E-ISSN: 2636-7688

year 2026
volume 33
issue 6

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: 6 | June 2026

Ann Med Res

Annals of Medical Research

The Official Journal of Inonu University Faculty of Medicine

Editorial Advisory Board | Volume: 33 | Issue: 6 | June 2026

Ann Med Res

Annals of Medical Research

The Official Journal of Inonu University Faculty of Medicine

Table of Contents | Volume: 33 | Issue: 6 | June 2026

Original Articles

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Investigation of the anticancer effect of Ramelteon on different human cancers: An in vitro study

Tuba Keskin[a], Guldeniz Sekerci[a], Senanur Ilikca[a], Cigdem Tekin[b], Suat Tekin[a],*

[a]*Inonu University, Faculty of Medicine, Department of Physiology, Malatya, Türkiye*
[b]*Inonu University, Vocational School of Health Services, Department of Home Care, Malatya, Türkiye*
**Corresponding author: suat.tekin@inonu.edu.tr (Suat Tekin)*

MAIN POINTS

- Antiproliferative effect of Ramelteon: It reduced cell viability in tumour cell lines (A2780, MCF-7, Caco-2, LNCaP).
- Ramelteon exhibited anti-genotoxic properties by suppressing DNA damage.
- Clinical potential: The demonstration of anti-cytotoxic and anti-genotoxic effects of Ramelteon across different tumor cell lines suggests that it may be a candidate for cancer therapy; however, advanced in vitro and in vivo studies are required to elucidate the underlying mechanisms.

Cite this article as: Keskin T, Sekerci G, Ilikca S, Tekin C, Tekin S. Investigation of the anticancer effect of Ramelteon on different human cancers: An in vitro study. *Ann Med Res*. 2026;33(6):245-252. doi: 10.5455/annalsmedres.2025.10.293.

ABSTRACT

Aim: To describe the various methods used to treat cancer, one of the diseases with the highest mortality rates worldwide. Among these methods, the use of agents exhibiting antioxidant activity has increased significantly. Ramelteon, a melatonin receptor agonist, stands out for its antioxidant and anti-inflammatory properties. In this study, we investigated the cytotoxic and genotoxic effects of Ramelteon in different human cancer cell lines.

Materials and Methods: In the study, Ramelteon was applied at concentrations of 5, 10, 25, 50, 100, and 500 µM to human ovarian (A2780), breast (MCF-7), colon (Caco-2), and prostate (LNCaP) cancer cell lines, which were exposed and incubated for 24 hours. The MTT assay was used to measure changes in viability. After obtaining data on dose-dependent effects, inhibitory concentration values were determined. Genotoxicity was examined using the Comet assay at the determined effective dose. Intergroup comparisons were performed using the Kruskal-Wallis H test. When statistically significant differences were found between groups, multiple comparisons were performed using the Bonferroni-corrected Mann-Whitney U tests ($p<0.05$).

Results: Ramelteon significantly reduced cell viability ($p<0.05$) and exhibited genotoxic effects in the A2780, MCF-7, Caco-2, and LNCaP cell lines ($p<0.05$).

Conclusion: Research findings indicate that Ramelteon exhibits anticancer potential in A2780, MCF-7, Caco-2, and LNCaP cell lines, suggesting that the drug may be considered a promising candidate for the development of new cancer treatment strategies.

Keywords: Colon cancer, Melatonin, Breast cancer, Prostate cancer, Ramelteon, Ovarian cancer

Received: Oct 01, 2025 **Accepted:** Dec 29, 2025 **Available Online:** Jun 25, 2026

INTRODUCTION

When a cell reaches a certain level of maturity, it divides to promote growth and development. When this process proceeds normally, the organism remains healthy; however, if control of cell division is lost for various reasons, cancer can develop [1]. Results indicate that anti-inflammatory and antioxidant compounds used to reduce inflammation closely associated with carcinogenesis promote programmed cell death (apoptosis) [2-4]. Therefore, the search for effective cancer treatments has led to speculation about the therapeutic potential of certain agents with anti-inflammatory properties. Investigating the anticancer potential of pharmacological agents currently in clinical use offers significant strategic and translational advantages, particularly with regard to compounds exhibiting anti-inflammatory and antioxidant properties. The pharmacokinetic and safety profiles of these agents are well understood; thus, safety concerns can be significantly reduced in early-phase clinical trials [5]. Because chronic inflammation and oxidative stress drive tumor growth and spread (formation, progression, and metastasis), these drugs may fundamentally affect the core determinants of anticancer biology [6].

Melatonin, which has demonstrated anti-inflammatory and antioxidant effects and is produced endogenously, has been studied extensively and shown to exert anticancer effects in multiple reports. Other findings indicate that melatonin promotes autophagy and induces apoptosis, in addition to exhibiting anticancer effects [7-9]. Currently, exogenous forms

of melatonin are used as receptor agonists in pharmacological treatments. The most important characteristics of these pharmacological agents that distinguish them from melatonin are their higher receptor affinity, their greater absorption, and their longer half-life [10]. Ramelteon, a tricyclic synthetic analog of melatonin, is used as a hypnotic agent for the treatment of insomnia. Approved for use in 2005, this agent has a higher affinity for melatonin receptor types 1a (MT1) and 1b (MT2) than melatonin [11]. Short- to intermediate-acting Ramelteon binds to MT1 and MT2 receptors in the suprachiasmatic nucleus, the central circadian pacemaker [12]. Other studies indicate that its half-life is approximately 1–2.6 hours, which is longer than melatonin's half-life of 30 minutes [13, 14]. Research findings indicate that Ramelteon also contributes to memory development by increasing neuronal activity during wakefulness [15]. In addition, Ramelteon has been shown to reduce postoperative delirium, which is commonly encountered in elderly patients with esophageal cancer [16]. Published studies have shown that Ramelteon reduces the viability of neuroblastoma and prostate cancer cell lines by 50% [17]. Another study also suggests that Ramelteon, with its melatonin-like activity, has the potential to be a candidate for the treatment of recurrent endometrial cancer [18].

Although Ramelteon has been reported to be clinically reliable and to provide indirect benefits in oncology practice, and the literature indicates its efficacy in certain cancer types, a detailed review of these studies reveals that its effects on other cancer types have not yet been investigated. This study investigates the effects of Ramelteon on several cancer types, including ovarian, breast, colon, and prostate cancer.

In this context, we aimed to investigate the effect of Ramelteon on ovarian, breast, colon, and prostate cancer cell lines.

MATERIALS AND METHODS

Cell culture

The present study was conducted in the Department of Physiology, Faculty of Medicine, Inonu University. Ovarian (A2780), breast (MCF-7), colorectal (Caco-2), and prostate (LNCaP) cancer cell lines were utilized in this study. The cells were multiplied under suitable conditions [19]. Cultures were maintained at 37°C with 5% CO_2 in an incubator (Esco, Singapore) and were provided with fresh media every three days. When the cells became confluent, they were harvested from the flasks using trypsin-EDTA and stained with 0.4% trypan blue. The number of viable cells was determined by counting using an inverted microscope, and the experiment was initiated only when cell viability was 90% or greater [20, 21]. For the cell viability experiments, 1.5×10^5 cells were plated into a 12 × 8–well plate. Following a 24-hour incubation period, the cells within the plate were exposed to culture medium, dimethyl sulfoxide (DMSO), and Ramelteon. Post hoc power analysis indicated a calculated power [1-beta] of approximately 1, with a type I error (alpha) of 0.05, a sample size of 10, and an effect size of 7.7 for cell viability.

MTT assay method

Following cell seeding, Ramelteon was prepared in 1% DMSO and added to the wells at final concentrations of 5 µM, 10 µM, 25 µM, 50 µM, 100 µM, and 500 µM [17, 18]. The control group received only the cells' own medium, while the solvent group received the same medium containing 1% DMSO. The cells were then incubated for an additional 24 hours, and cell proliferation was determined using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT). MTT penetrates the cell membrane of viable cells and, after various enzymatic reactions, is reduced to form a blue-violet product [22, 23]. The MTT assay is employed for cell viability testing precisely because this reaction is restricted to metabolically active cells. Spectrophotometric measurement of the resultant color allows correlation of absorbance with the number of living cells. For analysis, MTT was dissolved in phosphate buffer of 0.5 mg/mL and sterilized via filter. Cell culture media and Ramelteon were removed from the plates, and 50 µL of MTT solution was added to the wells, which were then incubated for 3 hours. After incubation, the MTT solution was removed from the wells, and 100 µL of DMSO was added [24, 25]. The light intensity from the cells in the wells was measured at a wavelength of 550 nm using a BioTek Synergy HTX plate reader (Netherlands) [26]. The average absorbance values obtained from the control wells were accepted as 100% viability, and the results were obtained by comparing the absorbance value obtained from the control group with the average [27]. The absorbance values obtained from the ramelteon and solvent groups were also compared with the control group to calculate the % cell viability [25]. The experimental design of the study is shown in Figure 1. These study steps were repeated at least 10 times on different occasions. After the effects of Ramelteon on cell viability were determined by the MTT assay, the inhibitory concentration 50 (IC_{50}) values of these compounds were calculated using GraphPad Prism 8 software and subsequently used for the Comet assay.

Single-cell gel electrophoresis comet assay methods

Comet analysis, also known as single-cell gel electrophoresis, is widely used to assess DNA damage (genotoxicity) [26]. Comet analysis was performed using the method developed by Devlin et al. [27]. In the first step, microscope slides were coated with 0.65% high-melting-point agarose prepared in phosphate buffer and then incubated in the dark for one day. Ramelteon at the determined IC_{50} concentrations was prepared and used to incubate A2780, MCF-7, Caco-2, and LNCaP cells. The cells were then spread onto slides prepared with low-melting-point agarose, and the slides were covered with coverslips. The protocol began by storing the prepared slides in the dark at +4°C for 20–25 minutes. After removing the coverslips, the slides were lysed for 1 hour at +4 °C

https://doi.org/10.5455/annalsmedres.2025.10.293

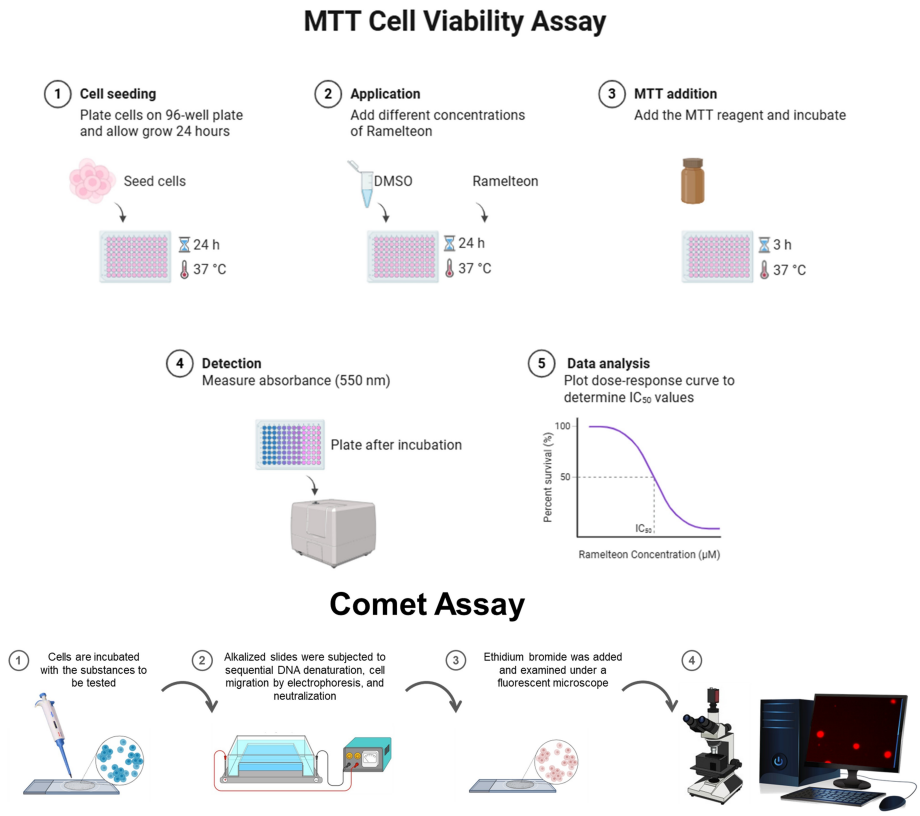

Figure 1. MTT and Comet Assay experimental design (DMSO: dimethyl sulfoxide, IC_{50}: Inhibitory concentration 50, MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)

in the dark using a cold lysis buffer (2.5 M NaCl, 100 mM EDTA, 10 mM Tris, pH 10) containing 1% Triton X-100 and 10% DMSO. Post-lysis, the slides were placed in a horizontal electrophoresis tank (Bio-Rad, USA) filled with cold electrophoresis buffer, and electrophoresis was performed in the dark. Neutralization was then carried out three times, for 5 minutes each, at +4 °C using neutralization buffer (0.4 M Tris, pH 7.5). The slides were then stained with 50 μL of ethidium bromide and incubated for 20–30 minutes. Imaging was performed using a Leica fluorescence microscope, and scoring was performed with Comet IV software. At least 25 cell images per slide were randomly selected for analysis, allowing determination of head intensity and length, tail intensity and length, and olive tail moment for each experimental group. Analyzing the changes in these five parameters was key to assessing the presence and degree of DNA damage [25]. Figure 1 shows the experimental design of the study.

Statistical analysis

Data analysis was performed using IBM SPSS Statistics for Windows, version 24.0 (Armonk, NY: IBM Corp.). Normality was confirmed by the Kolmogorov–Smirnov test. Based on the study results, comparisons between groups were conducted using a one-way ANOVA. When significant differences were identified among the groups, post hoc pairwise comparisons were performed using the Bonferroni test. All p-values < 0.05 were considered statistically significant. Based on the MTT assay results, the IC_{50} values of Ramelteon for all cell lines were determined using GraphPad Prism 8. Figures for the study were generated using the BioRender platform.

