gestational age at diagnosis of 21 weeks. Diagnostic methods included fine-needle aspiration biopsy, ultrasound, transurethral resection (TUR) biopsy, and MRI. Some patients underwent cancer-related surgery and chemotherapy during pregnancy. Two patients diagnosed with cancer during pregnancy died. However, when compared with other patient groups, no statistically significant difference was detected. NICU admission and maternal ICU requirements were found to be higher in patients diagnosed with cancer during pregnancy than in the other two groups. Additionally, neonatal birth weight was lower compared to healthy pregnancies.

Cancer during pregnancy is a complex condition that necessitates a multidisciplinary approach aimed at safeguarding the health of both the mother and the fetus. Based on this premise, we retrospectively analyzed cases of cancer diagnosed during pregnancy at one of Turkey's leading referral hospitals, which receives a high volume of patients from other provinces.

When considering the incidence of cancer during pregnancy, including diagnoses made up to 12 months postpartum, the rate is approximately 1 in 1,000 births; however, the incidence of cancer strictly during pregnancy is reported to be between 0.38% and 0.50% [10]. Breast cancer is the most frequently diagnosed malignancy [11], a finding that was consistent with our study.

In earlier periods, the management of cancer during pregnancy often involved either pregnancy termination or postponement of cancer treatment, as no standardized treatment approach was available [12]. Due to concerns about fetal harm, chemotherapy was administered reluctantly by physicians [13]. The absence of a standardized treatment protocol had adverse effects on both maternal and fetal health. However, in recent years, accumulating evidence from research has led to a shift in clinical practice [14]. Case-based data suggest that chemotherapy administered after the first trimester does not pose significant risks to the fetus [15]. Nevertheless, children exposed to cytotoxic treatment before birth should be monitored for long-term developmental follow-up [16]. Amant et al. showed in their study that children exposed to cytotoxic treatment in utero developed moderate cognitive developmental changes [17]. There is increasing evidence that abdominal surgery can be performed with relative safety in pregnant women [18]. Surgery is important

https://doi.org/10.5455/annalsmedres.2025.08.219

in the management of gynecological malignancies. However, surgical procedures performed during pregnancy should be performed by a surgical team with more experience in this area [18]. Despite all precautions, adverse obstetric outcomes such as preterm delivery, miscarriage, and fetal distress may occur after surgery during pregnancy [18]. Perioperatively, the mother's hemodynamic parameters should be closely monitored to ensure optimal fetal perfusion [18]. Additionally, the mother's risk of aspiration increases due to pregnancy-related gastroesophageal reflux [18]. In our study, some patients underwent chemotherapy and cancer-related surgical procedures during pregnancy, and no fetal loss was observed among these patients; however, their neonates required intensive care following birth.

The diagnostic methods used during pregnancy are also a source of concern for both patients and physicians. However, studies indicate that with appropriate protective measures, scattered radiation does not pose a significant risk to the fetus [19]. In our hospital, diagnostic procedures such as fine-needle aspiration biopsy, TUR biopsy, endoscopic biopsy, cervical biopsy, and MRI were utilized.

Maternal mortality among pregnant women diagnosed with cancer is closely associated with cancer type and stage [5]. In our study, the maternal mortality rate was 11.1%, a figure consistent with findings from a study conducted at another major hospital in Turkey [20]. The patients who died due to their illness had advanced-stage breast cancer and gastric cancer. In both cases, spontaneous labor occurred before death.

Cancer during pregnancy can influence the timing of delivery. A cohort study conducted in Italy found that cancer was associated with iatrogenic preterm birth, resulting in lower birth weight, reduced Apgar scores, and higher NICU admission rates [21]. Similarly, in our study, when comparing pregnant women diagnosed with cancer to those in remission and healthy pregnant women, neonates in the cancer group had lower birth weights and a higher need for intensive care.

Appropriate diagnostic and treatment approaches in patients diagnosed with cancer during pregnancy can be effective in protecting maternal and fetal health. Surgery and chemotherapy can be applied safely, especially in the second and third trimesters. Multidisciplinary teamwork and patient information positively affect treatment success and outcomes. Pregnancy follow-up and neonatal care services are critical in the management of these patients.

There is a pressing need for additional studies aimed at collecting more data on the diagnosis and treatment of cancer during pregnancy, which would help bridge existing knowledge gaps and contribute to improvements in clinical practice.

Single-center design, relatively small patient sample size, inability to match groups by age, lack of long-term follow-up data on newborns, and limited number of variables analyzed can be considered limitations of our study. This has constrained the depth of the statistical analysis. Furthermore, different cancer types and stages, as well as treatment methods, may have different effects on neonatal outcomes. The study data did not include information on the socioeconomic status of the mothers. This is one of the limitations of our study, and further research is needed to assess the impact of socioeconomic factors on neonatal and maternal outcomes. The retrospective design of the study resulted in some information being missing or limited in the data collection process. In particular, the limited data on socioeconomic status, detailed clinical stages, and treatment protocols are significant limitations of our study. Future prospective studies aim to address these shortcomings.In our study, these differences were not evaluated in detail; this is among the limitations of our study and is an important topic that should be investigated in future research.

Despite these limitations, this study provides valuable information about cancer incidence and outcomes during pregnancy in Turkey, is consistent with findings from recent studies, and contributes to the existing literature. In this study, clinical characteristics and maternal and fetal outcomes of cancer patients diagnosed during pregnancy or in remission were analyzed. Breast cancer is the most common malignancy and is usually diagnosed in mid-pregnancy. Treatment modalities such as surgery and chemotherapy during pregnancy can be safely performed with appropriate timing and multidisciplinary approach and have favorable or at least acceptable maternal-fetal outcomes. However, the need for neonatal intensive care and maternal intensive care is higher in pregnant women diagnosed with cancer, which is an important indicator for neonatal outcomes and maternal survival. It has been observed that maternal mortality is more common in advanced and aggressive tumors. In cases where cancer is diagnosed during pregnancy, multidisciplinary teamwork and personalized treatment plans are of vital importance. Early diagnosis and appropriate treatment can contribute to improved outcomes for both the mother and newborn.

CONCLUSION

This study highlights the need for caution in managing pregnancies undergoing oncological procedures, while emphasizing that multicenter studies can provide a broader data set and make significant contributions to future research.

Ethics Committee Approval: This study was approved by the Ethics Committee of Ankara City Hospital (TABED 1-25-868).

Informed Consent: Not necessary because the study is retrospective.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors state no conflicts of interest.

Author Contributions: BBO: Conceptualization, Methodology, Drafting the article, Supervision, Visualization; AT:

https://doi.org/10.5455/annalsmedres.2025.08.219

Conceptualization, Investigation, Drafting the article, Methodology, Visualization; AAB: Investigation, Data curation, Drafting the article; EA: Visualization; ŞB: Formal Analysis; ÖK: Data curation; EK: Visualization, Writing – original draft; DS: Analysis and interpretation of data, Validation, Visualization, Writing – review & editing.

Financial Disclosure: None declared.

Artificial Intelligence Disclosure: The authors declare that an artificial intelligence–based tool ChatGPT(GPT-4) was used solely for language editing and improving the clarity and readability of the manuscript. The AI tool did not contribute to the study design, data analysis, interpretation of results, or scientific conclusions. All content was critically reviewed and approved by the authors, who take full responsibility for the manuscript.

REFERENCES

1. Jamshidi Kerachi A, Shahlaee MA, Habibi P, et al. Global and regional incidence of intrahepatic cholestasis of pregnancy: a systematic review and meta-analysis. *BMC Med.* 2025;23(1):129. doi: 10.1186/s12916-025-03935-0.
2. Pawlus B, Zwolinski J, Koneczna U, et al. Neonatal outcomes for women diagnosed with cancer during pregnancy-single-center study. *Ginekol Pol.* 2024;95(6):443-450. doi: 10.5603/gpl.95610.
3. Schwab R, Anic K, Hasenburg A. Cancer and pregnancy: A comprehensive review. *Cancers.* 2021;13(12):3048. doi: 10.3390/cancers13123048.
4. de Haan J, Verheecke M, Van Calsteren K, et al. Oncological management and obstetric and neonatal outcomes for women diagnosed with cancer during pregnancy: a 20-year international cohort study of 1170 patients. *Lancet Oncol.* 2018;19(3):337-346. doi: 10.1016/S1470-2045(18)30059-7.
5. Van Calsteren K, Amant F. Cancer during pregnancy. *Acta Obstet Gynecol Scand.* 2014;93(5):443-6. doi: 10.1111/aogs.12380.
6. Barrois M, Anselem O, Pierga JY, et al. Cancer during pregnancy: Factors associated with termination of pregnancy and perinatal outcomes. *Eur J Obstet Gynecol Reprod Biol.* 2021;261:110-115. doi: 10.1016/j.ejogrb.2021.04.020.
7. Cardonick E, Iacobucci A. Use of chemotherapy during human pregnancy. *Lancet Oncol.* 2004;5(5):283-91. doi: 10.1016/s1470-2045(04)01466-4.
8. Kal HB, Struikmans H. Radiotherapy during pregnancy: fact and fiction. *Lancet Oncol.* 2005;6(5):328-33. doi: 10.1016/s1470-2045(05)70169-8.
9. Ring AE, Smith IE, Jones A, et al. Chemotherapy for breast cancer during pregnancy: an 18-year experience from five London teaching hospitals. *J Clin Oncol.* 2005;23(18):4192-7. doi: 10.1200/jco.2005.03.038.
10. Eibye S, Kjær SK, Mellemkjær L. Incidence of pregnancy-associated cancer in Denmark, 1977-2006. *Obstet Gynecol.* 2013;122(3):608-17. doi: 10.1097/AOG.0b013e3182a057a2.
11. Aktoz F, Yalcin AC, Yüzdemir HS, et al. Treatment of massive liver metastasis of breast cancer during pregnancy: first report of a complete remission with trastuzumab and review of literature. *J Matern Fetal Neonatal Med.* 2020;33(7):1266-1271. doi: 10.1080/14767058.2018.1517308.
12. Greiber IK, Viuff JH, Mellemkjaer L, et al. Cancer in pregnancy and the risk of adverse pregnancy and neonatal outcomes: A nationwide cohort study. *Bjog.* 2022;129(9):1492-1502. doi: 10.1111/1471-0528.17074.
13. Han SN, Kesic VI, Van Calsteren K, et al. Cancer in pregnancy: a survey of current clinical practice. *Eur J Obstet Gynecol Reprod Biol.* 2013;167(1):18-23. doi: 10.1016/j.ejogrb.2012.10.026.
14. Vandenbroucke T, Verheecke M, van Gerwen M, et al. Child development at 6 years after maternal cancer diagnosis and treatment during pregnancy. *Eur J Cancer.* 2020;138:57-67. doi: 10.1016/j.ejca.2020.07.004.
15. Bartsch E, Park AL, Young J, et al. Infant feeding practices within a large electronic medical record database. *BMC Pregnancy Childbirth.* 2018;18(1):1. doi: 10.1186/s12884-017-1633-9.
16. Amant F, Van Calsteren K, Halaska MJ, et al. Long-term cognitive and cardiac outcomes after prenatal exposure to chemotherapy in children aged 18 months or older: an observational study. *Lancet Oncol.* 2012;13(3):256-64. doi: 10.1016/S1470-2045(11)70363-1.
17. Vandenbroucke T, Verheecke M, van Gerwen M, et al. Child development at 6 years after maternal cancer diagnosis and treatment during pregnancy. *Eur J Cancer.* 2020 Oct;138:57-67. doi: 10.1016/j.ejca.2020.07.004.
18. Amant F, Halaska MJ, Fumagalli M, et al. Gynecologic cancers in pregnancy: guidelines of a second international consensus meeting. *Int J Gynecol Cancer.* 2014;24(3):394-403. doi: 10.1097/igc.0000000000000062.
19. Wieseler KM, Bhargava P, Kanal KM, et al. Imaging in pregnant patients: examination appropriateness. *Radiographics.* 2010;30(5):1215-29; discussion 1230-3. doi: 10.1148/rg.305105034.
20. Tunç SY. Kadın Hastalıkları ve Doğumda Güncel Konular III. Livre de Lyon; 2024.
21. Esposito G, Franchi M, Dalmartello M, et al. Obstetric and neonatal outcomes in women with pregnancy associated cancer: a population-based study in Lombardy, Northern Italy. *BMC Pregnancy Childbirth.* 2021;21(1):31. doi: 10.1186/s12884-020-03508-4.