RESULTS

Investigation of the effects of Ramelteon on cell viability

Cell viability (%) of cancer cell lines treated with Ramelteon at different concentrations after 24 hours of incubation is shown in Figure 2 and Table 1. Exposure of A2780 cells to Ramelteon reduced cell viability compared with the control, and this effect was most pronounced at 50, 100, and 500 μM (p = 0.001) (Figure 2A). Similarly, 24-hour Ramelteon exposure significantly decreased cell viability in the MCF-7 cell line at all tested doses compared with the control (p = 0.001) (Figure 2B). As shown in Figure 2C, Ramelteon also reduced cell viability in the Caco-2 cell line at all doses except the 5 μM dose (p = 0.02). Furthermore, in LNCaP cells, a statistically significant reduction in viability was observed at the higher doses (100 and 500 μM) (p = 0.02) (Figure 2D). Additionally, numerical data related to cell viability are presented in Table 1. Table 2 summarizes the IC_{50} values determined by the MTT assay for A2780, MCF-7, Caco-2, and LNCaP cells after 24 hours of Ramelteon treatment. These IC_{50} findings collectively demonstrate that Ramelteon reduces cell viability by 50% in various cancer cell lines at the lowest tested concentration (Table 2).

https://doi.org/10.5455/annalsmedres.2025.10.293

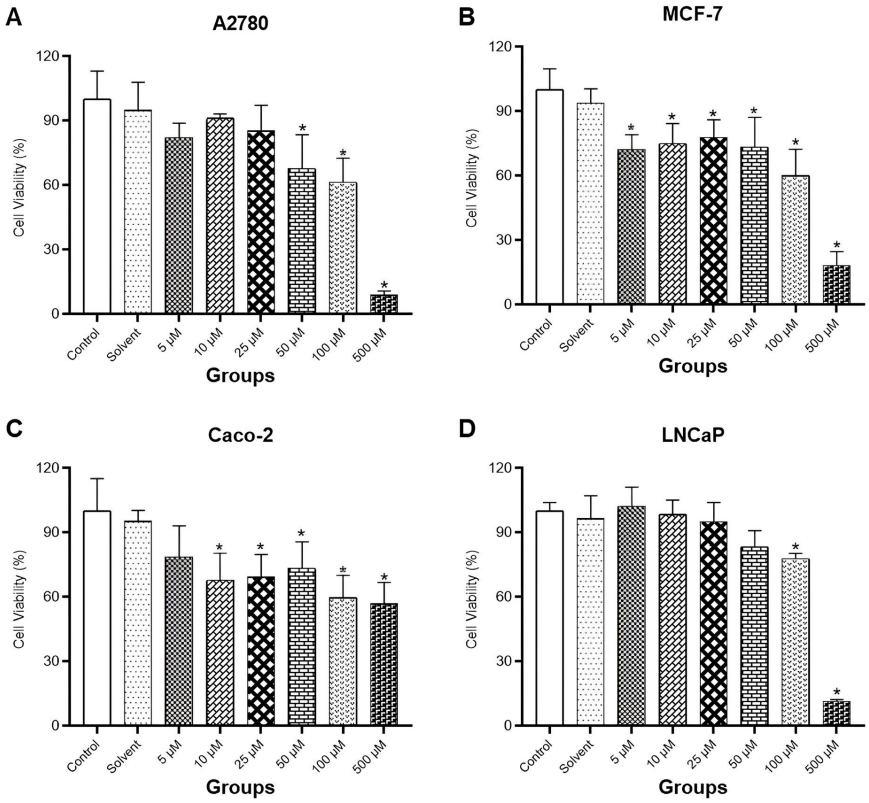

Figure 2. Dose-dependent cell viability results of A2780, MCF-7, Caco-2 and LNCaP cells after 24 h incubation of Ramelteon agents. Each data point is the average of 10 viability measurements. (*) shows p<0.05. (A, A2780 Cell Lines; B, MCF-7 Cell Lines; C, Caco-2 Cell Lines; D, LNCaP Cell Lines).

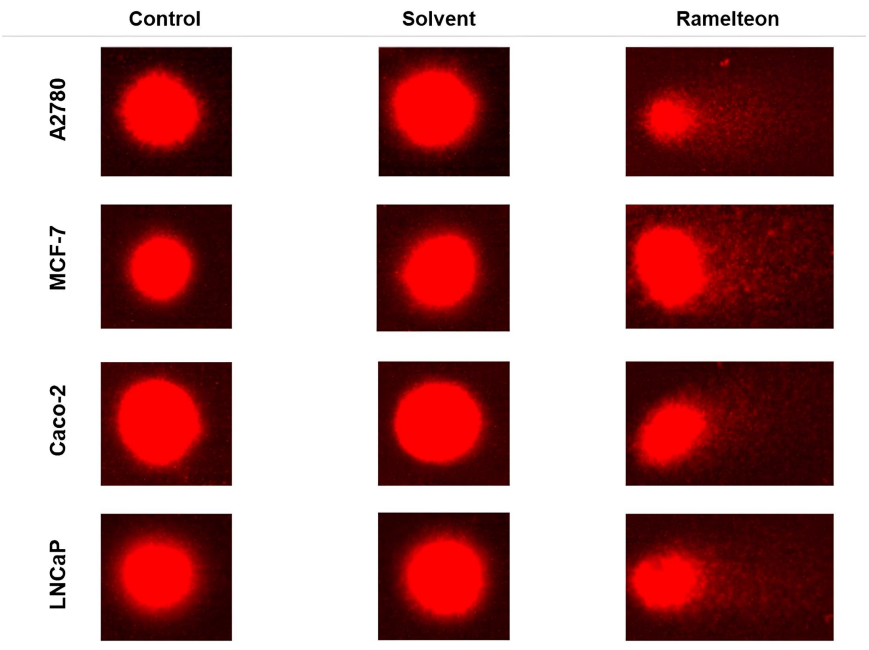

Figure 3. Comet assay images obtained from various cancer cells treated with Ramelteon.

Investigation of the effects of Ramelteon on DNA damage

The extent of DNA damage in cancer cell lines after Ramelteon treatment was quantified using the Comet assay, and the results are presented in Figures 3, 4 and Table 3. The microscopic data (Figure 3) provide qualitative evidence that Ramelteon induced comet formation in the cell lines. Comet assay results showed that Ramelteon significantly increased both head intensity (Figure 4A) and head length (Figure 4B) in A2780, MCF-7, Caco-2, and LNCaP cells ($p = 0.004$). Interestingly, the critical DNA damage parameters tail intensity

https://doi.org/10.5455/annalsmedres.2025.10.293

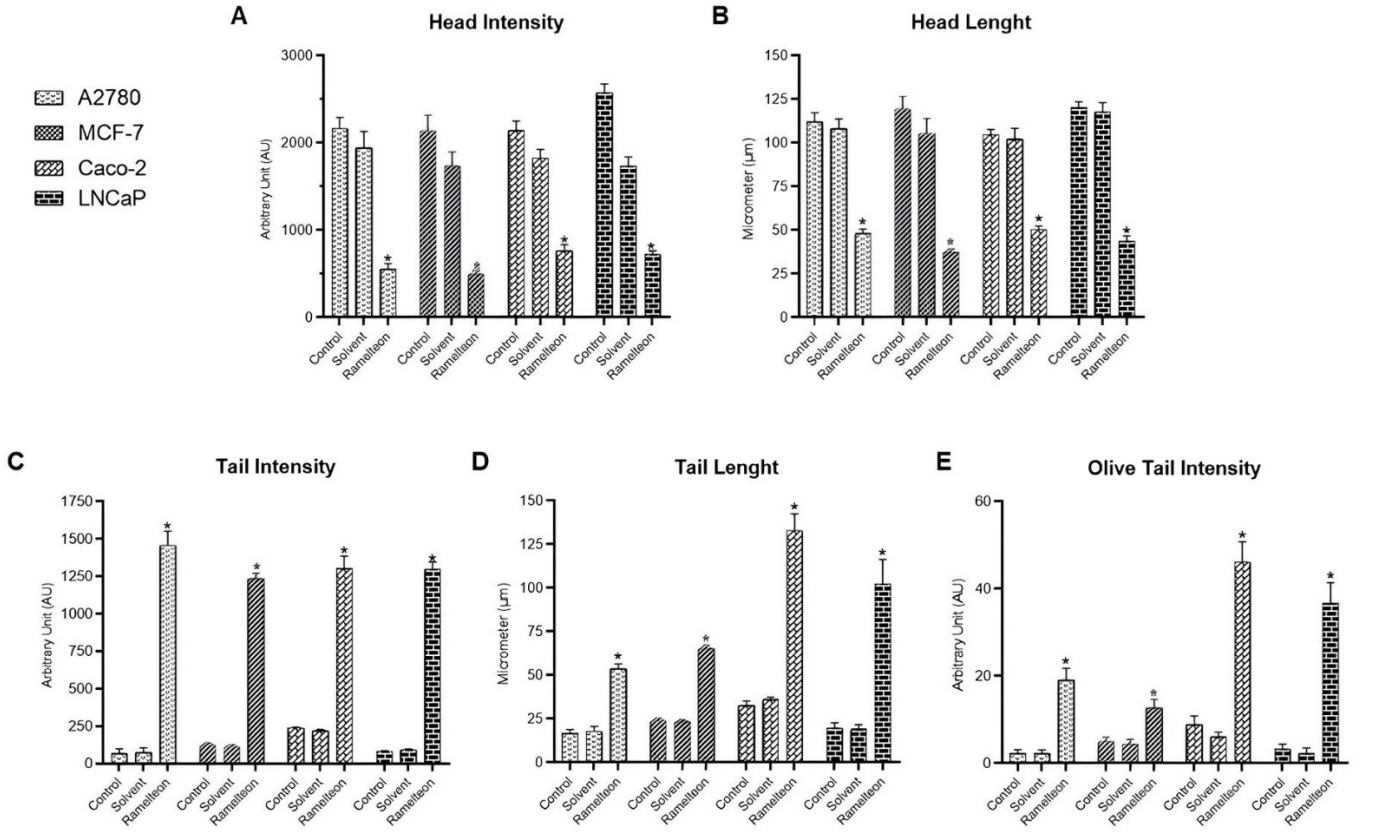

Figure 4. Comet assay obtained from various cancer cells treated with Ramelteon. (*) shows p<0.05.

Table 1. Cell viability results calculated for A2780, MCF-7, Caco-2 and LNCaP cells after a 24-hour incubation with ramelteon compounds.

	Cell Viability (%)			
Groups	A2780	MCF-7	Caco-2	LNCaP
Control	100.00±13.11	100.00±9.85	100.00±15.16	100.00±4.01
Solvent	94.94±13.05	93.94±6.62	95.23±5.12	96.51±10.59
5 µM	82.23±6.71	72.38±6.80*	78.80±14.38	102.30±8.84
10 µM	91.25±2.01	75.04±9.29*	67.96±12.46*	98.51±6.66
25 µM	85.36±11.89	77.83±8.30*	69.31±10.52*	95.13±8.87
50 µM	67.92±15.61*	73.59±13.63*	73.54±12.24*	83.40±7.57
100 µM	61.44±11.24*	60.06±12.28*	59.63±10.55*	77.86±2.51*
500 µM	8.83±1.93*	18.11±6.56*	56.86±10.01*	11.47±0.80*

Each data point is the average of 10 viability measurements. (*) shows p<0.05.

Table 2. IC_{50} (µM) concentrations calculated for A2780, MCF-7, Caco-2 and LNCaP cells in the GraphPad Prism 8 program of 24 h incubation of compounds Ramelteon.

	IC_{50} Concentrations			
Ramelteon (µM)	A2780	MCF-7	Caco-2	LNCaP
	114.8	111.3	186.6	205.2

(Figure 4C), tail length (Figure 4D), and olive tail moment (Figure 4E) were assessed.

The tail moment (Figure 4E) was observed to decrease significantly in all four cell lines following Ramelteon exposure compared with the control (p=0.004). Additionally, numerical data related to DNA damage are presented in Table 3.

DISCUSSION

Studies concerning the antineoplastic effects of melatonin and its analogues have proliferated recently [28]. The anticancer potential of melatonin has been documented; this stems from its ability to inhibit cancer cell proliferation, induce apoptosis, and increase DNA damage [28]. Numerous experimental studies have demonstrated that melatonin has oncostatic effects in several cancers, including breast, ovarian, prostate, oral, stomach, and colorectal cancer [29]. The mechanisms involved include numerous molecular pathways, such as antioxidant activity, modulation of MT1 and MT2, and control of crucial tumor behaviors. These behaviors include apoptosis, metabolism, angiogenesis, invasion, inhibition of metastasis, and induction of epigenetic alterations [30-32]. Melatonin receptor agonists (such as Agomelatin, Tasimelteon, and Ramelteon) are prominent not only are significant therapeutic options for sleep disorders but also for their antioxidant, anti-inflammatory, and antiproliferative effects [33]. In this context, Ramelteon, a selective agonist of the MT1 and MT2 melatonin receptors, was recently approved for the treatment of insomnia characterized by difficulty falling asleep [34]. A review of the literature revealed that Ramelteon reduced cell viability in neuroblastoma, prostate cancer [17], and endometrial cancer cells, and that it exerted its effects via MT1/MT2 receptors

https://doi.org/10.5455/annalsmedres.2025.10.293

Table 3. The extent of DNA damage in A2780, MCF-7, Caco-2 and LNCaP cancer cell lines head density, head length, tail density, tail length and olive tail moment were calculated for each experimental group after Ramelteon treatment.

Head Intensity				
Ramelteon (μM)	A2780	MCF-7	Caco-2	LNCaP
Control	2166.84±126.35	2136.180±181.74	2144.05±106.06	2577.52±98.03
Solvent	1941.62±190.47	1731.120±165.610	1827.41±98.22	1737.21±100.78
Effective Dose (IC_{50})	553.62±66.37*	494.240±44.110*	763.98±72.05*	727.12±36.45*
Head Length				
Ramelteon (μM)	A2780	MCF-7	Caco-2	LNCaP
Control	107.75±1.7	120.880±7.37	104.96±2.88	120.24±3.29
Solvent	107.38±3.04	112.300±8.7	102.04±6.3	117.96±5.24
Effective Dose (IC_{50})	47.23±2.14*	36.45±1.75*	50.32±2.05*	43.84±2.92*
Tail Intensity				
Ramelteon (μM)	A2780	MCF-7	Caco-2	LNCaP
Control	67.19 ±31.12	122.120±15.79	239.71±4.89	80.76±5.36
Solvent	73.04±32.17	112.200±14.65	219.78±8.27	92.38±6.28
Effective Dose (IC_{50})	1455.57±96.61*	1233.570±36.870*	1301.15±85.37*	1297.98±48.64*
Tail Length				
Ramelteon (μM)	A2780	MCF-7	Caco-2	LNCaP
Control	19.12±4.52	21.16±4.49	31.34±7.3	24.75±3.31
Solvent	21.91±3.36	20.8±5.74	29.11±6.01	25.08±5.20
Effective Dose (IC_{50})	57.85±5.15*	64.61±4.14*	137.52±10.64*	108.50±5.74*
Olive Tail Moment				
Ramelteon (μM)	A2780	MCF-7	Caco-2	LNCaP
Control	2.23±0.9	4.91±1.13	8.88±1.99	3.36±1
Solvent	2.29±0.8	4.39±1.12	6.07±1.05	2.25±1.3
Effective Dose (IC_{50})	19.01±2.83*	12.62±2*	46.16±4.64*	36.65±4.84*

[18]. Studies have reported that Ramelteon exhibits therapeutic effects on inflammation, apoptosis, oxidative stress markers, and DNA damage [35, 36]. Ramelteon's anticancer mechanism primarily revolves around the drug's agonistic effect on melatonin receptors (MT1 and MT2), although receptor-independent pathways have also been reported [37]. When MT1 and MT2 receptors are overexpressed in cancer cells, Ramelteon binding suppresses adenylate cyclase activity, thereby reducing cyclic adenosine monophosphate levels and slowing cell proliferation [18]. Furthermore, it has been suggested that Ramelteon may indirectly affect antioxidant pathways (e.g., the nuclear factor erythroid 2-related factor 2 signaling pathway), thereby reducing oxidative stress and DNA damage by exhibiting melatonin-like effects [38]. It has been shown that at high concentrations, Ramelteon can directly trigger mitochondrial dysfunction independent of the receptor by disrupting the balance of the Bcl-2 family proteins, activating apoptotic pathways, and thereby exerting cytotoxic effects [39]. All of these mechanisms support Ramelteon's potential to regulate metastasis, angiogenesis, and the cell cycle in different cancer cell types. Additionally, among other literature data, it has been reported that Ramelteon reduces cellular cytotoxicity in neuroblastoma cells in which an ischemic model has been established [35] and in which a neurodegenerative disease model associated with mitochondrial dysfunction and oxidative stress has been created [39, 40]. These preclinical findings necessitate a clinical re-evaluation of Ramelteon's effects in different disease models and may expand opportunities to develop new strategies for drug repositioning and cost considerations. Thus, drugs in this class may have synergistic effects with existing treatment regimens, improve side-effect profiles, and increase patient tolerability [5, 6]. In this regard, the existing literature is largely limited to toxicity and protective models and to studies using a single cell line; the direct antitumor activity of Ramelteon across different cancer types, in multiple human cell lines, in dose-response analyses, and especially in vivo tumor models remains insufficiently elucidated. Our study aims to investigate the effects of Ramelteon on cell lines derived from ovarian, breast, colon, and prostate cancers, which are among the most common cancers from an epidemiological perspective. In conclusion, Ramelteon has been shown to reduce cell viability significantly in ovarian, breast, colon, and prostate cancer cell lines, and to exhibit protective effects on DNA integrity, as demonstrated by comet assay analyses. In particular, the increase in head diameter and head density and the decrease in tail length, tail density, and olive tail moment suggest that Ramelteon exhibits both anti-cytotoxic and anti-genotoxic profiles. In this context, considering our findings alongside published literature, we predict that Ramelteon,

https://doi.org/10.5455/annalsmedres.2025.10.293

which exhibits antioxidant and anti-inflammatory properties, may also have anticancer properties and be a suitable candidate for cancer treatment. However, a more detailed understanding of the mechanisms of action and their validation in different in vivo models are necessary for the clinical evaluation of Ramelteon's anticancer potential. In this context, Ramelteon's current safety profile and pharmacokinetic properties allow for the evaluation of the drug's repositioning strategy in oncological applications.

CONCLUSION

In our study, the antiproliferative and genotoxic effects of Ramelteon, a melatonergic agonist, on four human cancer cell lines (A2780, MCF-7, Caco-2, LNCaP) were evaluated collectively for the first time. The findings revealed that Ramelteon reduced cell viability in a dose-dependent manner across all cell lines. Ramelteon has shown protective effects against genotoxicity by suppressing DNA damage. In this context, our study demonstrates that Ramelteon exhibits anti-cytotoxic and anti-genotoxic effects in different tumor cell lines, suggesting that Ramelteon may be a potential therapeutic agent for cancer treatment. However, in order to clarify the mechanisms of this effect, comprehensive in vitro and in vivo studies involving advanced molecular-level analyses are required.

Ethics Committee Approval: Ethical approval was not required because it was a cell culture study.

Informed Consent: Informed consent was not obtained due to the cell culture study.

Peer-review: Externally peer-reviewed.

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

Author Contributions: Conception: T.K, G.S; Design: T.K, S.I; Supervision: C.T, S.T, Materials: G.S, S.I; Data Collection and/or Processing: T.K, C.T; Analysis and/or Interpretation: T.K, S.I; Literature Review: G.S, S.I; Writing: T.K, G.S, S.T, Critical Review: C.T, S.T.

Financial Disclosure: This study was supported by the İnönü University Scientific Research Projects Unit with project number TSA-2024-3546.

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

REFERENCES

1. Harris F, Dennison SR, Singh J, Phoenix DA. On the selectivity and efficacy of defense peptides with respect to cancer cells. *Med Res Rev.* 2013;33(1):190-234. doi: 10.1002/med.20252.
2. Zappavigna S, Cossu AM, Grimaldi A, et al. Anti-inflammatory drugs as anticancer agents. *Int J Mol Sci.* 2020;21(7):2605. doi: 10.3390/ijms21072605.
3. Jemel-Oualha I, Elloumi-Mseddi J, Beji A, Hakim B, Aifa S. Controversial effect on Erk activation of some cytotoxic drugs in human LOVO colon cancer cells. *J Recept Signal Transduct Res.* 2016;36(1):21-25. doi: 10.3109/10799893.2014.975246.
4. Chen M, Xu S, Zhou P, He G, Jie Q, Wu Y. Desloratadine citrate disodium injection, a potent histamine H1 receptor antagonist, inhibits chemokine production in ovalbumin-induced allergic rhinitis guinea pig model and histamine-induced human nasal epithelial cells via inhibiting the ERK1/2 and NF-kappa B signal cascades. *Eur J Pharmacol.* 2015;767:98-107. doi: 10.1016/j.ejphar.2015.10.014.
5. Sousa SM, Xavier CPR, Vasconcelos MH, Palmeira A. Repurposing some of the well-known non-steroid anti-inflammatory drugs (NSAIDs) for cancer treatment. *Curr Top Med Chem.* 2023;23(13):1171-1195. doi: 10.2174/1568026623666230130150029.
6. Ioakeim-Skoufa I, Tobajas-Ramos N, Menditto E, et al. Drug Repurposing in Oncology: A Systematic Review of Randomized Controlled Clinical Trials. *Cancers.* 2023;15(11):2972. doi: 10.3390/cancers15112972.
7. Zhao Y, Wang C, Goel A. A combined treatment with melatonin and andrographis promotes autophagy and anticancer activity in colorectal cancer. *Carcinogenesis.* 2022;43(3):217-30. doi: 10.1093/carcin/bgac008.
8. Florido J, Martinez-Ruiz L, Rodriguez-Santana C, et al. Melatonin drives apoptosis in head and neck cancer by increasing mitochondrial ROS generated via reverse electron transport. *J Pineal Res.* 2022;73(3):e12824. doi: 10.1111/jpi.12824.
9. Minocha T, Das M, Rai V, et al. Melatonin induces apoptosis and cell cycle arrest in cervical cancer cells via inhibition of NF-κB pathway. *Inflammopharmacology.* 2022;30(4):1411-29. doi: 10.1007/s10787-022-00964-6.
10. Kast RE. Agomelatine or ramelteon as treatment adjuncts in glioblastoma and other M1-or M2-expressing cancers. *Contemp Oncol (Pozn).* 2015;19(2):157-62. doi: 10.5114/wo.2015.51421.
11. Cajochen C. TAK-375 takeda. *Curr Opin Investig Drugs.* 2005;6(1):114-21. PMID: 15675611.
12. Rawashdeh O, Hudson RL, Stepien I, Dubocovich ML. Circadian periods of sensitivity for Ramelteon on the onset of running-wheel activity and the peak of suprachiasmatic nucleus neuronal firing rhythms in C3H/HeN mice. *Chronobiol Int.* 2011;28(1):31-38. doi: 10.3109/07420528.2010.532894.
13. Stevenson S, Cornelissen K, Clarke E, Hibberd M. Study of the absorption, metabolism, and excretion of (14C)-ramelteon (TAK-375). *Clin Pharmacol Ther.* 2004;75(2):P22-P22. doi: 10.1016/j.clpt.2003.11.084.
14. Kato K, Hirai K, Nishiyama K, et al. Neurochemical properties of ramelteon (TAK-375), a selective MT1/MT2 receptor agonist. *Neuropharmacology.* 2005;48(2):301-10. doi: 10.1016/j.neuropharm.2004.09.007.
15. Kudara M, Kato-Ishikura E, Ikegaya Y, Matsumoto N. Ramelteon administration enhances novel object recognition and spatial working memory in mice. *J Pharmacol Sci.* 2023;152(2):128-35. doi: 10.1016/j.jphs.2023.04.002.
16. Mayanagi S, Haneda R, Inoue M, Ishii K, Tsubosa Y. Ramelteon and suvorexant for postoperative delirium in elderly patients with esophageal cancer. *Esophagus.* 2023;20(4):635-42. doi: 10.1007/s10388-023-01019-7.
17. Balpınar Ö, Nadaroğlu H, Hacımüftüoğlu A. Green synthesis, characterization of melatonin-like drug bioconjugated CoS quantum dots and its antiproliferative effect on different cancer cells. *Mol Biol Rep.* 2023;50(11):9143-51. doi: 10.1007/s11033-023-08817-3.
18. Osanai K, Kobayashi Y, Otsu M, Izawa T, Sakai K, Iwashita M. Ramelteon, a selective MT1/MT2 receptor agonist, suppresses the proliferation and invasiveness of endometrial cancer cells. *Hum Cell.* 2017;30(3):209-15. doi: 10.1007/s13577-017-0169-7.
19. Keser S, Çiftçi EB, Öz S, Tekin S, Kaygılı O. Investigation of in vitro genotoxic, cytotoxic, antiradical and phytochemical properties of endemic Verbascum diversifolium Hochst. *J Phys Chem Funct Mater.* 2025;8(1):15-22. doi: 10.54565/jphcfum.1612825.

https://doi.org/10.5455/annalsmedres.2025.10.293

20. Karataş MO, Tekin S, Alici B, Sandal S. Cytotoxic effects of coumarin substituted benzimidazolium salts against human prostate and ovarian cancer cells. *J Chem Sci.* 2019;131:1-12. doi: 10.1007/s12039-019-1647-0.
21. Özen F, Tekin S, Koran K, Sandal S, Görgülü AO. Synthesis, structural characterization, and in vitro anti-cancer activities of new phenylacrylonitrile derivatives. *Appl Biol Chem.* 2016;59(2):239-48. doi: 10.1007/s13765-016-0163-x.
22. Denizot F, Lang R. Rapid colorimetric assay for cell growth and survival: modifications to the tetrazolium dye procedure giving improved sensitivity and reliability. *J Immunol Methods.* 1986;89(2):271-77. doi: 10.1016/0022-1759(86)90368-6.
23. Horáková Kn, Šovčíková A, Seemannová Z, et al. Detection of drug-induced, superoxide-mediated cell damage and its prevention by antioxidants. *Free Radic Biol Med.* 2001;30(6):650-64. doi: 10.1016/s0891-5849(00)00508-6.
24. Beytur A, Tekin Ç, Çalışkan E, et al. Hexa-substituted cyclotriphosphazene derivatives containing hetero-ring chalcones: Synthesis, in vitro cytotoxic activity and their DNA damage determination. *Bioorg Chem.* 2022;127:105997. doi: 10.1016/j.bioorg.2022.105997.
25. Ozkaya SN, Keskin T, Tekin C, Tekin S, Beytur A. Investigation of the antitumor effects of anti-inflammatory desloratadine and trimebutine on different types of human cancer cells: An in vitro study. *Ann Med Res*. 2023;30(6):727-32. doi: 10.5455/annalsmedres.2023.05.122.
26. Klaude M, Eriksson S, Nygren J, Ahnström G. The comet assay: mechanisms and technical considerations. *Mutat Res.* 1996;363(2):89-96. doi: 10.1016/0921-8777(95)00063-1.
27. Devlin HL, Mack PC, Burich RA, et al. Impairment of the DNA repair and growth arrest pathways by p53R2 silencing enhances DNA damage–induced apoptosis in a p53-dependent manner in prostate cancer cells. *Mol Cancer Res.* 2008;6(5):808-18. doi: 10.1158/1541-7786.MCR-07-2027.
28. Gatti G, Lucini V, Dugnani S, et al. Antiproliferative and pro-apoptotic activity of melatonin analogues on melanoma and breast cancer cells. *Oncotarget.* 2017;8(40):68338-53. doi: 10.18632/oncotarget.20124.
29. Srinivasan V, Spence DW, Pandi-Perumal SR, Trakht I, Cardinali DP. Therapeutic actions of melatonin in cancer: possible mechanisms. *Integr Cancer Ther.* 2008;7(3):189-203. doi: 10.1177/1534735408322846.
30. Mediavilla MD, Sanchez-Barcelo EJ, Tan DX, Manchester L, Reiter RJ. Basic mechanisms involved in the anti-cancer effects of melatonin. *Curr Med Chem.* 2010;17(36):4462-81. doi: 10.2174/092986710794183015.
31. Sánchez-Hidalgo M, Guerrero JM, Villegas I, Packham G, de la Lastra CA. Melatonin, a natural programmed cell death inducer in cancer. *Curr Med Chem.* 2012;19(22):3805-21. doi: 10.2174/092986712801661013.
32. Li Y, Li S, Zhou Y, et al. Melatonin for the prevention and treatment of cancer. *Oncotarget.* 2017;8(24):39896-39921. doi: 10.18632/oncotarget.16379.
33. Liu J, Clough SJ, Hutchinson AJ, Adamah-Biassi EB, Popovska-Gorevski M, Dubocovich ML. MT1 and MT2 Melatonin Receptors: A Therapeutic Perspective. *Annu Rev Pharmacol Toxicol.* 2016;56:361-83. doi: 10.1146/annurev-pharmtox-010814-124742.
34. Devi V, Shankar PK. Ramelteon: a melatonin receptor agonist for the treatment of insomnia. *J Postgrad Med.* 2008;54(1):45-8. doi: 10.4103/0022-3859.39193.
35. Aslankoc R, Savran M, Doğuç DK, Sevimli M, Tekin H, Kaynak M. Ameliorating effects of ramelteon on oxidative stress, inflammation, apoptosis, and autophagy markers in methotrexate-induced cerebral toxicity. *Iran J Basic Med Sci.* 2022;25(10):1183-89. doi: 10.22038/ijbms.2022.62955.13913.
36. Çatakli D, Sevük MA, Coşan S, et al. Ramelteon protects intestinal tissue against injury caused by methotrexate via showing antiapoptotic, antiinflammatory and antioxidant properties. *Bezmialem Sci.* 2022;10(6):770-76. doi: 10.14235/bas.galenos.2022.70783.
37. Mehrzadi S, Pourhanifeh MH, Mirzaei A, Moradian F, Hosseinzadeh A. An updated review of mechanistic potentials of melatonin against cancer: pivotal roles in angiogenesis, apoptosis, autophagy, endoplasmic reticulum stress and oxidative stress. *Cancer Cell Int.* 2021;21(1):188. doi: 10.1186/s12935-021-01892-1.
38. Yang W, Zhang Y, Lu D, et al. Ramelteon protects against human pulmonary microvascular endothelial cell injury induced by lipopolysaccharide (LPS) via activating nuclear factor erythroid 2-related factor 2 (Nrf2)/heme oxygenase-1 (HO-1) pathway. *Bioengineered.* 2022;13(1):1518-29. doi: 10.1080/21655979.2021.2021065.
39. Li C, Zhang Y, Liu R, Mai Y. Ramelteon ameliorated 1-methyl-4-phenylpyridinium (MPP+)-induced neurotoxicity in neuronal cells in a mitochondrial-dependent pathway. *Bioengineered.* 2021;12(1):4868-77. doi: 10.1080/21655979.2021.1960767.
40. Qi X, Tang Z, Shao X, et al. Ramelteon improves blood-brain barrier of focal cerebral ischemia rats to prevent post-stroke depression via upregulating occludin. *Behav Brain Res.* 2023;449:114472. doi: 10.1016/j.bbr.2023.114472.