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Evaluation of conventional histopathological scoring in breast carcinoma using artificial intelligence technologies

Seda Cetinkaya Karabekir [a,*], Aylin Gokhan [b], Hilal Arslan [a], Burak Keskin [c], Adile Ferda Dagli [d]

[a] İzmir Bakırçay University, Faculty of Medicine, Department of Histology and Embryology, İzmir, Türkiye
[b] Ege University, Faculty of Medicine, Department of Histology and Embryology, İzmir, Türkiye
[c] İzmir Bakırçay University, Department of Management Information Systems, İzmir, Türkiye
[d] İzmir Bakırçay University, Faculty of Medicine, Department of Medical Pathology, İzmir, Türkiye
*Corresponding author: seda.karabekir@bakircay.edu.tr (Seda Cetinkaya Karabekir)

MAIN POINTS

- A deep learning–based model was developed to evaluate the Ki-67 proliferation index in breast carcinoma and compared with manual assessment.
- The model showed high accuracy and low error rates in low-proliferation cases (<20% Ki-67) (MAE: 5.31%; accuracy: 80%).
- AI-assisted evaluation has the potential to reduce interobserver variability, minimize human error, and ease the workload of pathologists during the preliminary assessment phase.
- Error analysis in high-proliferation samples revealed segmentation challenges due to dense nuclear clustering and faint immunonegative staining, indicating that advanced image-processing techniques could further improve model performance.

Cite this article as: Cetinkaya Karabekir S, Gokhan A, Arslan H, Keskin B, Dagli AF. Evaluation of conventional histopathological scoring in breast carcinoma using artificial intelligence technologies. *Ann Med Res*. 2026;33(1):33--42. doi: 10.5455/annalsmedres.2025.08.244.

ABSTRACT

Aim: Breast cancer is the most commonly diagnosed malignancy in women, and early detection plays a critical role in the success of treatment. The Ki-67 proliferation index is widely used to evaluate tumor cell proliferation; however, its manual scoring process is observer-dependent, time-consuming, and inherently subjective. This study aims to assess Ki-67 immunohistochemical staining using deep learning algorithms in an objective, rapid, and reproducible manner, and to compare the model's performance with conventional scoring methods.

Materials and Methods: In the first phase of the study, a dataset was created using digital images of Ki-67-stained histological sections obtained from patients diagnosed with breast cancer. These images were used to train a machine learning algorithm. In the second phase, 50 new Ki-67-stained tissue sections previously unseen by the model were digitized, and the model's predictions were compared with Ki-67 index values calculated by conventional manual assessment.

Results: The developed model achieved a mean absolute error (MAE) of 8.69%, a root mean square error (RMSE) of 13.00%, and a coefficient of determination (R^2) of 0.540 in overall prediction performance. For cases with low proliferation (Ki-67<20%), the model demonstrated high accuracy (MAE: 5.31%). Binary classification based on a 20% threshold yielded 80% accuracy, 80% sensitivity, 90% precision, and an F1 score of 0.84.

Conclusion: The use of artificial intelligence algorithms in Ki-67 assessment demonstrated successful performance, with an MAE of 8.69%, and has the potential to reduce pathologists' workload during the preliminary evaluation phase. The findings suggest that, with further refinement, the proposed model could contribute to more objective, consistent, and reproducible assessments in breast cancer diagnostics.

Keywords: Breast cancer, Ki-67, Artificial intelligence, Deep learning, Artificial neural network

Received: Aug 29, 2025 **Accepted:** Oct 31, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

Each year more than half a million women worldwide die from breast cancer. To reduce mortality associated with breast cancer, many countries have implemented screening programs over the past two decades. These programs have led to a nearly 30% reduction in breast cancer-related deaths by enabling earlier diagnosis, combined with advances in treatment. Nevertheless, breast cancer remains the leading cause of cancer-related mortality among women [1]. Breast cancer encompasses various subtypes [2]. According to the World Health Organization (WHO) Classification of Tumours of the Breast (5th edition, 2019), breast cancer is categorized into well-defined histological subtypes. These include ductal carcinoma in situ (DCIS), invasive carcinoma of no special type (NST), invasive lobular carcinoma, tubular carcinoma, mucinous carcinoma, medullary carcinoma, apocrine

https://doi.org/10.5455/annalsmedres.2025.08.244

carcinoma, metaplastic carcinoma, and other rare variants. In addition, molecular subtypes such as hormone receptor-positive/HER2-negative, HER2-positive, and triple-negative breast cancers are also recognized due to their prognostic and therapeutic significance. This classification provides a more standardized approach to diagnosis, research, and treatment planning [3].

Early and definitive diagnosis of breast cancer, including the identification of its specific subtype, is crucial in preventing disease progression and associated complications. Accurate diagnosis enables timely and effective treatment planning and ultimately reduces breast cancer-related mortality. Various imaging modalities are used for breast cancer detection, including mammography (MG) [4], breast thermography (BT) [5], magnetic resonance imaging (MRI) [6], positron emission tomography (PET), computed tomography (CT) [7], ultrasound (US) [8] and histopathology (HP) [9]. These methods are widely used for detecting breast cancer at early stages [10-12].

Among these, histopathological evaluation plays a critical role in diagnosing, staging, and determination of breast cancer by analyzing tissue specimens. Accurate staging of the cancer is essential for selecting the most appropriate treatment strategy. During this process, certain prognostic markers such as the proliferation index (PI) and mitotic activity are used to evaluate tumor biology. These markers help predict disease progression and guide treatment decisions. Ki-67 is a nuclear antigen expressed in proliferating but not quiescent cells, making it a reliable marker for calculating the PI. The Ki-67 proliferation index is determined by evaluating immunohistochemically stained tumor sections under a microscope. The most densely stained region, referred to as the "hot spot," is selected, and the proportion of Ki-67-positive nuclei relative to the total number of nuclei is calculated and reported as the proliferation index [13, 14].

However, histopathological assessment of breast tissue is highly subjective and depends on the experience of the pathologist [15-17]. Manual Ki-67 scoring is associated with significant interobserver variability, particularly around key clinical thresholds (e.g., 14% or 20%), and lacks standardized criteria for the number of high-power fields to be evaluated [18]. These limitations introduce uncertainty into molecular subtype classification and therapeutic decision-making [17]. Therefore, there is growing interest in the application of artificial intelligence (AI)-based approaches to reduce subjectivity in pathological assessment, particularly in breast pathology [15-17].

Numerous AI-based tools utilizing machine learning (ML) and deep learning (DL) have been developed for tasks such as breast cancer classification, detection, and segmentation [19].

Recent comprehensive reviews and applied studies highlight that artificial intelligence can be integrated into all stages of breast pathology, including classification, grading, biomarker quantification, and risk prediction; in particular, CNN and Vision Transformer (ViT) based approaches have been shown to provide robust feature learning at the whole-slide image (WSI) level [20-22]. In parallel, methods have been reported for automatic scoring and risk stratification in fine distinctions such as HER2-low status and in key biomarkers such as Ki-67 [23]. Nevertheless, significant barriers to routine clinical adoption remain, including validation, generalizability, and standardization. Studies published in 2024–2025 emphasize that digital image analysis tools especially for Ki-67 evaluation have not yet provided sufficient evidence to be deemed fully suitable for clinical practice, underscoring the need for large-scale, multicenter validation [24, 25].

Artificial intelligence was first conceptualized by Alan Turing in 1950 and later defined by John McCarthy in 1956 as “the science and engineering of making intelligent machines [26, 27]. Today, AI represents a broad domain within computer science aimed at developing machines capable of performing tasks that typically require human intelligence [28]. These systems can store experiences, learn from them, reason, create, judge, and make decisions. A critical subset of AI, machine learning (ML)—particularly in conjunction with deep learning (DL)—has enabled significant progress in image processing [29].

DL employs multilayered architectures such as convolutional neural networks (CNNs), which mimic human visual perception and excel at tasks such as classification and segmentation [30]. Due to their ability to extract meaningful features from raw data, DL-based models have proven highly effective in analyzing pathology images. In this context, AI applications in pathology are primarily based on the recognition of visual patterns via CNNs and the extraction of diagnostic information from digitized slides [29].