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Alpha-lipoic acid as a therapeutic agent in ibuprofen-induced liver fibrosis in a rat model

Gulhan Unlu[a], *, Kubra Tugce Kalkan[b]

[a]Kırşehir Ahi Evran University, Faculty of Medicine, Department of Medical Pharmacology, Kırşehir, Türkiye
[b]Kırşehir Ahi Evran University, Faculty of Medicine, Department of Histology and Embryology, Kırşehir, Türkiye
*Corresponding author: inonugulhan@gmail.com (Gulhan Unlu)

MAIN POINTS

- Ibuprofen-induced liver injury caused hepatocyte degeneration, sinusoidal dilatation, collagen accumulation, and increased the expression of TLR4, TGF-β, and α-SMA.
- ALA treatment significantly reduced TLR4 and α-SMA expression and alleviated sinusoidal dilatation and fibrosis, although its effect on TGF-β, was limited.
- The therapeutic potential of ALA lies in its ability to suppress inflammatory signaling and hepatic stellate cell activation, suggesting that ALA may serve as a protective strategy against NSAID-induced liver fibrosis.

Cite this article as: Unlu G, Kalkan KT. Alpha-lipoic acid as a therapeutic agent in ibuprofen-induced liver fibrosis in a rat model. *Ann Med Res*. 2026;33(6):253–260. doi: 10.5455/annalsmedres.2025.09.268.

ABSTRACT

Aim: Ibuprofen (IBU), a commonly used nonsteroidal antiinflammatory drug, may cause hepatotoxicity and fibrosis with prolonged or high-dose use. Alpha-lipoic acid a natural antioxidant, has shown protective effects against oxidative stress and liver injury. The potential therapeutic role of ALA in IBU-induced liver fibrosis was evaluated.

Materials and Methods: Thirty male Wistar Albino rats were divided into five groups (n = 6): control, sham (corn oil), IBU (200 mg/kg by oral gavage for 21 days), IBU + ALA (IBU for 21 days followed by ALA 100 mg/kg by oral gavage for 7 days), and ALA (100 mg/kg for 7 days). Liver tissue was analyzed histopathologically with hematoxylin-eosin and Masson's trichrome staining and immunohistochemically for TLR4, TGF-β, and α-SMA expression. Data were analyzed using GraphPad Prism 10, with normality assessed by Shapiro-Wilk, group comparisons by analysis of variance or Kruskal-Wallis, and $p < 0.05$ considered significant.

Results: IBU caused significant hydropic degeneration, sinusoidal dilatation, and perivascular collagen accumulation compared with controls ($p<0.001$). ALA treatment significantly reduced sinusoidal dilatation ($p<0.01$) and perivascular fibrosis ($p<0.001$) significantly but had no significant effect on hydropic degeneration. Immunohistochemistry showed increased TLR4, TGF-β, and α-SMA expression in the IBU group ($p<0.001$). ALA significantly decreased TLR4 ($p<0.001$) and α-SMA ($p<0.001$), while reduction in TGF-β was not statistically significant.

Conclusion: Although ALA alleviates IBU-induced liver fibrosis by suppressing inflammatory signaling (TLR4) and hepatic stellate cell activation (α-SMA), its effect on TGF-β-mediated fibrogenesis is limited. These findings suggest that ALA may help prevent nonsteroidal anti-inflammatory drug-induced liver injury, but further studies with longer treatment durations are needed.

Keywords: Ibuprofen, Alpha-lipoic acid, Fibrosis, Rat

Received: Oct 03, 2025 **Accepted:** Dec 11, 2025 **Available Online:** Jun 25, 2026

INTRODUCTION

Ibuprofen (IBU) was first introduced in the United Kingdom in 1969 and became internationally available as a prescription medication in the 1970s [1]. IBU is a nonsteroidal antiinflammatory drug that non-selectively inhibits cyclooxygenase (COX) enzymes, thereby reducing prostaglandin synthesis. This mechanism exerts anti-inflammatory, analgesic, and antipyretic effects [2,3]. IBU is rapidly absorbed from the gastrointestinal tract at approximately 80%, reaching a maximum plasma concentration within 1-2 hours after oral administration and demonstrating 99% plasma protein binding. It is metabolized in the liver and excreted by the kidneys, with an elimination half-life ranging from 1.6 to 2.5 hours [4]. While IBU is commonly used for musculoskeletal disorders, osteoarthritis, and rheumatoid arthritis, high doses and long-term use of international bitterness unit have been associated with hepatotoxicity, including drug-induced fibrosis [2,3,5]. Therefore, liver fibrosis may develop secondary to chronic IBU-related hepatic injury, impairing liver function through excessive extracellular matrix accumulation [6].

Fibrosis is characterized by abnormal and excessive extracellular matrix deposition during chronic tissue injury and represents a maladaptive wound-healing response extension [7]. Although initially protective, uncontrolled fibrosis disrupts organ structure and function. Hepatocytes, hepatic stellate cells, sinusoidal endothelial cells, and immune cells play ma-

jor roles in the progression of fibrosis in the liver. Advanced fibrosis and cirrhosis are associated with severe vascular remodeling, which results in impaired hepatocyte perfusion and liver dysfunction [8].

Natural compounds continue to attract increasing interest in biomedical research, as approximately 80% of the global population relies on natural or traditional medicines for health care purposes [9]. Alpha-lipoic acid (ALA) was first identified in the 1930s and later isolated from liver tissue in the 1950s as a thiol-containing molecule with strong antioxidant characteristics [10-12]. ALA is considered a potent antioxidant and mitochondrial cofactor with notable antiinflammatory properties [13]. It is naturally present in various vegetables and is particularly abundant in mitochondria-rich animal tissues such as the liver, heart, and kidneys [14,15]. In addition, all anatomical parts of plants contain polyphenolic compounds, particularly flavonoids, which constitute approximately 60% of naturally occurring polyphenols and display diverse pharmacological activities [16]. Multiple studies have demonstrated the therapeutic benefits of ALA in conditions such as diabetes, atherosclerosis, neurodegenerative diseases, joint disorders, and AIDS [17]. ALA exerts hepatoprotective effects by decreasing oxidative stress through free radical scavenging and regenerating endogenous antioxidants, thereby preventing cellular injury and maintaining hepatic function [18].

Toll-like receptor 4 (TLR4) and transforming growth factor-beta (TGF-β) signaling pathways play key roles in liver fibrogenesis. TLR4 contributes to hepatic injury under certain pathological conditions, such as alcoholic and non-alcoholic fatty liver disease, but its inactivation does not reduce injury in bile duct ligation or CCl4-induced fibrosis models [14, 19-21]. TGF-β is a major fibrogenic cytokine that activates hepatic stellate cells (HSCs) through the TGF-β/Smad pathway, leading to myofibroblast transformation and excessive extracellular matrix production [22-25]. Additionally, TGF-β interacts with Notch, Wnt/β-catenin, and YAP/TAZ pathways, further amplifying inflammation, cellular injury, and fibrosis progression [26-29]. HSCs, located in the space of Disse, lose stored vitamin A during activation and begin expressing α-smooth muscle actin (α-SMA), producing large amounts of extracellular matrix [30-33].

In this study, liver fibrosis will be experimentally induced in rats by high-dose IBU administration, and the potential therapeutic effects of ALA on this fibrotic process will be investigated. Histopathological evaluations and immunohistochemical analyses will be performed to assess morphological and cellular responses. This study aimed to comprehensively evaluate the ability of alpha-lipoic acid to attenuate IBU-induced hepatic injury and fibrosis.

MATERIALS AND METHODS

Animals and the experimental protocol

This study was conducted at the Kirsehir Ahi Evran University Experimental Research Center. Procurement, care, and experimental modeling of the animals were conducted in the same center. The Kirsehir Ahi Evran University Local Animal Experiments Ethics Committee approved the experimental studies prior to their initiation (Approval number: 5/10, Date: 06/03/2025). A total of 30 male Wistar Albino rats, weighing 220 ± 20 g were used in this study. Before and during the experiment, the animals were housed under a 12-hour light/12-hour dark cycle at a constant temperature of 22-24 °C. Standard pellet chow and tap water were provided ad libitum without any restriction. Animal experiments were performed according to the Animal Research: Reporting In Vivo Experiments guidelines.

Experimental groups and treatments used

The animals were randomly assigned to 5 groups using simple randomization, each consisting of 6 rats [34].

Group 1 (control group): No procedure was applied to this group for 4 weeks.

Group 2 (sham group): Corn oil, which was used as the solvent for ALA, was administered to the rats by oral gavage for 4 weeks.

Group 3 (IBU group): The rats received 200 mg/kg IBU [35] by oral gavage for 21 days, followed by a 7-day drug-free period.

Group 4 (IBU + ALA group): The rats received 200 mg/kg IBU for 21 days, followed by oral administration of 100 mg/kg ALA for 7 days [36].

Group 5 (ALA group): No procedure was performed on the rats during the first 21 days. For the last 7 days, 100 mg/kg ALA was administered orally.

At the end of the study, the rats were anesthetized with xylazine (10 mg/kg, intraperitoneally) and ketamine (60 mg/kg, intraperitoneally), tissue samples were collected, and the animals were sacrificed by cervical dislocation. Liver tissue were dissected under deep anesthesia. A portion of the liver tissue was fixed in formaldehyde for histopathological and immunohistochemical evaluations.

Histopathological evaluation of the liver

Tissue samples from each experimental group were preserved in 10% formaldehyde solution for histological examination. After 72 h in formaldehyde, the tissues were rinsed with running tap water, processed through a series of increasing alcohols, and cleared with xylene. Subsequently, they were embedded in paraffin, and paraffin blocks were formed. Polylysine-coated slides were used to mount 5-μm sections of the blocks. The prepared slides were then paraffin-free with xylene using a standard histological staining process,

https://doi.org/10.5455/annalsmedres.2025.09.268

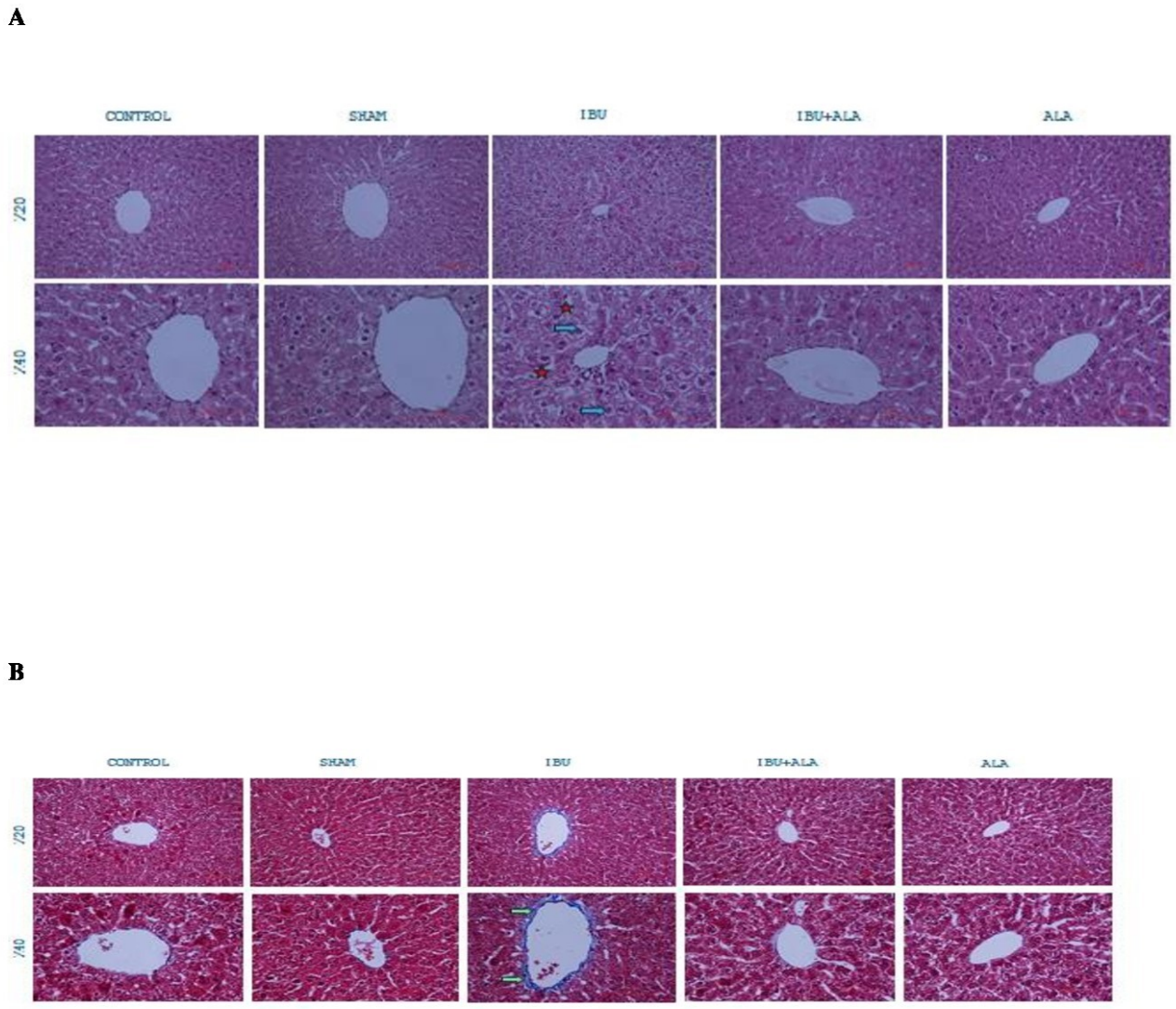

Figure 1. Light microscopic findings in rat liver tissue. (Nikon Eclipse Si, Tokyo, Japan,X200 and X400). The IBU group of HE stained histological section is shown as hydropicdegeneration (blue arrow), sinusoidal dilatation (red star). 1B. Light microscopic findings inrat liver tissue. (Nikon Eclipse Si, Tokyo, Japan, X200 and X400). The IBU group of MTstained histological section is shown as perivascular fibrosis (green arrow).

passed through a graded alcohol series (100%, 96%, 80%, 70%, and 50%), and rinsed in water. The sections were stained with hematoxylin and eosin (H&E) to detect the overall histopathological structure, and Masson's trichrome (MT) was used to assess connective tissue collagen development. They were run through a rising alcohol series, then xylene, and finally covered with a coverslip after dropping entella. They were examined under a light microscope (Nikon Eclipse Si, Tokyo, Japan). The liver tissue was examined for hydropic degeneration, sinusoidal dilatation, and perivascular fibrosis. Each criterion obtained a semi rating on a scale of 0 to 3 (0 =none, 1 = mild, 2 = moderate, and 3 = severe).

Liver immunohistochemical evaluation

The immunohistochemical staining method is performed by linking a marker with specialized antigenic identification characteristics to the cell surface. In this study, we examined the immunohistochemical expression of TLR4, TGF-β and α-SMA in liver tissue from rats with IBU-induced liver fibrosis model. TLR4 (Proteintech; Cat No. 19811-1 AP, 1:400), TGF-β (Proteintech, Cat No. 21898-1 AP, 1/250) and α-SMA (Proteintech, Cat No. 80008-1RR) primary antibodies were used at specific dilutions for the identified markers. The streptavidin-biotin kit (Ultravision Polyvalent (Rabbit-Mouse), Horseradish Peroxidase (HRP) Kit, 125 ml from Thermo Fisher Scientific (Waltham, Massachusetts, USA) was used to apply the avidin-biotin peroxidase method for the immunohistochemical staining method. The sections on polylysine slides were deparaffinized in xylene, dehydrated by maintaining them in each of the decreasing alcohol series, and then kept in distilled water for immunohistochemical staining. Subsequently, the sections were boiled in a microwave oven at 600 W with citrate buffer for antigen retrieval, washed again with phosphate-buffered saline (PBS), and treated with 3% hydrogen peroxide (H_2O_2) to inhibit endogenous per-

https://doi.org/10.5455/annalsmedres.2025.09.268

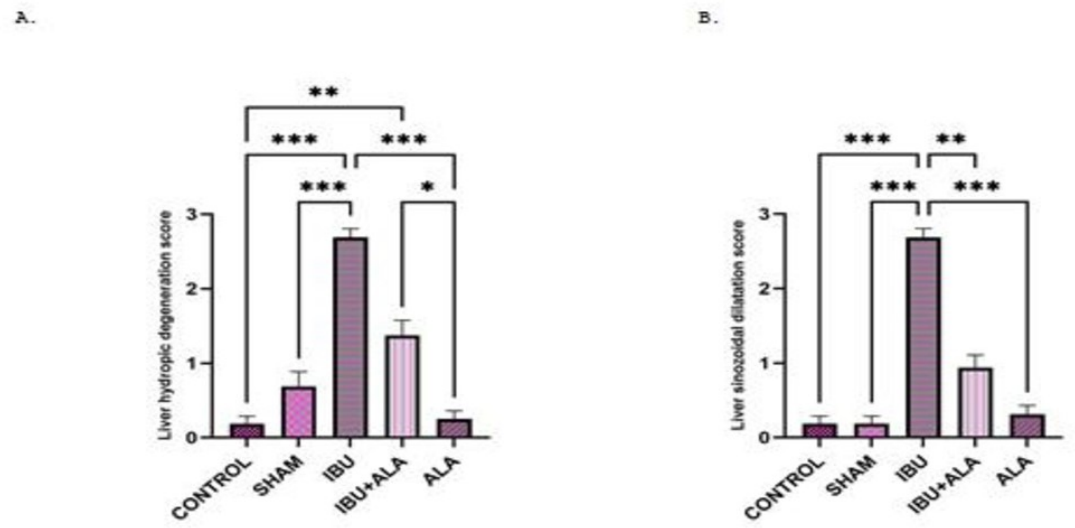

Figure 2. Graphical representation of hydropic degeneration in rat liver tissue groups. *;p<0.05, **; p<0.01, ***; p<0.001. Values are represented as mean ± SD using ANOVA andKruskal-Wallis followed as a post-hoc test Bonferroni and Dunn's test, respectively. 2B.Graphical representation of sinusoidal dilatation in rat liver tissue groups. *; p<0.05, **;p<0.01, ***; p<0.001. Values are represented as mean ± SD using ANOVA and Kruskal-Wallis followed as a post-hoc test Bonferroni and Dunn's test, respectively.

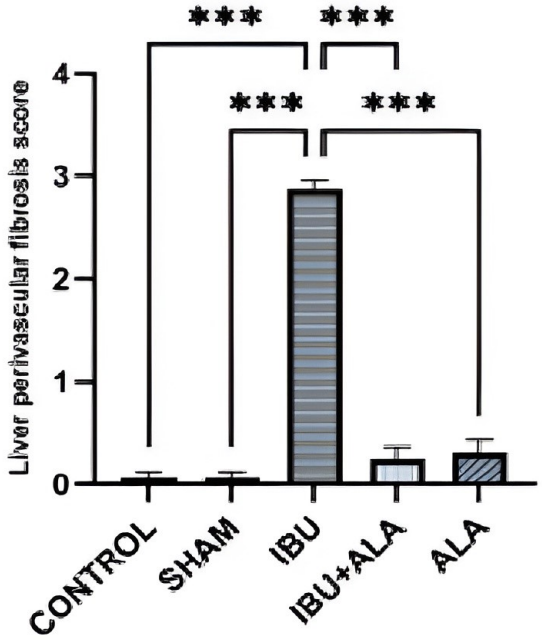

Figure 3. Graphical representation of perivascular fibrosis in rat liver tissue groups. *;p<0.05, **; p<0.01, ***; p<0.001. Values are represented as mean ± SD using ANOVA andKruskal-Wallis followed as a post-hoc test Bonferroni and Dunn's test, respectively.

oxidase activity. Block serum was then applied to the non-antigenic areas for 10 min at room temperature. The tissues in the preparations were covered with the prepared primary antibody, which was then incubated at +4°C overnight. After washing, the sections were treated with a secondary antibody, including biotin. To make immunoreactivity apparent, the sections were next coated with diaminobenzidine (DAB) substrate (Diaminobenzidine chromogen and substrate system, 125 ml from Thermo Fisher Scientific (Waltham, Massachusetts, USA). Liver sections stained with immunohistochemistry were examined under a light microscope, and microscopic images of a number of randomly chosen areas were obtained. The immunoreactivity intensity of the markers that appeared in the images was measured using ImageJ software. The total immunoreactivity was evaluated using the color thresholding feature of ImageJ.

Statistical analysis

All statistical analyses were performed using GraphPad Prism version 10. The Shapiro–Wilk test was used to assess the normality of the data distribution. For comparisons among multiple groups, one-way analysis of variance or the Kruskal-Wallis test was performed depending on the distribution characteristics. Post hoc analyses were performed using the Bonferroni test following analysis of variance and the Dunn test following Kruskal–Wallis analysis. Data are presented as mean ± SD and median (Q1-Q3). A p-value < 0.05 was considered statistically significant.

RESULTS

Light microscopy findings

Figure 1 illustrates 2 different scale levels (X20 and X40), and H&E staining was used to evaluate liver histology in different groups. In the control group, hepatocytes in the liver tissue were coordinated with the central vein, maintaining the

https://doi.org/10.5455/annalsmedres.2025.09.268

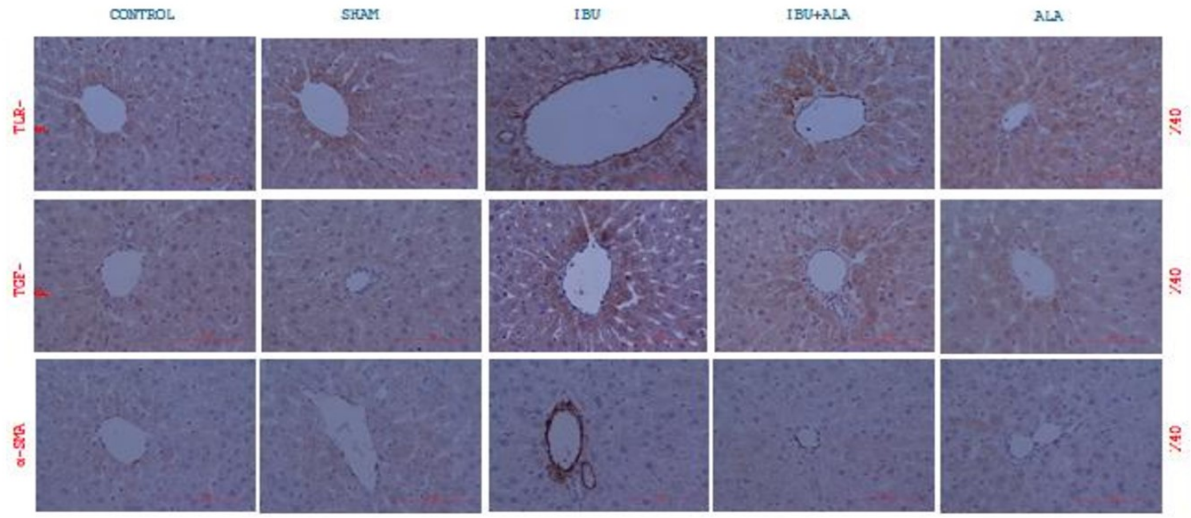

Figure 4. Immunohistochemical light microscopic findings of TLR-4, TGF-β and α-SMAmarkers in rat liver tissue of all groups. (Nikon Eclipse Si, Tokyo, Japan, X400). Browncolored areas with DAB indicate positivity.

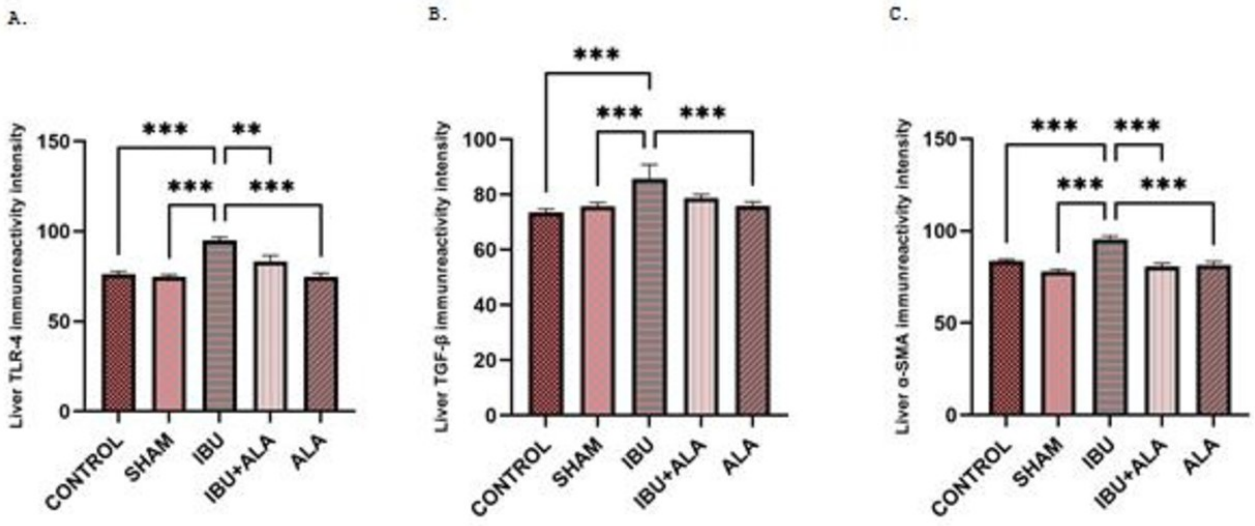

Figure 5. Graphical representation of TLR-4 (A), TGF-β (B) and α-SMA (C) markersreactivity of rat liver tissue groups. *; p<0.05, **; p<0.01, ***; p<0.001. Values arerepresented as mean ± SD using ANOVA and Kruskal-Wallis followed as a post-hoc testBonferroni and Dunn's test, respectively.

Table 1. Liver histopathological scores in experimental groups.

Liver Histoscore	Control (n=6)	SHAM (n=6)	IBU (n=6)	ALA (n=6)	IBU+ALA (n=6)	p value
Hydropic Degeneration	0.0 (0--1)	0.5 (0--2)	3.0 (2--3) ***	1.0 (0--1)	1.0 (0--3)	<0.001
Sinusoidal Dilatation	0.0 (0--1)	0.0 (0--1)	3.0 (2--3) ***	0.0 (0--1)	1.0 (0--2) **	<0.001
Perivascular Fibrosis	0.0 (0--1)	0.0 (0--1)	3.0 (2--3) ***	0.0 (0--1)	0.0 (0--1) ***	<0.001

Values are expressed as median (min–max). Statistical analyses were performed using ANOVA and Kruskal–Wallis tests, followed by Bonferroni and Dunn's post-hoc analyses. *** p<0.001 vs Control and SHAM. ** p<0.01 vs IBU.

cord structure. Sinusoids were clearly visible, and the liver tissue had a normal histological composition. The sham group had an identical structure to the control group, but there was slight vacuolization. The general structure of the ALA group was similar to that of the control group. In the IBU group, hepatocytes exhibited rich cytoplasmic transparency

https://doi.org/10.5455/annalsmedres.2025.09.268

Table 2. Immunoreactivity intensity of fibrosis markers in experimental groups.

Immunreactivity Intensity	Control (n=6)	SHAM (n=6)	IBU (n=6)	ALA (n=6)	IBU+ALA (n=6)	p value
TLR-4	76.17 ± 4.79	74.67 ± 3.86	95.02 ± 5.33	83.22 ± 10.07	74.69 ± 6.38	<0.001
TGF-β	72.69 (64.33–82.22)	75.79 (64.90–88.65)	89.90 (86.20–104.7)	75.65 (64.99–84.45)	79.73 (67.82–86.71)	<0.001
α-SMA	83.67 ± 2.70	78.07 ± 3.35	95.57 ± 4.91	81.54 ± 5.49	80.79 ± 5.08	<0.001

Values are expressed as mean ± SD or median (min–max). Statistical analyses were performed using ANOVA and Kruskal–Wallis followed by Bonferroni and Dunn's post-hoc tests. TLR-4: IBU group showed significantly higher values compared with Control and SHAM (***p<0.001). IBU+ALA group showed a significant reduction compared with IBU (**p<0.01). TGF-β: IBU group was significantly higher than Control and SHAM (***p<0.001). No significant difference between IBU and IBU+ALA groups (p>0.05). α-SMA: IBU group showed significantly higher values compared with Control and SHAM (***p<0.001). IBU+ALA group showed a significant reduction compared with IBU (***p<0.001).

and hydropic degeneration, represented as ballooned cells. In this group, hydropic degeneration showed a statistically significant increase compared to the Control (p<0.001), Sham (p<0.001) and ALA (p<0.001) groups. ALA treatment reduced this score compared with the IBU group, but this decrease was not significant (p>0.05) (Figures 1 and 2A) (Table 1).

The IBU group exhibited considerable sinusoidal dilatation. The outcome was significantly higher in the IBU group than in the Control (p<0.001), Sham (p<0.001), and ALA (p<0.001) groups. ALA treatment reduced this score compared with the IBU group, and the decrease was statistically significant (p<0.01) (Figures 1 and 2B) (Table 1).

MT staining was used to demonstrate the accumulation of collagen in the tissue. As shown in Figure 1B, an increase in connective tissue was observed in the perivascular section in the IBU group. This boost was especially noteworthy compared to the Control (p<0.001), Sham (p<0.001), and ALA (p<0.001) groups. The perivascular fibrosis score dropped significantly in the IBU + ALA group compared with the IBU group (p<0.001) (Figure 3; Table 1).

Immunohistochemical findings

When TLR4 reactivity in liver tissue was examined between groups, the Control, Sham, and ALA groups showed minimal reactivity, whereas TLR4 reactivity was noticeable in the portal areas and vascular walls in the IBU group (Figure 4). TLR4 reactivity was statistically significantly stronger in the IBU group than in the Control (p<0.001), Sham (p<0.001), and ALA (p<0.001) groups. In other words, with ALA treatment, the TLR4 intensity was significantly diminished compared to the IBU group (Figure 5A) (Table 2).

The control, sham, and ALA groups displayed low TGF-β reactivity when compared to the IBU group, which had particularly high levels in hepatocytes surrounding the sinusoidal and portal areas (Figure 4). The IBU group had significantly higher TGF-β reactivity than the Control (p<0.001), Sham (p<0.001), and ALA (p<0.001) groups at the same time. Compared with the IBU group, ALA treatment lessened TGF-β intensity, although this difference was not statistically significant (p>0.05) (Figure 5B) (Table 2).