The classification, detection, and staging of breast cancer are traditionally conducted by pathologists based on theoretical knowledge and clinical experience. However, this process is labor-intensive and time-consuming. Although accumulated observational expertise is invaluable, the potential for human error and time constraints associated with manual protocols cannot be overlooked. In this context, the present study aims to investigate the feasibility of evaluating Ki-67 a key biomarker in breast carcinoma diagnosis using artificial intelligence algorithms after initial manual assessment by pathologists. This approach aims to minimize human error and enhance the reliability of pathological evaluations.

MATERIALS AND METHODS

This study was approved by the Non-Interventional Clinical Research Ethics Committee of İzmir Bakırçay University (Decision No: 1357, Date: 13.12.2023). The research was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki, with full respect for human rights.

https://doi.org/10.5455/annalsmedres.2025.08.244

Between January 2021 and March 2025, a total of 150 patients with histopathologically confirmed breast carcinoma whose Ki-67 index had been assessed at the Department of Pathology, İzmir Bakırçay University Çiğli Training and Research Hospital, were included. All slides were digitized at 40× magnification, and only cases meeting predefined technical quality criteria (adequate focus, contrast, and background) were analyzed. Inclusion in the training set additionally required the availability of archival FFPE (formalin-fixed, paraffin-embedded) tissue blocks. Cases with pronounced staining artifacts, excessive background, or other technical defects rendering images unsuitable for analysis were excluded. An a priori power analysis (G*Power 3.1) for a two-tailed Pearson correlation with $\alpha = 0.05$, power = 0.90, and an anticipated effect size $r \approx 0.26$ indicated a required total sample size of 150.

Dataset construction

The study focused on the Ki-67 proliferation index of patients with breast cancer evaluated during routine diagnostic procedures at the Department of Pathology. A total of 100 digital images of breast cancer tissue sections immunohistochemically stained for Ki-67 were used. Imaging was performed using a ZEISS Axiolab 5 digital laboratory microscope integrated with a Zeiss AxioCam ERc 5s camera, and image data were recorded using ZEN 3.4 (Blue Edition) software. These images formed the digital dataset used to train the machine learning (ML) model.

Histopathological evaluation

Immunohistochemically stained breast cancer tissue sections were evaluated manually by experienced pathologists. During the scoring process, a total of 1,000 tumor cells were counted across the tissue section by identifying nuclei showing Ki-67 positivity, and the Ki-67 proliferation index was calculated. This method, which is widely accepted in literature, is considered reliable for assessing tumor cell proliferative activity and predicting breast cancer prognosis. Counting was performed in the area of highest labeling intensity, known as the 'hot spot,' at 40× magnification.

Use of deep learning algorithms

A deep learning-based model was developed to facilitate faster, more objective, and reproducible evaluation of cell proliferation indices derived from Ki-67 immunohistochemistry (IHC) images, which are widely used in breast cancer diagnostics. For the task of cell segmentation, Cellpose—a general-purpose cell segmentation algorithm built on a convolutional neural network (CNN) architecture was employed [31]. Cellpose is capable of accurately distinguishing cellular structures such as nuclei and cytoplasm in biological images.

The images were examined by an expert pathologist, and 'hot-spot' areas with the highest labeling intensity were included in the analysis; stromal and inflammatory cells were excluded. Using the Cellpose interface, two separate models were trained: one configured to segment all cells and the other to detect only Ki-67–stained cells. During model training, all nuclei were manually annotated, and positive/negative classification was performed based on staining intensity and nuclear morphology. No pre-processing (e.g., color normalization, contrast enhancement) was applied. The hyperparameters used during training included a learning rate of 0.1, a weight decay of 0.0001, and 100 epochs. After training, both models were run for each image in the test dataset and the results were recorded. Validation of the test data was performed by the expert pathologist, after which the differences between the ground truth and the model predictions were calculated as percentages. Finally, statistical analyses were applied to demonstrate the reliability of these differences.

Model training and optimization

The model was trained using high-resolution digital images of Ki-67-stained tissue sections obtained from retrospectively confirmed cases of breast carcinoma. No preprocessing techniques (e.g., contrast enhancement, color normalization, noise reduction) were applied to maintain the visual fidelity of images as perceived by pathologists and to ensure fair comparison between manual and automated assessments.

Cellpose, an open-source segmentation algorithm, was deployed in a local Python environment on a personal computer. Leveraging its flexibility, two separate models were trained using both pre-trained weights and task-specific datasets:

1. A general model for detecting and counting all cells.

2. A specialized model for identifying and counting only Ki-67-positive cells (Figure 1).

The Ki-67 proliferation index for each image was calculated as the ratio between the outputs of these two models.

Training was conducted using a supervised learning approach on 100 Ki-67-stained histological images acquired at 40x magnification. All cells and Ki-67-positive nuclei in these images were manually annotated, and the resulting segmentation masks were formatted in .npy files for model input (Figure 2).

During training, model weights and hyperparameters were optimized to enhance segmentation performance. Training was executed on a personal computer equipped with an Intel i7-13650HX processor, 32 GB RAM, 512 GB SSD, and an NVIDIA RTX 4060 GPU (Figure 2).

Model validation with test dataset

Model validation was performed using an independent test dataset curated and verified by expert pathologists. This dataset comprised previously unseen samples not used during training. Accordingly, paraffin blocks from 50 histologically confirmed breast cancer cases were retrieved from the

https://doi.org/10.5455/annalsmedres.2025.08.244

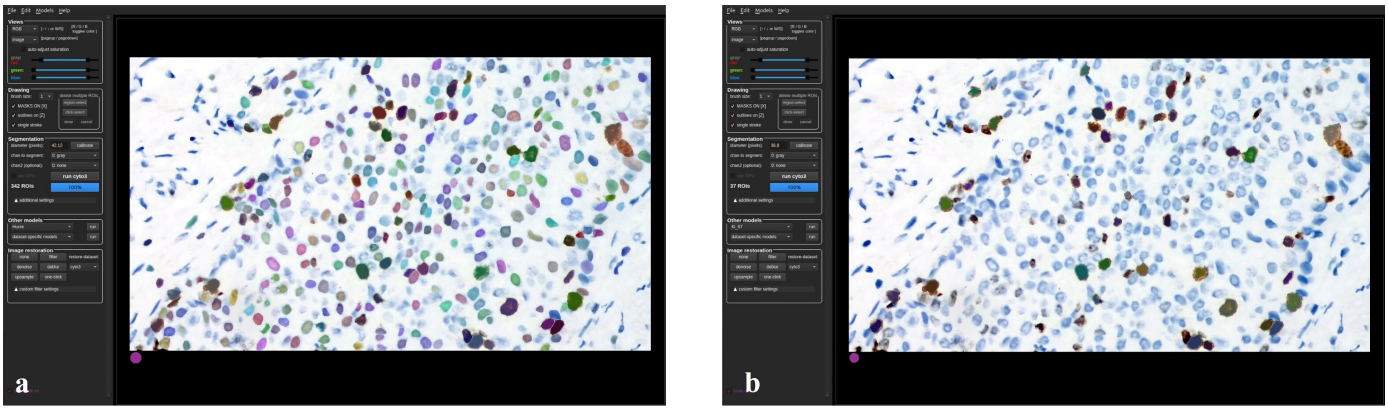

Figure 1. Screenshots of model outputs: (a) model detecting and counting all cells in the image, (b) specialized model detecting and counting only Ki-67–positive cells.

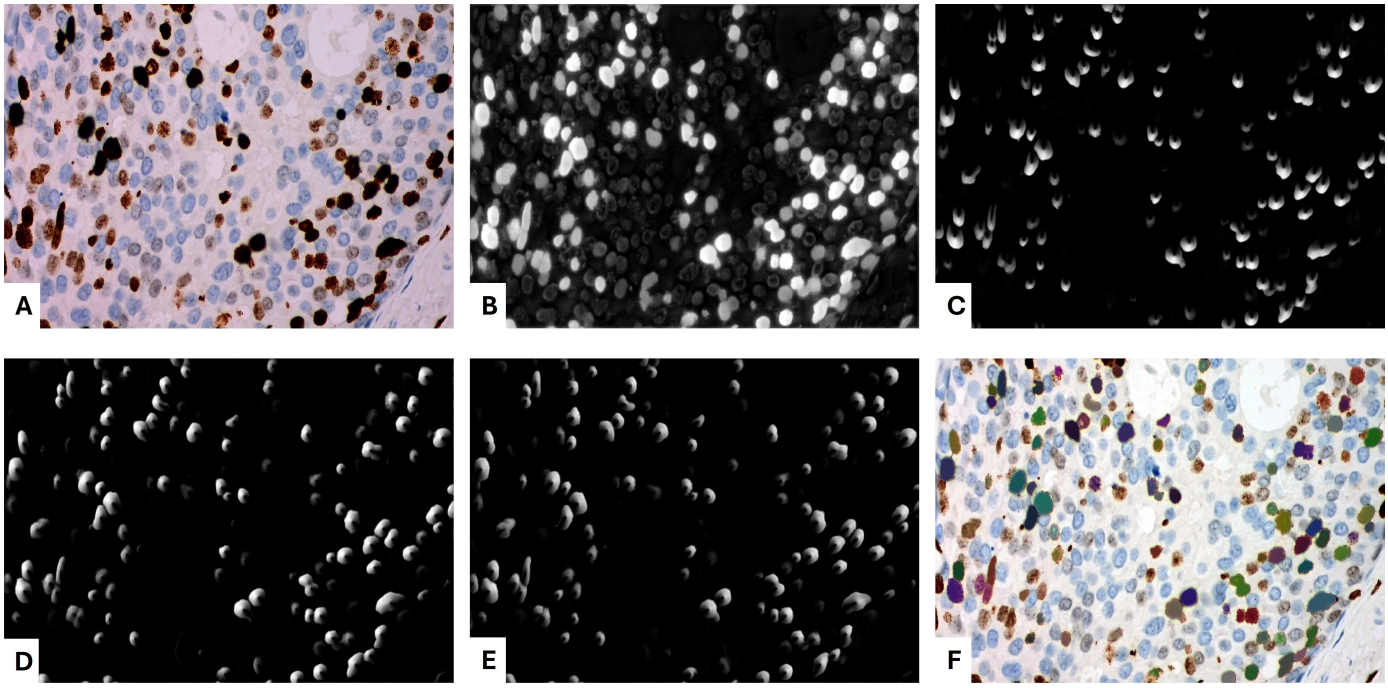

Figure 2. Diagram illustrating the workflow of the AI model from patch extraction to cell classification. (A) Original image; (B–E) segmentation steps including patch extraction and masking; (F) final output image showing the detection and separation of Ki-67–positive (brown-stained) and negative cells.

pathology archives of İzmir Bakırçay University Çiğli Training and Research Hospital, and tissue sections were prepared via microtomy. These sections were then subjected to Ki-67 immunohistochemical staining in the Histology and Pathology Laboratory of the university.