The analysis of liver tissue α-SMA levels revealed weak expression in the control, sham, and ALA groups, but α-SMA positivity was identified in the IBU group, especially in the vessels (Figure 5). The IBU group demonstrated significantly higher α-SMA reactivity compared to the Control (p<0.001), Sham, (p<0.001) and ALA (p<0.001) groups. In the IBU+ ALA group α-SMA reactivity was significantly reduced compared with that in the IBU group (p<0.001) (Figure 5C) (Table 2).

DISCUSSION

IBU administration caused marked hepatocellular injury, including hydropic degeneration, sinusoidal dilation, and increased perivascular collagen accumulation. These histopathological changes were accompanied by significant increases in TLR4, TGF-β, and α-SMA immunoreactivity, indicating the activation of inflammatory and fibrotic pathways. Overall, these findings confirm that prolonged exposure to IBU damages hepatocytes and triggers fibrosis-associated signaling.

Alpha-lipoic acid treatment notably improved IBU-induced liver injury. Reductions in sinusoidal dilatation, perivascular collagen deposition, and TLR4 and α-SMA expression suggest that ALA protects the liver through multiple mechanisms. Although the decrease in TGF-β reactivity was not statistically significant, the downward trend indicates partial modulation of profibrotic activity. These results suggest that ALA attenuates IBU-mediated liver toxicity by targeting oxidative stress, inflammatory signaling, and fibrogenesis. ALA acts as a potent antioxidant, modulates NF-κB-dependent inflammatory cascades, and inhibits hepatic stellate cell activation [35]. Our findings support this multilevel hepatoprotective effect.

The histopathological changes align with those reported in previous studies. Panchal et al. (2022) observed portal inflammation, hepatocyte degeneration, central vein congestion, and neutrophil infiltration after exposure to IBU. Protective effects were observed with T. zeylanicus extract or silymarin [37]. Similarly, Gola et al. (2013) reported mild sinusoidal degeneration after 6 h of acute hypobaric hypoxia, which progressed to moderate dilation and vascular congestion after 24 h [38]. Kravchuk et al. (2004) found that ALA reduced fatty degeneration and inflammatory changes in alco-

https://doi.org/10.5455/annalsmedres.2025.09.268

holic hepatitis, while fibrotic changes stabilized or decreased [39]. These studies support our findings and highlight the protective potential of ALA against drug-induced liver injury.

Perivascular collagen accumulation, assessed with Masson's trichrome staining, increased significantly in the IBU group, whereas ALA markedly reduced it. The TGF-β pathway, a key fibrosis driver, promotes hepatic stellate cell transdifferentiation into myofibroblasts via Smad-dependent and -independent pathways, enhances extracellular matrix synthesis, and sustains fibrogenesis through positive feedback [23, 24, 36]. The rise in TGF-β in the IBU group aligns with this mechanism. ALA reduced TGF-β levels, although not significantly, suggesting a limited effect on TGF-β-mediated fibrosis.

TLR4 links inflammation and fibrosis from an immune perspective. TLR4 is expressed in Kupffer cells, hepatocytes, and HSCs and recognizes PAMPs and DAMPs, triggering proinflammatory cytokine production [21]. TLR4 activation also sensitizes HSCs to TGF-β-induced fibrosis [40]. Elevated TLR4 expression in the IBU group supports the role of microbial and endogenous triggers in fibrosis. The significant reduction in TLR4 with ALA suggests suppression of inflammatory signaling and slowed fibrosis progression.

α-Smooth muscle actin (α-SMA) is a key marker of HSC activation. α-SMA rises during HSC transdifferentiation into myofibroblast-like cells, reflecting collagen-producing activity [30, 31]. IBU significantly increased α-SMA positivity, confirming HSC activation. ALA significantly reduced α-SMA, indicating suppression of HSC activation and limited fibrotic matrix deposition.

Overall, ALA provides multilevel hepatoprotection against IBU-induced liver injury through its combined antioxidant, antiinflammatory (via TLR4/NF-κB), and antifibrotic effects. This underscores its potential as a therapeutic adjuvant to mitigate NSAID-related hepatotoxicity. Future studies using oxidative stress markers, gene expression analysis, and dose-response evaluations will clarify the translational potential of this study in clinical settings.

Limitations

The current study provides important evidence for the hepatoprotective effects of alpha-lipoic acid against IBU-induced liver injury, although several limitations should be noted. The acute animal model may not fully replicate the complex progression of chronic liver fibrosis seen in clinical cases, potentially limiting the translational relevance of the findings. While ALA significantly suppressed key fibrotic markers, including TLR4 and α-SMA, its more modest effect on TGF-β signaling suggests that it may not completely block all fibrotic pathways. The absence of functional liver tests represents another limitation, as incorporating biochemical parameters, such as ALT and AST levels, could have strengthened the correlation between histological findings and actual liver function. Additionally, the study's single-dose design leaves questions unanswered regarding potential dose-dependent effects and the optimal therapeutic regimen. The exclusive use of an animal model also introduces uncertainties about human applicability due to known species differences in drug metabolism and fibrotic responses. These limitations highlight important areas for future investigation to better establish ALA's potential clinical utility in preventing drug-induced liver fibrosis while not diminishing the value of the current findings.

CONCLUSION

Our findings demonstrate that ALA exerts a notable therapeutic effect on IBU-induced liver fibrosis, particularly in inflammatory (TLR4) and HSC activation (α-SMA) markers, while its impact on TGF-β-mediated fibrogenesis appears more limited. Given the pathophysiological interplay between TGF-β, TLR4, and α-SMA in the literature, the differential effects of ALA on these pathways highlight the importance of targeted combination therapies in treatment strategies.

Ethics Committee Approval: Ethical approval for this study was obtained from the Ethics Committee of Kirsehir Ahi Evran University Local Animal Experiments (date: 06/03/2025, approval number: 5/10).

Informed Consent: Not necessary for this manuscript.

Peer-review: Externally peer-reviewed.

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

Author Contributions: Concept: GÜ.; Design: GÜ.; Supervision: GÜ.; Materials: KTK.; Data Collection and/or Processing: KTK.; Analysis and/or Interpretation: GÜ., KTK.; Literature Review: GÜ.; Writing: GÜ., KTK. Critical Review: GÜ.

Financial Disclosure: Kirsehir Ahi Evran University Scientific Research Projects Unit supported this project under grant number TIP.A3.25.003.

Artificial Intelligence Disclosure: The authors used an AI-based language model, ChatGPT (OpenAI, GPT-5.3), solely for language editing and improving clarity. It had no role in study design, data analysis, interpretation, or conclusions. The authors take full responsibility for the content.

REFERENCES

1. Rainsford KD. Ibuprofen: pharmacology, efficacy and safety. *Inflammopharmacol.* 2009;17(6):275-342. doi: 10.1007/s10787-009-0016-x.
2. Bashyal S. Ibuprofen and its different analytical and manufacturing methods: a review. *Asian J Pharm Clin Res.* 2018;11(7):25-9. doi: 10.22159/ajpcr.2018.v11i7.24484.
3. Ghanim AM, Girgis AS, Kariuki BM, et al. Design and synthesis of ibuprofen-quinoline conjugates as potential antiinflammatory and analgesic drug candidates. *Bioorg Chem.* 2022;119:105557. doi: 10.1016/j.bioorg.2021.105557.

https://doi.org/10.5455/annalsmedres.2025.09.268

4. Kayaalp SO. Akılcı tedavi yönünden tıbbi farmakoloji. *Pelikan Yayıncılık*. 2012.
5. Zoubek ME, Lucena MI, Andrade RJ, Stephens C. Systematic review: ibuprofen-induced liver injury. *Aliment Pharmacol Ther*. 2020;51(6):603-11. doi: 10.1111/apt.15645.
6. Chládek J, Simková M, Vanecková J, Hroch M, Vávrová J, Hůlek P. Assessment of methotrexate hepatotoxicity in psoriasis patients: a prospective evaluation of four serum fibrosis markers. *J Eur Acad Dermatol Venereol*. 2013;27(8):1007-14. doi: 10.1111/j.1468-3083.2012.04643.x.
7. Jarmakiewicz-Czaja S, Gruszecka J, Filip R. What Do NAFLD, Liver Fibrosis, and Inflammatory Bowel Disease Have in Common? Review of the Current Literature. *Metabolites*. 2023;13(3):378. doi: 10.3390/metabo13030378.
8. Bataller R, Brenner DA. Liver fibrosis. *J Clin Investig*. 2005;115(2):209-18. doi: 10.1172/JCI24282.
9. Yunusoglu O, Kalfa I, Demirel ME, et al. A comprehensive review of the pharmacological, therapeutic, and toxicological properties of boric acid and other boron-containing compounds: current landscape and future perspectives. *Pharmacy and Pharmacology*. 2025;13(3):202-238. doi: 10.19163/2307-9266-2025-13-3-202-238.
10. Sneli EE, Strong FM, Peterson WH. Growth Factors for Bacteria: VIII. Pantothenic and Nicotinic Acids as Essential Growth Factors for Lactic and Propionic Acid Bacteria. *J Bacteriol*. 1939;38(3):293-308. doi: 10.1128/jb.38.3.293-308.1939.
11. Morris TW, Reed KE, Cronan Jr, JE. Lipoic acid metabolism in Escherichia coli: the lplA and lipB genes define redundant pathways for ligation of lipoyl groups to apoprotein. *J Bacteriol*. 1995;177(1):1-10. doi: 10.1128/jb.177.1.1-10.1995.
12. Reed LJ, DeBusk BG, Gunsalus IC, Hornberger CS Jr. Crystalline α-lipoic acid: a catalytic agent associated with pyruvate dehydrogenase. *Science*. 1951;114(2952):93-4. doi: 10.1126/science.114.2952.93.
13. Yunusoglu O, Koyuncu E. Alpha-lipoic acid in pharmaceutical development: a comprehensive review of its therapeutic potential and molecular mechanisms. *Prospects Pharm Sci*. 2025;23(4):67–86. doi: 10.56782/pps.456.
14. Nanji AA, Khettry U, Sadrzadeh SH. Lactobacillus feeding reduces endotoxemia and severity of experimental alcoholic liver (disease). *Proc Soc Exp Biol Med*. 1994;205(3):243-7. doi: 10.3181/00379727-205-43703.
15. Wollin SD, Jones PJ. Alpha-lipoic acid and cardiovascular disease. *J Nutr*. 2003;133(11):3327-30. doi: 10.1093/jn/133.11.3327.
16. Yunusoğlu O, Ayaz İ, Dovankaya EH. Pharmacological, medicinal and biological properties of flavonoids: a comprehensive review. *J Res Pharm*. 2025;29(2):561–584. doi: 10.12991/jrespharm.1661054.
17. Golbidi S, Badran M, Laher I. Diabetes and alpha lipoic Acid. *Front Pharmacol*. 2011;2:69. doi: 10.3389/fphar.2011.00069.
18. Yürük AA, Ayaz A. The Health Effects of Alpha Lipoic Acid. *HU Faculty of Health Sci J*. 2014;1(1):11-23.
19. Adachi Y, Moore LE, Bradford BU, Gao W, Thurman RG. Antibiotics prevent liver injury in rats following long-term exposure to ethanol. *Gastroenterology*. 1995;108(1):218-24. doi: 10.1016/0016-5085(95)90027-6.
20. Li Z, Yang S, Lin H, Huang J, et al. Probiotics and antibodies to TNF inhibit inflammatory activity and improve nonalcoholic fatty liver disease. *Hepatology*. 2003;37(2):343-50. doi: 10.1053/jhep.2003.50048.
21. Seki E, De Minicis S, Österreicher CH, et al. TLR4 enhances TGF-β signaling and hepatic fibrosis. Nat Med. 2007;13(11):1324-32. doi: 10.1038/nm1663.
22. Bedair Dewidar CM, Dooley S, Meindl-Beinker N. TGF-β in Hepatic Stellate Cell Activation and Liver Fibrogenesis—Updated 2019. *Cells*. 2019;8(11):1419. doi: 10.3390/cells8111419.
23. Higashi T, Friedman SL, Hoshida Y. Hepatic stellate cells as key target in liver fibrosis. *Adv Drug Deliv Rev*. 2017;121:27-42. doi: 10.1016/j.addr.2017.05.007.
24. Xu F, Liu C, Zhou D, Zhang L. TGF-β/SMAD Pathway and Its Regulation in Hepatic Fibrosis. *J Histochem Cytochem*. 2016;64(3):157-67. doi: 10.1369/0022155415627681.
25. Dooley S, Hamzavi J, Breitkopf K, et al. Smad7 prevents activation of hepatic stellate cells and liver fibrosis in rats. *Gastroenterology*. 2003;125(1):178-91. doi: 10.1016/s0016-5085(03)00666-8.
26. Zhang K, Han X, Zhang Z, et al. The liver-enriched lnc-LFAR1 promotes liver fibrosis by activating TGFβ and Notch pathways. *Nat Commun*. 2017;8(1):144. doi: 10.1038/s41467-017-00204-4.
27. Wang J-n, Li L, Li L-y, Yan Q, Li J, Xu T. Emerging role and therapeutic implication of Wnt signaling pathways in liver fibrosis. *Gene*. 2018;674:57-69. doi: 10.1016/j.gene.2018.06.053.
28. Weiskirchen R, Weiskirchen S, Tacke F. Organ and tissue fibrosis: Molecular signals, cellular mechanisms and translational implications. *Mol Aspects Med*. 2019;65:2-15. doi: 10.1016/j.mam.2018.06.003.
29. Wang C, Ma C, Gong L, et al. Macrophage polarization and its role in liver disease. *Front Immunol*. 2021;12:803037. doi: 10.3389/fimmu.2021.803037.
30. Asahina K. Hepatic stellate cell progenitor cells. *J Gastroenterol Hepatol*. 2012;27(Suppl 2):80-4. doi: 10.1111/j.1440-1746.2011.07001.x.
31. Abd El-Rahman SS, Fayed HM. Targeting AngII/AT1R signaling pathway by perindopril inhibits ongoing liver fibrosis in rat. *J Tissue Eng Regen Med*. 2019;13(12):2131-2141. doi: 10.1002/term.2940.
32. Hessin AF, Hegazy RR, Hassan AA, Yassin NZ, Kenawy SA. Resveratrol prevents liver fibrosis via two possible pathways: modulation of alpha fetoprotein transcriptional levels and normalization of protein kinase C responses. *Indian J Pharmacol*. 2017;49(4):282-9. doi: 10.4103/ijp.IJP_299_16.
33. Weiskirchen R, Tacke F. Liver fibrosis: from pathogenesis to novel. *Dig Dis*. 2016;34(4):410-22. doi: 10.1159/000444556.
34. Charan J, Kantharia ND. How to calculate sample size in animal studies? *J Pharmacol Pharmacother*. 2013;4(4):303-6. doi: 10.4103/0976-500X.119726.
35. Kumar G, Hota D, Saikia UN, Pandhi P. Evaluation of analgesic efficacy, gastrotoxicity, and nephrotoxicity of fixed-dose combinations of nonselective, preferential, and selective cyclooxygenase inhibitors with paracetamol in rats. *Exp Toxicol Pathol*. 2010;62(6):653-62. doi: 10.1016/j.etp.2009.08.011.
36. Dik B, Hatipoglu D, Ates MB. Potential effects of Resatorvid and alpha lipoic acid on gentamicin-induced nephrotoxicity in rats. *Pharmacol Res Perspect*. 2024;12(4):e1222. doi: 10.1002/prp2.1222.
37. Panchal NK, Swarnalatha P, and Prince SE. Trichopus zeylanicus ameliorates ibuprofeninebriated hepatotoxicity and enteropathy: an insight into its modulatory impact on pro/anti-inflammatory cytokines and apoptotic signaling pathways. *Inflammopharmacology*. 2022;30(6):2229-2242. doi: 10.1007/s10787-022-01052-5.
38. Gola S, Keshri GK, Gupta A. Hepatic metabolism of ibuprofen in rats under acute hypobaric hypoxia. *Exp Toxicol Pathol*. 2013;65(6):751-8. doi: 10.1016/j.etp.2012.11.001.
39. Kravchuk IuA, Mekhtiev SN, Uspenskiĭ IuP, Grinevich VB, Koblov SV. Device laboratory and postmortem parallels in alcoholic hepatitis during combined therapy using thioctic (alpha-lipoic) acid. *Klin Med (Mosk)*. 2004;82(6):55-7. PMID: 15344693.
40. Paik YH, Schwabe RF, Bataller R, Russo MP, Jobin C, Brenner DA. Toll-like receptor 4 mediates inflammatory signaling by bacterial lipopolysaccharide in human hepatic stellate cells. *Hepatology*. 2003;37(5):1043-55. doi: 10.1053/jhep.2003.50182.