Immunohistochemical staining protocol

Paraffin-embedded tissue blocks from breast cancer patients were sectioned at 5 μm thickness using a microtome. Ki-67 immunostaining was performed using a primary antibody (Abcam, ab16667). Deparaffinization was achieved with xylene for 30 minutes, followed by hydration through graded alcohols. To block non-specific binding, sections were incubated in citrate buffer (pH 6.0) and 3% hydrogen peroxide, then treated with Ultra V Block. Subsequently, sections were incubated overnight with the Ki-67 primary antibody, followed by a 10-minute incubation with the secondary antibody. Signal development was achieved using streptavidin-peroxidase and diaminobenzidine (DAB). After chromogenic reaction, Mayer's hematoxylin was used as a counterstain. Coverslips were mounted with Entellan, and sections were visualized using a Zeiss Lab.A1 light microscope and photographed with the Zeiss AxioCam ERc 5s imaging system.

Preparation of validation dataset for performance evaluation

Following IHC staining, the most intensely labeled nuclear regions ("hot spots") were selected from each sample, and 40x magnification digital images were acquired. These 50 images were reserved exclusively for the validation phase of the study.

https://doi.org/10.5455/annalsmedres.2025.08.244

Following segmentation and classification, the Ki-67 index values predicted by the model were compared against manually calculated values obtained by an experienced pathologist from the same images. This comparison served to assess the predictive accuracy of the model.

The validation phase was designed to evaluate the clinical reliability, validity, and applicability of the proposed model. The results demonstrated the model's predictive performance and highlighted the potential of AI-based systems in histopathological assessment.

Statistical analysis

The deep learning model's performance was evaluated using both regression and classification metrics. Prediction accuracy was quantified using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Pearson Correlation Coefficient (r), alongside a scatter plot to visualize the relationship between predicted and actual values. For binary classification, a clinical Ki-67 threshold of 20% was applied. A confusion matrix was generated to calculate Accuracy, Sensitivity (Recall), Precision, Specificity, and F1 Score, assessing the model's capacity to differentiate between low and high proliferative index groups.

RESULTS

Model performance and overall statistical evaluation

The performance of the Ki-67 proliferation index prediction model was evaluated using five fundamental regression metrics. All statistical calculations were performed using Microsoft Excel, and the selected metrics were used to assess the model's predictive capacity from multiple dimensions.

A quantitative summary of model performance is presented in Table 1.

The mean absolute error (MAE) was calculated as 8.69%, indicating an average deviation of ±8.69 percentage points in the model's predictions. The mean squared error (MSE) was found to be 169.07, while the root mean square error (RMSE) was 13.00%, suggesting that predicted values deviated from actual values by an average of 13 points, representing the standard deviation of the error distribution.

The coefficient of determination (R^2), calculated to evaluate the explanatory power of the model, was 0.540. This value indicates that the model explains 54% of the total variance in the Ki-67 proliferation index. Furthermore, the Pearson correlation coefficient (r) between the predicted and actual values was calculated as 0.753, indicating a statistically significant and strong positive correlation between the two variables.

Overall, the developed model demonstrated consistent predictive performance, particularly in low-to-moderate Ki-67 value ranges, and was capable of establishing a strong linear relationship with approximately 75% correlation strength.

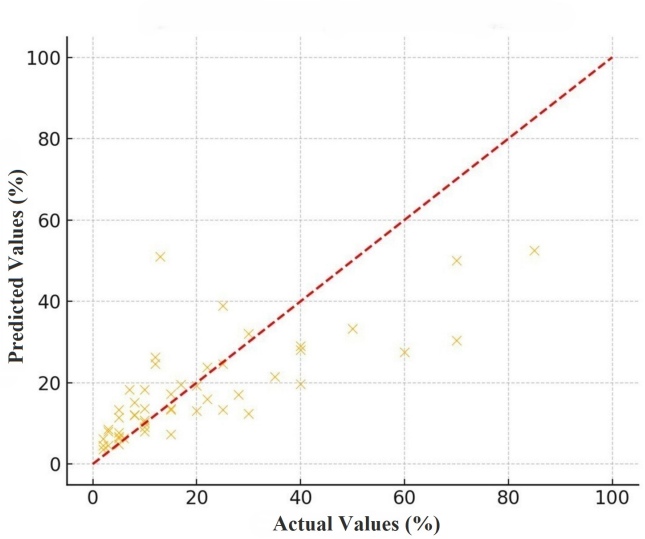

Figure 3. Visual analysis of model performance: relationship between actual Ki-67 index values and those predicted by the model.

Analysis of predicted and actual values

A visual analysis of the model's performance revealed the relationship between predicted and actual Ki-67 values. The results indicated that the model produced generally consistent predictions (Figure 3).

To further investigate model performance, the five most accurate and five most erroneous predictions were analyzed. The analysis of the top five predictions showed high accuracy, particularly in cases with low Ki-67 levels. Conversely, the most erroneous predictions revealed systematic issues in estimating high Ki-67 values (Figure 4).

Error analysis demonstrated that the model performed more accurately in the 10–30% prediction range, while substantial underestimation was observed for values ≥50% (Table 2).

Class-based error analysis

Given the clinical significance of Ki-67 values, a class-based performance analysis was performed. For this purpose, true Ki-67 values were dichotomized using a 20% threshold, separating the cases into two groups: low proliferation (<20%) and high proliferation (≥20%). This threshold aligns with the commonly used luminal subtype classification in the literature and is critical for tumor aggressiveness and prognosis evaluation. For each class, MAE and RMSE were calculated.

The error levels for low Ki-67 values (<20%) were significantly lower, suggesting that the model had better predictive accuracy for this group. In contrast, prediction errors increased in the high Ki-67 group (≥20%), with the MAE reaching 12.32 (Table 3).

Trend and subgroup-based evaluation of model predictions

The Ki-67 prediction model was evaluated using 50 paired observations, yielding mean values of 19.86 for real and 17.92

https://doi.org/10.5455/annalsmedres.2025.08.244

Table 1. General evaluation of model performance metrics.

Metric	Value	Description
Mean Absolute Error (MAE)	8.69%	Average prediction error
Mean Squared Error (MSE)	169.07	Weighting of large errors
Root Mean Squared Error (RMSE)	13.00%	Measure of standard deviation
R^2 Score	0.540	Proportion of explained variance
Correlation Coefficient (r)	0.753	Strength of linear relationship

Table 2. The actual values, model predictions, and absolute error percentages of the five best and five worst predictions of the artificial intelligence model regarding the Ki-67 index.

Best 5 Predictions	Actual Value (%)	Predicted Value (%)	Absolute Error (%)
a	10	10.14	+0.14
b	5	4.83	-0.17
c	6	6.35	+0.35
d	25	24.64	-0.36
e	10	10.76	+0.76
Worst 5 Predictions	**Actual Value (%)**	**Predicted Value (%)**	**Absolute Error (%)**
f	70	30.42	-39.58
g	13	51.04	+38.04
h	85	52.50	-32.50
i	60	27.58	-32.42
j	40	19.71	-20.29

Table 3. Class-based performance analysis of the model according to Ki-67 proliferation levels. The model demonstrated low error rates and high accuracy in the <20% class, whereas its performance declined in the ≥20% class.

Class	Sample Size	MAE	RMSE	Performance Evaluation
<20%	31	5.31	8.79	High performance
≥20%	19	12.32	14.19	Moderate performance

Table 4. Performance metrics used in the binary classification analysis of the Ki-67 proliferation index.

Metric	Value	Definition	Description
Accuracy	0.80	Proportion of correctly classified samples	Overall success rate
Error Rate	0.20	Proportion of misclassified samples	Overall error rate
Sensitivity (Recall)	0.80	Proportion of correctly detected <20 cases	Detection of low Ki-67
Precision	0.90	Proportion of correct <20 predictions	Reliability of positive predictions
Specificity	0.80	Proportion of correctly detected ≥20 cases	Detection of high Ki-67
F1 Score	0.84	Harmonic mean of sensitivity and precision	Balanced performance measure

for predicted data. The results demonstrated a strong correlation (r = 0.753) and no significant overall bias according to the paired t-test (mean difference = 1.94, p = 0.296), suggesting that the model's predictions were statistically consistent with the actual values.

Subgroup analysis revealed that the model's predictive accuracy varied across Ki-67 value ranges. In the low-proliferation group (<20%, n = 31), prediction errors were small (MAE ≈ 5.3, RMSE ≈ 8.8) and statistically aligned with real values (p ≈ 0.003), indicating that the model performed reliably in this clinically relevant range. For the moderate-high group (≥20%, n = 19), errors increased (MAE ≈ 12.3, RMSE ≈ 14.2) but still followed the correct statistical trend, showing no major deviation in direction. However, in high-proliferation subgroups (≥50%, n = 5 and ≥70%, n = 3), significant underestimation was observed, with higher error magnitudes (MAE ≈ 28.2–30.7; RMSE ≈ 29.5–31.7) and corresponding p-values of ≈0.003 and 0.03.

These findings indicate that while the model captured the correct overall trend between predicted and real Ki-67 values, its accuracy decreased in extreme ranges, primarily due to the limited number of high-proliferation samples. The reduced sample size in these subgroups likely contributed to increased error variance and lower statistical stability. Overall, the model performed well for low-to-moderate Ki-67 indices, demonstrating reliable predictive capacity, but tended to underestimate values in clinically critical high-proliferation cases.

https://doi.org/10.5455/annalsmedres.2025.08.244

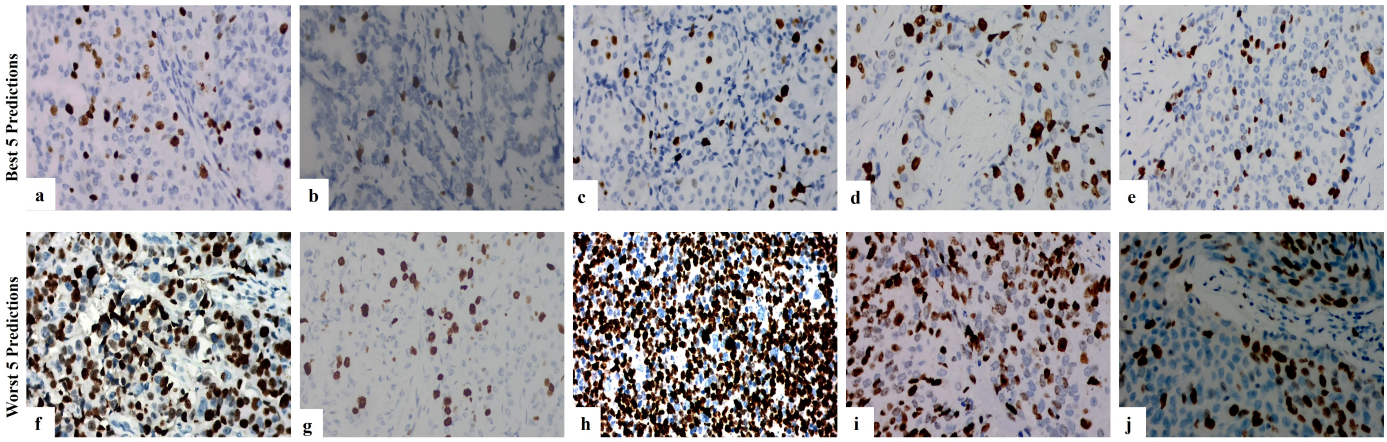

Figure 4. Example images of the five best (a, b, c, d, e) and five worst (f, g, h, i, j) predictions of the AI model for the Ki-67 index. (The actual values, model predictions, and absolute error percentages for these examples are presented in Table 2).

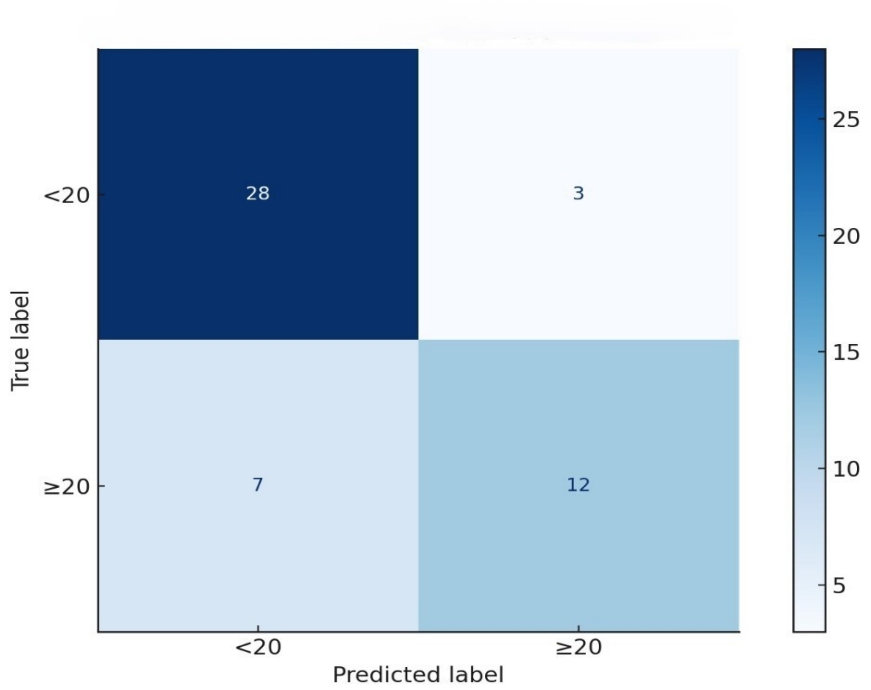

Figure 5. Confusion matrix of the binary classification analysis using the 20% threshold for Ki-67.

Confusion matrix and classification performance

To evaluate the model's performance in binary classification, confusion matrix–based performance metrics were calculated. Classification was performed using the 20% Ki-67 threshold to divide cases into two categories (Figure 5).

The model achieved an accuracy of 0.80, indicating that 80% of predictions were correct overall. Sensitivity was also 0.80, meaning that the model correctly identified 80% of low Ki-67 cases. Consequently, 20% of high Ki-67 cases were missed. The precision was calculated as 0.90, showing that among the cases predicted to have <20% Ki-67, 90% were indeed true positives. This high precision underscores the model's reliability in estimating low proliferation indices (Table 4).

DISCUSSION

The deep learning-based model developed in this study yielded promising results for the objective, reproducible, and rapid assessment of the Ki-67 proliferation index in breast cancer samples. The model demonstrated acceptable diagnostic sensitivity, particularly in low-proliferation cases (<20%), with a mean absolute error (MAE) of 5.31% and a high F1 score of 0.84.

Conventional pathological assessment of the Ki-67 index via manual counting is subject to interobserver variability and human error, as previously reported in the literature [32]. Furthermore, this variability particularly at clinically relevant thresholds such as 10–30% can significantly influence treat-

https://doi.org/10.5455/annalsmedres.2025.08.244

ment decisions [33]. In this context, the potential of digital image analysis and AI-assisted systems to improve objectivity and promote standardization in pathological evaluation is noteworthy [34].

Recent studies have demonstrated the efficacy of digital pathology and deep learning approaches for similar tasks. For instance, Cireşan et al. (2013) achieved high accuracy in mitotic figure detection using a CNN-based model [35]. Similarly, Veta et al. (2019), through the TUPAC16 challenge, showcased the effectiveness of machine learning algorithms in predicting breast tumor proliferation [34].

The sensitivity of the model in binary molecular subtype classification using the 20% Ki-67 threshold was found to be 80%, suggesting that some highly proliferative tumors may be missed. Given the clinical importance of accurately identifying high-risk cases, the incorporation of secondary control mechanisms such as pathologist review is recommended before integration into clinical workflows.

The observed high accuracy of the model in cases with low proliferation aligns with previous findings. Matsumoto et al. (2025) reported that their deep learning model demonstrated strong agreement (q = 0.961) with manual scoring in 494 breast cancer cases and showed low error rates in low Ki-67 samples. This supports our results, particularly the low MAE of 5.31% for <20% Ki-67 values [36]. Likewise, Dy et al. (2024) showed that AI-assisted scoring significantly reduced error rates (from 5.9% to 2.1%), improved interobserver agreement (ICC: 0.70 → 0.92), and decreased evaluation time in a study involving 90 international pathologists. Notably, AI applications were especially effective in standardizing assessments within the 5–30% Ki-67 range, thereby reducing variability [37].

These findings support the low error rates observed in our study for low Ki-67 values.

Additionally, error analysis of cases with Ki-67 positivity ≥20% revealed that the model's limitations were not solely due to class imbalance in the training dataset, but also to morphological and staining-related segmentation challenges. Analysis of the five most erroneous predictions revealed two major issues. First, in regions with 85–90% immunopositive nuclei, intense DAB staining and nuclear clustering impaired the model's ability to differentiate individual nuclei, leading to systematic underestimation of cell counts. This finding is consistent with prior reports indicating diminished segmentation performance in densely packed cell regions [38, 39]

Second, some immunonegative nuclei were too faintly stained (pale blue) to be recognized by the segmentation algorithm. This issue is commonly observed in threshold-based or traditional intensity-mapping segmentation approaches [31]. Moreover, our model intentionally excluded any preprocessing such as color filtering to preserve the visual appearance of slides as perceived by pathologists. While this approach ensured objectivity in comparative evaluation, studies suggest that integrating color space optimization may enhance classification accuracy [40].

Limitations

Although the model developed in this study demonstrated high accuracy in low-proliferation cases, a decline in performance was observed in high Ki-67 values (≥20%). This limitation is not solely attributable to class imbalance in the training dataset but is also associated with the absence of preprocessing steps (e.g., color normalization, contrast adjustment) and the challenges posed by intense DAB staining and nuclear clustering, which complicated segmentation. In addition, the very faint staining of some immunonegative nuclei restricted the model's discriminative capacity. These methodological and technical constraints may limit the generalizability of the findings and their applicability to clinical practice.

CONCLUSION

This study demonstrated that a deep learning-based model can provide rapid, objective, and reproducible assessments of the Ki-67 proliferation index in breast carcinoma samples. The model's high accuracy in low-proliferation cases supports its potential role as a clinical decision-support tool. However, segmentation errors observed in highly proliferative samples underscore the need for advanced image processing techniques to improve the model's sensitivity to morphological variability.

Future work will aim to integrate multi-channel color normalization, histogram equalization, and adaptive thresholding techniques into the segmentation pipeline. Additionally, data augmentation strategies will be employed to ensure a more balanced representation of complex and challenging images in the training dataset. These enhancements are expected to improve overall model performance and facilitate the broader and more reliable application of AI-assisted systems in histopathological diagnostics.

Acknowledgement: We would like to thank the Department of Pathology at İzmir Bakırçay University Çiğli Training and Research Hospital for providing the opportunity to conduct our study.

Ethics Committee Approval: This study was conducted in accordance with the ethical principles stated in the Declaration of Helsinki. The research protocol was reviewed and approved by the Non-Interventional Clinical Research Ethics Committee of İzmir Bakırçay University (Approval No: 1357; Date: 13.12.2023). All procedures were carried out in line with ethical standards, and patient confidentiality was strictly protected.

Informed Consent: Due to the retrospective nature of the study, the ethics committee waived the requirement for informed consent.

Peer-review: Externally peer-reviewed.

https://doi.org/10.5455/annalsmedres.2025.08.244

Conflict of Interest: The authors declare no conflict of interest.

Author Contributions: Concept/Design: SCK, AFD; Data Acquisition: HA; Data Analysis and Interpretation (including AI-based analyses): SCK, BK, HA, AG, AFD; Drafting Manuscript: SCK, BK, HA; Critical Revision of Manuscript: AFD, SCK, AG; Final Approval and Accountability: SCK, BK, HA, AG, AFD; Technical or Material Support: SCK, BK, HA, AG, AFD; Supervision: SCK, AFD.

Financial Disclosure: This study was supported by the İzmir Bakırçay University Scientific Research Projects Coordination Unit (Project No: BBAP.2023.014).

Artificial Intelligence Disclosure: The authors declare that an artificial intelligence–based tool (ChatGPT, version 5.2) was used solely for language editing to improve the clarity and readability of the manuscript. The AI tool did not contribute to the study design, data analysis, interpretation of results, or scientific conclusions. All content was critically reviewed, verified, and approved by the authors, who retain full responsibility for the manuscript.