https://doi.org/10.5455/annalsmedres.2025.09.268

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Primary cicatricial alopecia subtypes: Clinical and histopathological patterns in a large Turkish cohort

Zeynep Karaca Ural[a], *, Zeynep Utlu[a], Merve Hatun Erkayman[a], Handan Bilen[a], Bilal Akif Babacan[a], Nese Gocer Gurok[b], Serkan Naktiyok[c]

[a]Atatürk University, Faculty of Medicine, Department of Dermatology, Erzurum, Türkiye
[b]University of Health Sciences, Elazığ Fethi Sekin City Health Application and Research Center, Department of Dermatology, Elazığ, Türkiye
[c]Atatürk University, Faculty of Economics and Administrative Sciences, Department of Labor Economics and Industrial Relations, Erzurum, Türkiye
*Corresponding author: zeynepkaraca.zk90@gmail.com (Zeynep Karaca Ural)

MAIN POINTS

- The largest series of PCA from Türkiye (n = 116, histopathologically confirmed).
- DLE was the most common (60.3%), followed by LPP (28.4%); FFA was rare (4.3%).
- DLE showed basement membrane thickening and dermal mucin; whereas LPP showedsebaceous gland atrophy and epidermal hypertrophy.
- Clinicopathological correlation is essential for accurate diagnosis.

Cite this article as: Karaca Ural Z, Utlu Z, Erkayman MH, Bilen H, Babacan BA, Gocer Gurok N, Naktiyok S. Primary cicatricial alopecia subtypes: Clinical and histopathological patterns in a large Turkish cohort. *Ann Med Res*. 2026;33(6):261–266. doi: 10.5455/annalsmedres.2025.09.260.

ABSTRACT

Aim: Primary cicatricial alopecias (PCAs) are a group of rare dermatological disorders characterized by irreversible destruction of hair follicles due to inflammatory processes, leading to permanent hair loss. The diagnostic process is often delayed because of nonspecific clinical presentation and overlapping histopathological features. This study aimed to analyze the clinical and histopathological characteristics of PCA subtypes diagnosed in a tertiary dermatology clinic in Türkiye over a 17-year period.

Materials and Methods: A retrospective analysis was conducted on 116 cases of histopathologically confirmed PCA between January 2008 and January 2025. Clinical diagnoses included discoid lupus erythematosus (DLE), lichen planopilaris (LPP), frontal fibrosing alopecia (FFA), pseudopelade of Brocq (PPB), and folliculitis decalvans (FD). Histopathological features were systematically reviewed and statistically analyzed using the Statistical Package for the Social Sciences version 25.0.

Results: The most frequent subtype was DLE (60.3%), followed by LPP (28.4%). DLE was associated with basal membrane thickening (82.9%) and dermal mucin deposition (71.4%), whereas LPP showed higher rates of sebaceous gland loss (51.5%) and interface dermatitis (54.5%). FFA exhibited overlapping features with LPP, including 100% basal cell degeneration. Statistically significant associations were found between specific histopathological patterns and clinical subtypes.

Conclusion: This study provides one of the comprehensive PCA datasets in Türkiye, highlighting distinct histopathological patterns associated with specific clinical subtypes. These findings underscore the importance of clinicopathological correlation in PCA diagnosis and contribute valuable data for improving diagnostic accuracy and patient management.

Keywords: Cicatricial alopecia, Lichen planopilaris, Discoid lupus erythematosus, Histopathology, Scalp diseases

Received: Sep 16, 2025 **Accepted:** Dec 18, 2025 **Available Online:** Jun 25, 2026

INTRODUCTION

Primary cicatricial alopecias (PCAs) are rare dermatological disorders that cause irreversible damage to the hair follicle and the surrounding structures due to inflammatory processes. The epithelial stem cells of the hair follicle are irreversibly destroyed, and their replacement by fibrous tissue leads to permanent hair loss. Psychological distress and reduced quality of life are associated with PCAs [1,2]. The prevalence of PCAs among all patients presenting to dermatology outpatient clinics is approximately 0.09% [1], and they constitute 2.1%–7.3% of all hair diseases [3–5]. This disease group exhibits a complex etiopathogenesis and a variable clinical course.

Although different classification systems have been proposed for PCAs, the North American Hair Research Society (NAHRS) classification is the most widely accepted model, which categorizes PCAs according to the predominant inflammatory infiltrate observed in histopathology. Ac-

cording to this classification, PCAs are classified into 4 main groups. The lymphocyte-predominant group includes patients with discoid lupus erythematosus (DLE), lichen planopilaris (LPP) and its subtypes, pseudopelade of Brocq (PPB), and central centrifugal cicatricial alopecia (CCCA). The neutrophil-predominant group includes folliculitis decalvans (FD) and dissecting cellulitis (DC). The third group, which shows mixed inflammatory cell infiltrates, includes acne keloidalis nuchae (AKN). Cases that do not show a specific inflammatory cell type are classified as nonspecific [6].

The diagnosis of PCAs can be challenging for both dermatologists and pathologists, particularly when patients present at advanced stages in which histopathological characteristics may no longer be specific. Therefore, early diagnosis largely depends on the combined evaluation of clinical findings and histopathological data [7]. Although many studies have investigated the distribution of PCA subtypes and their histopathological features in different geographical regions, case series from Türkiye remain limited. This retrospective study aimed to examine the demographic characteristics, clinical subtypes, and histopathological features of patients who presented to the dermatology outpatient clinic of a tertiary healthcare center and were definitively diagnosed with PCA based on clinical and histological findings. In addition, this study aims to provide local epidemiological data that may support diagnostic processes.

MATERIALS AND METHODS

In this study, the clinical and histopathological characteristics of patients diagnosed with cicatricial alopecia who presented to the dermatology outpatient clinic of a tertiary healthcare center between January 2008 and January 2025 were evaluated. A total of 119 patients were evaluated histopathologically; however, 3 patients were excluded due to missing histopathology slides and incomplete demographic data that prevented accurate classification. Consequently, 116 patients with confirmed histopathological diagnoses were included in the final analysis.

This was a descriptive, non-interventional study based on a retrospective review of existing patient records. In retrospective studies, the sample size depends on the suitability of the available records rather than the researchers' choice. Studies with similar designs can be conducted with smaller samples [8]. Therefore, in this study, the sample size was determined based on the accessible clinical records belonging to the study period.

The study was approved by the Atatürk University Non-Interventional Clinical Research Ethics Committee (date: 27.06.2025, decision no: 64). The principles of the Declaration of Helsinki were adhered to throughout the study; patient confidentiality was maintained, data were anonymized, and no identifying information was recorded. Clinical photographs were included with written informed consent and anonymized to ensure patient confidentiality. No age or sex restrictions were applied.

In our clinic, scalp biopsies are performed in patients with suspected cicatricial alopecia in the presence of clinical ambiguity, cases requiring differential diagnosis, such as suspected discoid lupus erythematosus, atypical or early-stage lichen planopilaris/pseudopelade of Brocq lesions, or lesions that are refractory to treatment.

The inclusion criteria were as follows: presentation to the dermatology outpatient clinic between January 2008 and January 2025 and confirmation of the clinical diagnosis of cicatricial alopecia by dermatopathological evaluation. Patients without a definitive diagnosis or with incomplete records were excluded from the study.

Dermatology clinic researchers retrospectively collected data using the hospital information management system, pathology reports, and clinical examination notes. For each patient, age, sex, year of diagnosis, clinical diagnosis (LPP, FFA, DLE, and PPB), and histopathological findings (epidermal atrophy and hypertrophy, follicular plugging, basal cell degeneration, basement membrane thickening, interface dermatitis, dermal mucin deposition, and sebaceous gland atrophy) were systematically recorded.

The study was conducted and reported in accordance with the STROBE guidelines.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics version 25.0 (Armonk, NY: IBM Corp.). Descriptive statistics included mean, median, standard deviation, and percentage distributions. Pearson's Chi-square test was used for comparisons between subgroups for categorical variables. The likelihood ratio and linear-by-linear association tests were applied when appropriate. A p-value of <0.05 was considered statistically significant.

RESULTS

Among the 116 CAS included in the study, 51.7% were female (n = 60) and 48.3% were male (n = 56). most patients (52.6%) were aged between 31 and 50 years. When analyzed according to the year of diagnosis, 52.6% of the cases was diagnosed in 2020 or later.

Regarding clinical subtypes, the most common diagnosis was DLE with 60.3%, followed by LPP with 28.4%. Frontal FFA (4.3%), PPB (3.4%), FD (1.7%), and unclassified cases (1.7%) were observed less frequently. Figure 1 shows representative clinical photographs illustrating the characteristic features of the subtypes.

A statistically significant association was found between sex and clinical diagnosis (p = 0.048). LPP was more common in men, whereas FFA and PPB were detected only in women. No significant association was found between age groups and

https://doi.org/10.5455/annalsmedres.2025.09.260

Table 1. Demographic characteristics of the participants and distribution of clinical diagnoses.

Variable	DLE (n=70)	LPP (n=33)	FD (n=2)	PPB (n=4)	FFA (n=5)	Unclassified (n=2)	p-value	χ^2	d
Sex							**0.048***	10.689	5
Male	34 (48.6%)	19 (57.6%)	0	1 (25%)	0	2 (100%)			
Female	36 (51.4%)	14 (42.4%)	2 (100%)	3 (75%)	5 (100%)	0			
Age Group							0.647	7.809	10
≤30	14 (20%)	6 (18.2%)	1 (50%)	2 (50%)	2 (40%)	0			
31–50	36 (51.4%)	18 (54.%5)	1 (50%)	1 (25%)	3 (60%)	2 (100%)			
≥51	20 (28.6%)	9 (27.3%)	0	1 (25%)	0	0			
Distribution by Year							**0.014***	22.211	10
2008-2013	9 (12.9%)	5 (15.2%)	0	1 (25%)	0	0			
2014-2019	32 (45.7%)	3 (9.1%)	0	3 (75%)	1 (20%)	1 (50%)			
≥2020	29 (41.4%)	25 (75.8%)	2 (100%)	0	4 (80%)	1 (50%)			

Table 2. Histopathological findings by clinical diagnosis groups.

Histopathological Features	DLE	LPP	Folliculitis decalvans	PPB	FFA	p	χ^2	d
Epidermal atrophy	39 (55.7)	18 (54.5)	0 (0.0)	4 (100.0)	2 (40.0)	0.286	6.212	5
Epidermal hypertrophy	2 (2.9)	7 (21.2)	0 (0.0)	0 (0.0)	0 (0.0)	**0.038***	11.789	5
Follicular plugging	38 (54.3)	13 (39.4)	2 (100.0)	0 (0.0)	3 (60.0)	0.073	10.073	5
Basal cell degeneration	54 (77.1)	25 (75.8)	0 (0.0)	0 (0.0)	5 (100.0)	**<0.001***	23.873	5
Basement membrane thickening	58 (82.9)	13 (39.4)	0 (0.0)	0 (0.0)	0 (0.0)	**<0.001***	40.943	5
Interface dermatitis	29 (41.4)	18 (54.5)	0 (0.0)	0 (0.0)	3 (60.0)	0.225	6.941	5
Mucindeposition dermis	50 (71.4)	1 (3.0)	0 (0.0)	0 (0.0)	0 (0.0)	**<0.001***	54.521	5
Sebaceous gland loss	14 (20.0)	17 (51.5)	2 (100.0)	4 (100.0)	3 (60.0)	**0.002***	24.630	5

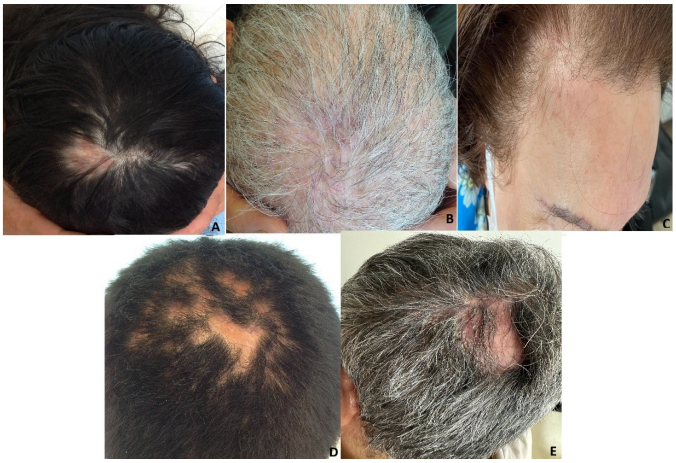

Figure 1. Representative clinical images of primary cicatricial alopecia subtypes: A) discoid lupus erythematosus: atrophic, scarring, erythematous alopecic plaques; B) Lichen planopilaris: perifollicular erythema and scaling; C) frontal fibrosing alopecia: frontal hairline recession and eyebrow loss; D) pseudopelade of Brocq: scarring alopecic patches lacking erythema; E) folliculitis decalvans: tufted appearance characterized by pustules, crusting, and multiple hairs emerging from single follicular openings.

clinical diagnoses (p = 0.647). Of the LPP cases, 75.8% was diagnosed in 2020 or later. A statistically significant association was found between the year of diagnosis and clinical subtypes (p = 0.014). Table 1 presents the demographic characteristics of the participants and the distribution of clinical diagnoses.

The most frequently observed histopathological findings were basal cell degeneration (72.4%), basement membrane thickening (61.2%), epidermal atrophy (55.2%), and follicular plugging (48.3%).