REFERENCES

1. Sechopoulos I, Teuwen J, Mann R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Semin Cancer Biol.* 2021;72:214-225. doi: 10.1016/j.semcancer.2020.06.002.
2. Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: How special are they? *Mol Oncol.* 2010;4(3):192-208. doi: 10.1016/j.molonc.2010.04.004.
3. Tan PH, Ellis I, Allison K, et al. The 2019 World Health Organization classification of tumours of the breast. *Histopathology.* 2020;77(2):181-5. doi: 10.1111/his.14091.
4. Moghbel M, Ooi CY, Ismail N, Hau YW, Memari N. A review of breast boundary and pectoral muscle segmentation methods in computer-aided detection/diagnosis of breast mammography. Artificial Intelligence Review. 2020;53(3):1873-1918. doi: 10.1007/s10462-019-09721-8.
5. Moghbel M, Mashohor S. A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection. *Artificial Intelligence Review.* 2013;39(4):305-13. doi: 10.1007/s10462-011-9274-2.
6. Murtaza G, Shuib L, Wahab AWA, et al. Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. *Artificial Intelligence Review.* 2020;53(3):1655-720. doi: 10.1007/s10462-019-09716-5.
7. Domingues I, Pereira G, Martins P, Duarte H, Santos J, Abreu PH. Using deep learning techniques in medical imaging: a systematic review of applications on CT and PET. *Artificial Intelligence Review.* 2020;53:4093-160. doi: 10.1007/s10462-019-09788-3.
8. Kozegar E, Soryani M, Behnam H, Salamati M, Tan T. Computer aided detection in automated 3-D breast ultrasound images: a survey. *Artificial Intelligence Review.* 2020;53(3):1919-41. doi: 10.1007/s10462-019-09722-7.
9. Saha M, Chakraborty C, Racoceanu D. Efficient deep learning model for mitosis detection using breast histopathology images. *Comput Med Imaging Graph.* 2018;64:29-40. doi: 10.1016/j.compmedimag.2017.12.001.
10. Cheng HD, Cai XP, Chen XW, Hu LM, Lou XL. Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern Recognition.* 2003;36(12):2967-91. doi: 10.1016/S0031-3203(03)00192-4.
11. Cheng HD, Shi XJ, Min R, Hu LM, Cai XR, Du HN. Approaches for automated detection and classification of masses in mammograms. *Pattern Recognition.* 2006;39(4):646-68. doi: 10.1016/j.patcog.2005.07.006.
12. Suh YJ, Jung J, Cho BJ. Automated Breast Cancer Detection in Digital Mammograms of Various Densities via Deep Learning. *J Pers Med.* 2020;10(4):211. doi: 10.3390/jpm10040211.
13. Rindi G. Nomenclature and classification of neuroendocrine neoplasms of the digestive system. WHO classification of tumours of the digestive system. 2010:13-4.
14. Yamaguchi T, Fujimori T, Tomita S, et al. Clinical validation of the gastrointestinal NET grading system: Ki67 index criteria of the WHO 2010 classification is appropriate to predict metastasis or recurrence. *Diagn Pathol.* 2013;8:65. doi: 10.1186/1746-1596-8-65.
15. Tellez D, Balkenhol M, Otte-Höller I, et al. Whole-Slide Mitosis Detection in H&E Breast Histology Using PHH3 as a Reference to Train Distilled Stain-Invariant Convolutional Networks. *IEEE Transactions on Medical Imaging.* 2018;37(9):2126-36. doi: 10.1109/TMI.2018.2820199.
16. Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers. *Cytometry A.* 2017;91(6):566-73. doi: 10.1002/cyto.a.23065.
17. Lu C, Romo-Bucheli D, Wang X, et al. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab Invest.* 2018;98(11):1438-48. doi: 10.1038/s41374-018-0095-7.
18. Kuş Silav Z. Meme Kanserinde Ki67 İndeks Ölçümlerinin Manuel ve Dijital Yöntemler Açısından Kıyaslanması. *İstanbul Gelişim Üniversitesi Sağlık Bilimleri Dergisi.* 2023.
19. Biswal A. Top 10 deep learning algorithms You should Know in 2021. https://www simplilearn com/tutorials/deep-learning-tutorial/deep-learning-algorithm, accessed. 2021;5:26.
20. Datwani S, Khan H, Niazi MKK, Parwani AV, Li Z. Artificial intelligence in breast pathology: Overview and recent updates. *Hum Pathol.* 2025;162:105819. doi: 10.1016/j.humpath.2025.105819.
21. Jiang B, Bao L, He S, Chen X, Jin Z, Ye Y. Deep learning applications in breast cancer histopathological imaging: diagnosis, treatment, and prognosis. *Breast Cancer Res.* 2024;26(1):137. doi: 10.1186/s13058-024-01895-6.
22. Soliman A, Li Z, Parwani AV. Artificial intelligence's impact on breast cancer pathology: a literature review. *Diagn Pathol.* 2024;19(1):38. doi: 10.1186/s13000-024-01453-w.
23. Valieris R, Martins L, Defelicibus A, et al. Weakly-supervised deep learning models enable HER2-low prediction from H &E stained slides. *Breast Cancer Res.* 2024;26(1):124. doi: 10.1186/s13058-024-01863-0.
24. Chaurasia AK, Toohey PW, Bennett MT, Harris HC, Hewitt AW. Automated quantification of Ki-67 expression in breast cancer from H&E-stained slides using a transformer-based regression model. *Breast Cancer Res.* 2025;27(1):206. doi: 10.1186/s13058-025-02149-9.
25. Goyal M, Marotti JD, Workman AA, et al. A multi-model approach integrating whole-slide imaging and clinicopathologic features to predict breast cancer recurrence risk. *NPJ Breast Cancer.* 2024;10(1):93. doi: 10.1038/s41523-024-00700-z.
26. Mehta N, Devarakonda MV. Machine learning, natural language programming, and electronic health records: The next step in the artificial intelligence journey? *J Allergy Clin Immunol.* 2018;141(6):2019-21.e1. doi: 10.1016/j.jaci.2018.02.025.
27. Li S, Deng YQ, Zhu ZL, Hua HL, Tao ZZ. A Comprehensive Review on Radiomics and Deep Learning for Nasopharyngeal Carcinoma Imaging. *Diagnostics (Basel).* 2021;11(9):1523. doi: 10.3390/diagnostics11091523.
28. Castiglioni I, Rundo L, Codari M, et al. AI applications to medical images: From machine learning to deep learning. *Phys Med.* 2021;83:9-24. doi: 10.1016/j.ejmp.2021.02.006.
29. Şensu S, Erdoğan N, Gürbüz YS. Patolojide dijital çağ ve yapay Zekâ: temel bilgiler. Türkiye Klinikleri Tip Bilimleri Dergisi. 2020;40(1):104-12. doi: 10.5336/medsci.2019-72835.

https://doi.org/10.5455/annalsmedres.2025.08.244

30. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60-88. doi: 10.1016/j.media.2017.07.005.
31. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods.* 2021;18(1):100-6. doi: 10.1038/s41592-020-01018-x.
32. Polley M-YC, Leung SC, McShane LM, et al. An international Ki67 reproducibility study. *J Natl Cancer Inst.* 2013;105(24):1897-906. doi: 10.1093/jnci/djt306.
33. Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst.* 2011;103(22):1656-64. doi: 10.1093/jnci/djr393.
34. Veta M, Heng YJ, Stathonikos N, et al. Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge. *Med Image Anal.* 2019;54:111-21. doi: 10.1016/j.media.2019.05.008.
35. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J, editors. Mitosis detection in breast cancer histology images with deep neural networks. International conference on medical image computing and computer-assisted intervention; 2013: Springer.
36. Matsumoto H, Miyata R, Tsuruta Y, et al. Ki-67 evaluation using deep-learning model-assisted digital image analysis in breast cancer. *Histopathology.* 2025;86(3):460-71. doi: 10.1111/his.15356.
37. Dy A, Nguyen NNJ, Meyer J, et al. AI improves accuracy, agreement and efficiency of pathologists for Ki67 assessments in breast cancer. *Sci Rep.* 2024;14(1):1283. doi: 10.1038/s41598-024-51723-2.
38. Graham S, Vu QD, Raza SEA, et al. Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med Image Anal.* 2019;58:101563. doi: 10.1016/j.media.2019.101563.
39. Trullo R, Bui Q-A, Tang Q, Olfati-Saber R, editors. Image translation based nuclei segmentation for immunohistochemistry images. MICCAI Workshop on Deep Generative Models. 2022. doi: 10.48550/arXiv.2208.06202.
40. Gonzalez RC. Digital image processing: Pearson education india; 2009.

https://doi.org/10.5455/annalsmedres.2025.08.244

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Effect of adding bone marrow aspirate injection to core decompression on clinical outcomes in the treatment of femoral head avascular necrosis

Murat Taşçı [a],*, Mustafa Murat Hakyoldaş [a]

[a]*University of Health Sciences, Ümraniye Training and Research Hospital, Department of Orthopaedics and Traumatology, Istanbul, Türkiye*

**Corresponding author: drtascimurat@gmail.com (Murat Taşçı)*

MAIN POINTS

- This study compares clinical and radiological outcomes of core decompression (CD) alone versus CD combined with bone marrow aspirate concentrate (BMAC) injection in patients with avascular necrosis of the femoral head (ANFH).
- BMAC group showed a higher proportion of patients with improved Harris Hip Scores and decreased Visual Analog Scale scores at 1-year follow-up.
- Although differences were not statistically significant, BMAC injection demonstrated a trend toward better joint preservation and functional outcomes.
- The procedure is simple, safe, and may be recommended as a routine adjunct to CD in selected ANFH patients.

Cite this article as: Taçcı M, Hakyoldaş MM. Effect of adding bone marrow aspirate injection to core decompression on clinical outcomes in the treatment of femoral head avascular necrosis. *Ann Med Res*. 2026;33(1):43–47. doi: 10.5455/annalsmedres.2025.09.275.

ABSTRACT

Aim: Avascular necrosis of the femoral head (ANFH) is a progressive condition that often affects young individuals and may ultimately require total hip arthroplasty (THA). Core decompression (CD) is a widely used joint-preserving surgical technique, and bone marrow aspirate concentrate (BMAC) injection has been proposed as an adjunct to enhance outcomes by promoting vascularization and cartilage repair.

To compare the clinical and radiological outcomes of patients treated with CD alone versus CD combined with BMAC injection for ANFH.

Materials and Methods: Patients diagnosed with ANFH who underwent CD with or without BMAC injection were retrospectively analyzed. Demographic characteristics, preoperative and 1-year Harris Hip Score (HHS), Visual Analog Scale (VAS) score, Steinberg stage, and the need for THA during follow-up were recorded. All diagnoses were confirmed by MRI and radiography.

Results: A total of 42 hips (30 males, 12 females; mean age 41.8 years, range 14–66) were analyzed. BMAC injection was performed in 28 hips (66.7%) and not performed in 14 hips (33.3%). At 1-year follow-up, Steinberg stage progression was observed in 73.8% of hips, while 26.2% showed regression. THA was required in 14.3% of patients overall—7% in the BMAC group versus 28% in the CD-only group ($p = 0.1$). VAS scores decreased in 32% of BMAC-treated hips versus 21% in CD-only hips ($p>0.05$). Improvement in HHS was seen in 60% of the BMAC group versus 50% of the CD-only group ($p>0.05$).

Conclusion: Although statistical significance was not reached, the addition of BMAC to core decompression showed a trend toward reduced THA conversion and improved hip function. These findings suggest a potential benefit that warrants further investigation in larger, prospective studies before routine use can be recommended.

Keywords: Avascular necrosis, Femoral head, Core decompression, Bone marrow aspirate concentrate, Total hip arthroplasty, Hip preservation

Received: Sep 20, 2025 **Accepted:** Oct 31, 2025 **Available Online:** Jan 26, 2026

INTRODUCTION

Avascular necrosis of the femoral head (ANFH), which develops due to impaired circulation in the femoral head, is a common condition in orthopedics [1]. This disease, frequently encountered in young individuals, occurs as a result of disrupted blood supply to the femoral head, leading to collapse of the articular cartilage, hip pain, and restricted joint motion [2]. Although its exact etiology has not been clearly established, total hip arthroplasty (THA) has shown successful outcomes in treating the condition after articular cartilage deterioration. However, due to the relatively young age of most patients and the potential need for revision surgeries after THA, joint-preserving surgical interventions, as well as conservative approaches such as physical therapy and weight-bearing restrictions, are also widely employed [3, 4].

Among joint-preserving surgeries, core decompression (CD) is the most frequently performed technique, aiming to reduce intraosseous pressure and improve the blood supply to

the articular cartilage. While successful outcomes have been achieved in early stages, CD alone is often insufficient to prevent progression to THA, leading to the use of additional procedures to preserve joint function. Various treatment modalities, including vascularized fibular grafting (as a standalone surgical approach), stem cell therapies, and pharmacologic adjuncts such as bisphosphonates, have also been explored for joint-preserving treatment in early-stage ANFH [4,5,6]. Previous studies have demonstrated a reduction in osteoprogenitor cells in the pathogenesis of ANFH, and it has been suggested that stem cells may contribute to recovery by differentiating into both vascular structures and chondroblasts, thereby improving femoral head pathology [5, 7].

However, despite promising early results, the magnitude of clinical benefit from adding BMAC to CD remains controversial due to variability in techniques, disease stages, and follow-up durations. Therefore, further studies are needed to determine whether this combined approach significantly improves outcomes compared to CD alone.

The aim of this study was to clinically and radiologically compare the outcomes of patients treated with CD alone and those treated with CD combined with autologous BMAC injection, which is an adjunctive procedure frequently performed in our clinic for ANFH.

We hypothesized that the addition of BMAC to CD would enhance hip function, reduce pain, and decrease the rate of progression to total hip arthroplasty compared with CD alone.

MATERIALS AND METHODS

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Institutional Ethics Committee (Approval No: B.10.1.TKH.4.34.H.GP.0.01/226, Date: 10.07.2025). As this was a retrospective analysis of routinely performed surgical procedures, a specific informed consent form was not required; the standard preoperative surgical consent obtained from all patients was considered sufficient.

This study was designed as a retrospective comparative study conducted in a single tertiary orthopedic center. This retrospective comparative study was conducted and reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines. The medical records and radiological data of patients who underwent core decompression (CD) with or without autologous bone marrow aspirate concentrate (BMAC) injection for avascular necrosis of the femoral head (ANFH) were reviewed and analyzed.

A non-probability (convenience) sampling method was used. All patients who met the inclusion criteria and had complete clinical and radiological follow-up data were included in the study.

Inclusion criteria

Patients diagnosed with avascular necrosis of the femoral head (ANFH) who underwent core decompression (CD) with or without autologous bone marrow aspirate concentrate (BMAC) injection and had a minimum follow-up of 1 year were included in the study.

Exclusion criteria

- Patients with incomplete clinical or radiological data,
- Patients with less than 1 year of follow-up,
- Those with secondary osteonecrosis due to infection, or malignancy,

Patients with advanced femoral head collapse (beyond Steinberg stage IV) at presentation.

No randomization method was used in this study. Patients were included based on the inclusion criteria, and grouping was determined according to whether bone marrow aspirate concentrate (BMAC) injection was performed. The decision to perform BMAC injection was made at the discretion of the operating surgeon during surgery. Since this was not an interventional study but rather a retrospective analysis of clinical outcomes, the grouping was considered independent of treatment effect bias.

As this was a retrospective study based on the review of patient records and radiological data, no blinding or masking method was used. The surgeons performing the procedures and the researchers analyzing the data were aware of the treatment groups (CD only vs. CD + BMAC). A prior sample size calculation was not performed. All eligible patients who met the inclusion criteria and had a minimum of 1-year follow-up were included in the analysis.

Demographic data, preoperative and 1-year postoperative Steinberg stage, Harris Hip Score (HHS), Visual Analog Scale (VAS) score, and whether total hip arthroplasty (THA) was performed during follow-up were recorded. The diagnosis in all patients was confirmed by magnetic resonance imaging (MRI) and direct radiography, and differential diagnoses were excluded. All measurements were performed together with the senior author to minimize errors. Harris Hip Scores and VAS assessments were recorded preoperatively and repeated at the 1-year follow-up.

The primary outcome parameters that were investigated included pre- and postoperative changes in HHS and VAS scores, while the secondary outcome was the change in Steinberg stage on radiological evaluation.

Surgical procedure

In all patients, the necrotic area of the femoral head was localized preoperatively using MRI. With the patient in the supine position, under fluoroscopic guidance, multiple drill holes

https://doi.org/10.5455/annalsmedres.2025.09.275

were created using a 3 mm K-wire. Anteroposterior and lateral fluoroscopic views were used to confirm that the preoperatively identified target region had been reached. In patients undergoing BMAC injection in addition to CD, bone marrow was aspirated from the iliac crest using a bone marrow biopsy needle and syringe. The surgeon advanced the needle through the cortex into the medullary cavity, a step that can be clearly felt once the cortex is penetrated. Approximately 15–20 mL of bone marrow aspirate was obtained for each hip. The aspirate was injected immediately into the decompression tract under fluoroscopic control without delay to prevent clot formation and maintain the viability of the concentrate. The syringe was held in position for about two minutes to allow partial coagulation within the tract and minimize backflow or leakage. No special technique was used to prevent BMAC discharge, and no leakage was observed during the procedure. All patients were discharged after 1 day of postoperative observation in the ward. The decision to perform BMAC injection was entirely at the discretion of the operating surgeon, and patient selection for the procedure was also based on the surgeon's preference.

Statistical analysis

Data were analyzed using SPSS software (ver. 22.0; IBM Corp., Armonk, NY, USA). Quantitative variables are expressed as mean ± standard deviation and minimum and maximum values. Qualitative variables are expressed as frequencies or ratios. Continuous variables were analyzed using the Student's t-test when normally distributed and the Mann–Whitney U test when non-normally distributed.The Chi-square test was used to assess the qualitative parameters. P-values < 0.05 were considered to indicate statistical significance.

RESULTS

A total of 42 hips were included in the study. The mean patient age was 41.8 years (range: 14–66). The operated side was the right hip in 24 patients (57.1%) and the left hip in 18 patients (42.9%). Of the patients, 30 (71.4%) were male and 12 (28.6%) were female. BMAC injection was not performed in 14 hips (33.3%), whereas it was performed in 28 hips (66.7%). The mean preoperative VAS score was 4.6 (range: 2–6), while the mean VAS score at 1 year was 5.1 (range: 2–10). The mean preoperative HHS was 67.6 (range: 16–93), while the mean HHS at 1 year was 63.2 (range: 23–97). At 1-year follow-up, Steinberg stage had progressed in 31 patients (73.8%) and regressed in 11 patients (26.2%). THA was performed in 6 patients (14.3%) during postoperative follow-up due to persistent symptoms and development of joint arthrosis. VAS scores decreased in 12 patients (28.6%), increased in 14 patients (33.3%), and remained unchanged in 16 patients (38.1%) at 1 year. HHS increased in 24 patients (57.1%) and decreased in 18 patients (42.9%) at 1 year (Table 1).

Table 1. Demographic characteristics of the patients included in the study.

		N (%)
Age		41.8
Gender	Female	12 (28%)
	Male	30 (72%)
Treatment	CD	14 (33%)
	CD and BMAI	28 (67%)
		Mean
Age		41.8
Preoperative VAS	CD	4.5
	BMAI	4.7
VAS in the 1st year	CD	5.4
	BMAI	5
Preoperative Harris Hip Score	CD	67.1
	BMAI	67.9
Harris Hip Score in the 1st year	CD	63
	BMAI	63.3

CD: Core decompression; BMAI: Bone Marrow Aspirate Injection.

Table 2. Comparison of demographic and clinical outcomes between the study groups.

Parameter	BMAC Group (n = 28)	Non-BMAC Group (n = 14)	p-value
Age (years)	45.1 (± 14.40)	40.1 (± 15.51)	>0.05
Steinberg stage decreased (%)	28%	21%	>0.05
Preoperative VAS score	4.7	4.5	>0.05
VAS score in the 1st year	5.0	5.4	>0.05
Improvement in the VAS (%)	32%	21%	>0.05

The mean age was 45.1 years in the BMAC group and 40.1 years in the non-BMAC group, with no significant demographic differences between the two. Improvement in Steinberg stage was observed in 28% of the BMAC group compared to 21% of the non-BMAC group. Mean VAS scores in the BMAC group changed from 4.7 preoperatively to 5.0 at 1 year, whereas in the non-BMAC group, they changed from 4.5 to 5.4. Improvement in VAS scores was noted in 32% of the BMAC group and 21% of the non-BMAC group ($p > 0.05$) (Table 2).

Regarding functional outcomes, the mean HHS decreased from 67.9 to 63.3 at 1 year in the BMAC group, and from 67.1 to 63.0 in the non-BMAC group. Although the difference between groups was not statistically significant, HHS improved in 60% of the BMAC group and 50% of the non-BMAC group at 1 year ($p>0.05$). Conversion to THA occurred in 7% of the BMAC group and 28% of the non-BMAC group ($p = 0.1$).

DISCUSSION

In the present study, we compared the outcomes of patients treated with CD alone versus those treated with CD com-

https://doi.org/10.5455/annalsmedres.2025.09.275

bined with BMAC injection for ANFH. Although statistical significance was not reached, the BMAC group demonstrated a trend toward greater improvement in HHS and VAS scores, as well as a lower rate of conversion to THA compared to the CD-only group.

ANFH is a serious orthopedic condition that predominantly affects young adults and frequently necessitates THA as the definitive treatment. Poor prognostic factors reported in the literature include male sex, prolonged symptom duration prior to treatment, lower HHS, and higher VAS scores [8]. Core decompression is the most common joint-preserving surgical procedure, and numerous clinical studies have investigated the use of stem cell therapy as an adjunct to CD [9]. Song et al. [10] reported survival rates of 79%, 77%, and 35% at a minimum 5-year follow-up for patients with Ficat stage I, II, and III disease, respectively, following CD.

Stem cell therapy has gained traction in regenerative medicine due to its potential to facilitate remodeling and repair [11]. When combined with CD, stem cell therapy has shown promise in improving graft survival. Gangji et al. [12] demonstrated that stem cell application improved hip function and slowed radiological progression. Similarly, a meta-analysis by Li et al. [13] reported that the stem cell group exhibited longer survival, slower radiological progression, and a 2.85-fold reduction in conversion to THA. Consistent with the literature, our study observed better outcomes regarding Steinberg stage and a lower rate of conversion to THA in the injection group, although these findings were not statistically significant. Notably, the BMAC used in this study was not processed to isolate or purify stem cells; therefore, the procedure represents a biological augmentation technique using unprocessed autologous bone marrow aspirate rather than a standardized stem cell therapy.

Regarding Harris Hip Scores, previous studies have indicated that lower preoperative values are associated with a negative impact on survival [8, 14]. In our study, the injection group showed greater improvement when comparing pre- and postoperative HHS changes. In patients who required conversion to THA, the mean preoperative VAS score was 5.0, and the mean preoperative HHS was 55.3. Furthermore, at the 1-year follow-up, VAS scores increased and HHS decreased in all these patients. Conversely, patients who did not undergo THA had lower mean preoperative VAS scores (4.6) and higher mean HHS (76.4). These findings align with Migliorini et al. [8], suggesting that preoperative VAS and HHS are critical prognostic indicators. While the BMAC group showed greater HHS improvement than the CD-only group, the difference did not reach statistical significance. Thus, while BMAC may provide a functional benefit, larger studies are required to confirm its clinical relevance.

Hauzeur et al. [15] compared changes in VAS scores between a stem cell injection group and a CD-only group and found no significant difference. Similarly, our study did not demonstrate a statistically significant reduction in VAS scores; however, a higher proportion of patients in the injection group experienced decreased pain. Despite the lack of statistical significance, we believe that BMAC injection may contribute to pain reduction.

Although the BMAC group demonstrated a trend toward reduced THA conversion and improved hip function, these differences were not statistically significant and should be interpreted with caution regarding clinical impact. Nevertheless, when CD is planned as a joint-preserving procedure, we recommend the routine addition of BMAC injection, as it presents no adverse effects and may positively contribute to hip preservation.

Limitations

The main limitations of this study include its retrospective design, relatively small sample size, and short follow-up period. Additionally, the lack of histological evaluation and the absence of a standardized protocol for BMAC preparation may limit the generalizability of the findings. Further prospective studies with larger patient cohorts and longer follow-up periods are needed to validate these results.

CONCLUSION

While ANFH remains a serious health problem in young patients, the addition of BMAC injection to core decompression may contribute to improved joint preservation and a lower tendency toward conversion to THA. However, this finding did not reach statistical significance, and further studies with larger cohorts are warranted to confirm these results.

Ethics Committee Approval: This study was conducted in accordance with the Declaration of Helsinki and received approval from local Instutional Review Board (University of Health Sciences Umraniye Training and Research Hospital Scientific Research Ethics Committee, Approval No: B.10.1.TKH.4.34.H.GP.0.01/226, Date: 10.07.2025).

Informed Consent: Not necessary for this manuscript.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors declare that they have no competing interests.

Author Contributions: Concept: M.T; Design: M.M.H; Supervision: M.T; Materials: M.M.H; Data Collection and/or Processing: M.M.H; Analysis and/or Interpretation: M.T; Literature Review: M.T; Writing: M.T; Critical Review: M.T, M.M.H. All authors critically revised the manuscript, agree to be fully accountable for ensuring the integrity and accuracy of the work, and read and approved the final manuscript.

Financial Disclosure: No funding has been received for this study.

Artificial Intelligence Disclosure: The authors declare that an artificial intelligence–based tool (ChatGPT, OpenAI, version

https://doi.org/10.5455/annalsmedres.2025.09.275

[4o]) was used solely for language editing and improving the clarity and readability of the manuscript. The AI tool did not contribute to the study design, data analysis, interpretation of results, or scientific conclusions. All content was critically reviewed and approved by the authors, who take full responsibility for the manuscript.

REFERENCES

1. Zhao D, Zhang F, Wang B, Liu B, Li L, et al. Guidelines for clinical diagnosis and treatment of osteonecrosis of the femoral head in adults (2019 version). *J Orthop Translat.* 2020;21:100-110. doi: 10.1016/j.jot.2019.12.004.
2. Hua KC, Yang XG, Feng JT, Wang F, Yang L, et al. The efficacy and safety of core decompression for the treatment of femoral head necrosis: a systematic review and meta-analysis. *J Orthop Surg Res.* 2019;14(1):306. doi: 10.1186/s13018-019-1359-7.
3. Xu Y, Jiang Y, Xia C, Wang Y, Zhao Z, Li T. Stem cell therapy for osteonecrosis of femoral head: Opportunities and challenges. *Regen Ther.* 2020;15:295-304. doi: 10.1016/j.reth.2020.11.003.
4. Migliorini F, La Padula G, Oliva F, Torsiello E, Hildebrand F, Maffulli N. Operative Management of Avascular Necrosis of the Femoral Head in Skeletally Immature Patients: A Systematic Review. *Life (Basel).* 2022;12(2):179. doi: 10.3390/life12020179.
5. Hernigou P, Bernaudin F, Reinert P, Kuentz M, Vernant JP. Bone-marrow transplantation in sickle-cell disease. Effect on osteonecrosis: a case report with a four-year follow-up. *J Bone Joint Surg Am.* 1997;79(11):1726-30. doi: 10.2106/00004623-199711000-00015.
6. Hadighi P, Faezi ST, Mortazavi SMJ, Rokni M, Aghaghazvini L, et al. Evaluation of core decompression outcome in systemic lupus erythematosus with hip osteonecrosis: a retrospective cohort study. *Adv Rheumatol.* 2024;64(1):4. doi: 10.1186/s42358-023-00345-9.
7. Emadedin M, Karimi S, Karimi A, Labibzadeh N, Niknejadi M, et al. Autologous bone marrow-derived CD133 cells with core decompression as a novel treatment method for femoral head osteonecrosis: a pilot study. *Cytotherapy.* 2019;21(1):107-112. doi: 10.1016/j.jcyt.2018.10.005.
8. Migliorini F, Maffulli N, Baroncini A, Eschweiler J, Tingart M, Betsch M. Prognostic factors in the management of osteonecrosis of the femoral head: A systematic review. *Surgeon.* 2023;21(2):85-98. doi: 10.1016/j.surge.2021.12.004.
9. Piuzzi NS, Chahla J, Jiandong H, Chughtai M, LaPrade RF, et al. Analysis of Cell Therapies Used in Clinical Trials for the Treatment of Osteonecrosis of the Femoral Head: A Systematic Review of the Literature. *J Arthroplasty.* 2017;32(8):2612-2618. doi: 10.1016/j.arth.2017.02.075.
10. Song WS, Yoo JJ, Kim YM, Kim HJ. Results of multiple drilling compared with those of conventional methods of core decompression. *Clin Orthop Relat Res.* 2007;454:139-46. doi: 10.1097/01.blo.0000229342.96103.73.
11. Goodman SB. The biological basis for concentrated iliac crest aspirate to enhance core decompression in the treatment of osteonecrosis. *Int Orthop.* 2018;42(7):1705-1709. doi: 10.1007/s00264-018-3830-1.
12. Gangji V, De Maertelaer V, Hauzeur JP. Autologous bone marrow cell implantation in the treatment of non-traumatic osteonecrosis of the femoral head: Five year follow-up of a prospective controlled study. *Bone.* 2011;49(5):1005-9. doi: 10.1016/j.bone.2011.07.032.
13. Li M, Chen D, Ma Y, Zheng M, Zheng Q. Stem cell therapy combined with core decompression versus core decompression alone in the treatment of avascular necrosis of the femoral head: a systematic review and meta-analysis. *J Orthop Surg Res.* 2023;18(1):560. doi: 10.1186/s13018-023-04025-8.
14. Martinot P, Dartus J, Justo A, Riouach H, Cremer P, et al. French Society of Orthopaedic Surgery and Traumatology (SoFCOT). Does augmented core decompression decrease the rate of collapse and improve survival of femoral head avascular necrosis? Case-control study comparing 184 augmented core decompressions to 79 standard core decompressions with a minimum 2 years' follow-up. *Orthop Traumatol Surg Res.* 2020;106(8):1561-1568. doi: 10.1016/j.otsr.2020.03.040.
15. Hauzeur JP, De Maertelaer V, Baudoux E, Malaise M, Beguin Y, Gangji V. Inefficacy of autologous bone marrow concentrate in stage three osteonecrosis: a randomized controlled double-blind trial. *Int Orthop.* 2018;42(7):1429-1435. doi: 10.1007/s00264-017-3650-8.

https://doi.org/10.5455/annalsmedres.2025.09.275