Differences were observed in the distribution of histopathological findings among PCA subtypes. Changes were present in both epidermal and dermal structures in DLE cases, including basement membrane thickening (82.9%), mucin deposition (71.4%), and basal cell degeneration (77.1%). Basal cell degeneration (75.8%), epidermal atrophy (54.5%), interface dermatitis (54.5%), and sebaceous gland atrophy (51.5%) were predominant in LPP. In addition, basement membrane thickening was less frequent in LPP (39.4%) than in DLE. In FFA cases, basal cell degeneration was present in all patients (100%). Other common findings included follicular plugging (60%), interface dermatitis (60%), and epidermal atrophy (40%) (Table 2).

Basal cell degeneration was significantly more frequent in both the LPP and DLE groups (p<0.001). Sebaceous gland atrophy and epidermal hypertrophy were also significantly more common in the LPP group (p = 0.002 and p = 0.038, respectively). In contrast, basement membrane thickening and mucin deposition were significantly more frequent in patients with DLE (both p<0.001). No statistically significant differences in epidermal atrophy (p = 0.286), interface dermatitis (p = 0.225), or follicular plugging (p = 0.073) were observed among the groups.

https://doi.org/10.5455/annalsmedres.2025.09.260

DISCUSSION

This study constitutes one of the comprehensive series addressing the histopathological features of PCAs in Türkiye. The large sample of 116 patients provides significant support for correlating clinical subtypes with different histopathological patterns and for diagnostic differentiation.

In our study, DLE was the most frequently observed subtype, followed by LPP. FFA was detected in only 4.3% of the patients. PCA subtype prevalence has been reported at different frequencies in different populations. In addition to studies reporting FFA as the most common subtype and LPP as the second most common subtype [9], studies from Iran and Pakistan have reported that DLE is the most frequently observed subtype [10,11]. In other studies conducted in Türkiye, LPP has been stated as the most common subtype, whereas the frequency of DLE shows considerable variability [4,12]. The high rate of DLE in our cohort may be attributed to the frequent preference for biopsy in cutaneous lupus cases, whereas the low frequency of FFA may be explained by patients presenting to our clinic at advanced stages of the disease, at which point biopsy may no longer be considered necessary.

The frequency of FD in our study and in other series reported from our country (1.7%–2.1%) is significantly lower than that reported in studies conducted in China (40.7%) when the rarer subtypes are examined [3,4]. FD is usually clinically recognizable, and in our center, direct treatment is generally preferred over biopsy in inflammatory or infectious lesions. This clinical approach may explain the low frequency of FD observed in our study. In the literature, the prevalence of PPB has also been reported at highly variable frequencies (3.3%–40%) [1,5,7,10,13]. Among our patients, this rate was 3.4%, whereas it ranges between 5% and 8% in other Turkish case series [4,12]. This variability may be due to differences in diagnostic criteria and the conceptual ambiguity surrounding PPB. While some authors consider PPB as the end-stage manifestation of other PCAs [14], the North American Hair Research Society classifies it as a separate subtype [6]. Since biopsies in our center are generally obtained from active lesions, late-stage cases may have been excluded from histopathological evaluation.

When all findings regarding subtype distribution are evaluated, the factors influencing the distribution of PCA subtypes include not only disease prevalence but also patient presentation profile, geographical characteristics, clinical practice patterns, and institutional diagnostic approaches.

Regarding the relationship between sex and PCA subtypes, most studies report that LPP is more common in women [13,15]. Similar findings have also been reported in prevalence studies conducted in neighboring countries, such as Iran and Türkiye [4,11,12]. Contrary to the female predominance frequently described in the literature, LPP was more commonly observed in men in our study. This inconsistency may be associated with sampling characteristics specific to the center, differences in diagnostic practices, or the possibility that LPP variants observed in men may exhibit more prominent clinical features.

FFA and PPB were observed only in female patients in our study, which is largely consistent with the literature. FFA is particularly common in postmenopausal women, and this may be attributed to hormonal, immunological, or genetic factors [16,17]. Similarly, female predominance has also been reported in the literature [1,4,7,10].

No statistically significant relationship was found between age groups and clinical subtypes among the patients examined. However, FFA and LPP are more common in older adults, whereas DLE and PPB are more prevalent in younger adults [1,4,10].

FFA is considered a clinical variant or subtype of LPP [18], and several studies have emphasized a marked increase in FFA incidence over the past decade [5,9,16]. This increase may be attributed to the widespread use of noninvasive diagnostic tools, such as trichoscopy, increased awareness of hair disorders, earlier dermatological consultations, better clinical recognition of LPP, and improved sensitivity in histopathological evaluations [19,20]. In our study, 75.8% of patients diagnosed with LPP received their diagnosis in 2020 or later.

The most frequently reported histopathological findings were basal cell degeneration, basement membrane thickening, epidermal atrophy, and follicular plugging. Sardar et al. [1] reported that lymphocytic infiltration was the predominant feature and that basal cell degeneration and follicular plugging were the most frequently observed findings. Similarly, Hashmi et al. [10] described lymphocytic infiltration, perivascular infiltrates, sebaceous gland loss, perifollicular fibrosis, and epidermal atrophy. The histopathological features frequently observed in patients with DLE and LPP were largely consistent with the classical histopathological patterns described in the literature [1,10].

Basal cell degeneration was significantly higher in both the LPP and DLE groups. Follicular lichenoid inflammation and vacuolar degeneration are defined as characteristic findings in LPP [21]. Bernárdez et al. [22] reported that vacuolar interface dermatitis is commonly observed in DLE, whereas lichenoid interface changes are more typical in LPP, usually accompanied by basal cell degeneration.

In our study, basement membrane thickening and dermal mucin deposition were significantly more frequently observed in patients with DLE. Sardar et al observed these findings only in patients with DLE, whereas Hashmi et al. reported basement membrane thickening in all patients with DLE [1,9]. A PAS-positive thickened basement membrane has been defined as a classic feature of DLE [22], and this finding has been emphasized as being rarely observed in LPP [23]. Similarly, dermal mucin deposition is considered distinctive in lupus-related diseases [22,24].

In our study, sebaceous gland atrophy and epidermal hypertrophy were significantly more common in the LPP group

https://doi.org/10.5455/annalsmedres.2025.09.260

than in the other cicatricial alopecia subtypes. However, the literature indicates that sebaceous gland loss is commonly observed not only in LPP but also in other PCA subtypes such as FFA, FD, and PPB [1]. Loss of sebaceous glands is commonly observed not only in LPP but also in other PCA subtypes, such as FFA, FD, and PPB [1]. Although Hashmi et al. [9] reported this finding most prominently in PPB cases, Inchara et al. [24] questioned its utility as a distinguishing diagnostic feature, as it can be seen in both LPP and DLE. Sebaceous gland atrophy may appear in the early stages of the disease [22].

Although epidermal hypertrophy is more commonly associated with DLE, FFA, and acne keloidalis nuchae (AKN), it has been reported in 20% of patients with LPP, similar to our study [1]. This rate was found to be statistically significant. Methodological factors such as sample size or variation in biopsy site selection may be responsible.

Limitations

This study has several limitations, such as its relatively small sample size, which included only histopathologically confirmed PCA cases in a tertiary healthcare center, and its cross-sectional design, which prevented the evaluation of the clinical course. Other important limitations include limited clinical data and the absence of trichoscopic evaluation. Methodologically, multiple comparison corrections were not applied because the comparison of more than two groups was descriptive in nature; this constitutes a limitation in the interpretation of the p-values. In addition, the small sample sizes in some subtypes partially restricted the chi-square test assumptions. Because the groups reflect the true distribution, they were not combined; therefore, Fisher's exact test would have been more appropriate in some analyses.

CONCLUSION

This retrospective study comprehensively analyzed the clinical and histopathological features associated with the subtypes of PCA. Findings such as basement membrane thickening and dermal mucin deposition in DLE and basal cell degeneration and sebaceous gland atrophy in LPP were identified as relatively specific features. Our findings emphasize that histopathological evaluation alone may not be sufficient for an accurate diagnosis and that diagnostic accuracy increases when histological findings are interpreted in conjunction with the clinical context. In this regard, our study represents one of the largest PCA case series reported from Türkiye.

Ethics Committee Approval: The study was approved by the Atatürk University Non-Interventional Clinical Research Ethics Committee (date: 27.06.2025, decision no: 64).

Informed Consent: Clinical photographs used in the manuscript were obtained after written informed consent from the patients. For the retrospective chart review component, informed consent was waived due to the anonymized nature of the data and the retrospective study design.

Peer-review: Externally peer-reviewed.

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

Author Contributions: Z.K.U.: conception, materials, data collection and/or processing, literature review, writing; Z.U.: design, materials, data collection and/or processing, critical review; M.H.E.: design, data collection and/or processing, literature review, critical review; H.B.: design, data collection and/or processing, literature review, critical review; B.A.B.: conception, data collection and/or processing, literature review; N.G.G.: design, materials, data collection and/or processing; S.N.: analysis and/or interpretation, literature review.

Financial Disclosure: The authors declare that no funding was received for this study.

Artificial Intelligence Disclosure: ChatGPT (OpenAI, GPT-5 series) and NotebookLM (Google) were used solely for linguistic editing, readability enhancement, literature organization, and content structuring support. All scientific interpretation, data analysis, intellectual content, and final conclusions were independently generated, verified, and approved by the authors.

REFERENCES

1. Sardar S, Sarkar S, Podder I, Banerjee G. Clinico-epidemiology and histopathologic spectrum of primary scarring alopecia: a cross-sectional study. *Indian J Dermatol.* 2022;67(5):624. doi: 10.4103/ijd.ijd_540_21.
2. Hordinsky M. Scarring alopecia: diagnosis and new treatment options. *Dermatol Clin.* 2021;39(3):383-388. doi: 10.1016/j.det.2021.05.001.
3. Qi S, Zhao Y, Zhang X, Li S, Cao H, Zhang X. Clinical features of primary cicatricial alopecia in Chinese patients. *Indian J Dermatol Venereol Leprol.* 2014;80(4):306-312. doi: 10.4103/0378-6323.136833.
4. Kurt B, Şenol Y, Dicle Ö. Clinical Evaluation of Primary Cicatricial Alopecias from Turkey: A Retrospective Analysis of 97 Cases. *Turkısh J Dermatol.* 2018;12(4):177-182. doi: 10.4274/tdd.3679.
5. Whiting DA. Cicatricial alopecia: clinico-pathological findings and treatment. *Clin Dermatol.* 2001;19(2):211-225. doi: 10.1016/s0738-081x(00)00132-2.
6. Olsen EA, Bergfeld WF, Cotsarelis G, et al. Summary of North American Hair Research Society (NAHRS)-sponsored Workshop on Cicatricial Alopecia, Duke University Medical Center, February 10 and 11, 2001. *J Am Acad Dermatol.* 2003;48(1):103-110. doi: 10.1067/mjd.2003.68.
7. Su HJ, Cheng AY, Liu CH, et al. Primary scarring alopecia: A retrospective study of 89 patients in Taiwan. *J Dermatol.* 2018;45(4):450-455. doi: 10.1111/1346-8138.14217.
8. Hess DR. Retrospective studies and chart reviews. *Respir Care.* 2004;49(10):1171-1174. PMID: 15447798.
9. Vañó-Galván S, Saceda-Corralo D, Blume-Peytavi U, et al. Frequency of the types of alopecia at twenty-two specialist hair clinics: a multicenter study. *Skin Append Disord.* 2019;5(5):309-315. doi: 10.1159/000496708.
10. Hashmi AA, Rashid K, Ali R, et al. Clinicopathological features of alopecia with an emphasis on etiology and histopathological characteristics of scarring alopecia. *Cureus.* 2022;14(8):e27661. doi: 10.7759/cureus.27661.

https://doi.org/10.5455/annalsmedres.2025.09.260

11. Beheshtiroy A, Hajmanoochehri F, Hossienghamar F. An Epidemiological Study of 97 Cases of Primary Cicatricial Alopecia in Iran. *Dermatol Reports.* 2015;7(2):5960. doi: 10.4081/dr.2015.5960.
12. Akdoğan N, Karaduman A, Evans SE, et al. Primary lymphocytic cicatricial alopecia: a retrospective analysis of 36 patients. *Istanbul Med J.* 2017;18(4):218-221. doi: 10.5152/imj.2017.32767.
13. Moure ER, Romiti R, Machado MC, Valente NY. Primary cicatricial alopecias: a review of histopathologic findings in 38 patients from a clinical university hospital in Sao Paulo, Brazil. *Clinics (Sao Paulo).* 2008;63(6):747-752. doi: 10.1590/s1807-59322008000600007.
14. Amato L, Mei S, Massi D, Gallerani I, Fabbri P. Cicatricial alopecia; a dermatopathologic and immunopathologic study of 33 patients (pseudopelade of Brocq is not a specific clinico-pathologic entity). *Int J Dermatol.* 2002;41(1):8-15. doi: 10.1046/j.1365-4362.2002.01331.x.
15. Nasimi M, Ahangari N, Lajevardi V, Mahmoudi H, Ghodsi SZ, Etesami I. Quality of life and mental health status in patients with lichen planopilaris based on Dermatology Life Quality Index and General Health Questionnaire-28 questionnaires. *Int J Womens Dermatol.* 2020;6(5):399-403. doi: 10.1016/j.ijwd.2020.09.005.
16. Tan E, Martinka M, Ball N, Shapiro J. Primary cicatricial alopecias: clinicopathology of 112 cases. *J Am Acad Dermatol.* 2004;50(1):25-32. doi: 10.1016/j.jaad.2003.04.001.
17. Doche I, Romiti R, Hordinsky M.K, Valente NS. "Normal-appearing" scalp areas are also affected in lichen planopilaris and frontal fibrosing alopecia: An observational histopathologic study of 40 patients. *Exp Dermatol.* 2020;29(3):278-281. doi: 10.1111/exd.13834.
18. Moreno-Ramírez D, Camacho Martínez F. Frontal fibrosing alopecia: a survey in 16 patients. *J Eur Acad Dermatol Venereol.* 2005;19(6):700-705. doi: 10.1111/j.1468-3083.2005.01291.x.
19. Zhang M, Zhang L, Rosman IS, Mann CM. Frontal fibrosing alopecia demographics: a survey of 29 patients. *Cutis.* 2019;103(2):E16-E22. PMID: 30893399.
20. Boghosian T, Bermudez N, Tosti A. The 100 Most-Cited Publications in Lichen Planopilaris: A Bibliometric Analysis. *Skin Appendage Disord.* 2025;11(3):245-254. doi: 10.1159/000542898.
21. Özcan D, Özen Ö, Seçkin D. Diagnostic value of histopathologic examination in alopecias. *Turkderm.* 2014;48(Suppl 1):24-30. doi: 10.4274/turkderm.48.s6.
22. Bernárdez C, Molina-Ruiz AM, Requena L. Histologic features of alopecias: part II: scarring alopecias. *Actas Dermosifiliogr.* 2015;106(4):260-270. doi: 10.1016/j.ad.2014.06.016.
23. Rigopoulos D, Stamatios G, Ioannides D. Primary scarring alopecias. *Curr Probl Dermatol.* 2015;47:76-86. doi: 10.1159/000369407.
24. Inchara YK, Tirumalae R, Kavdia R, Antony M. Histopathology of scarring alopecia in Indian patients. *Am J Dermatopathol.* 2011;33(5):461-467. doi: 10.1097/DAD.0b013e318201abcd.

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

A comparative analysis of large language models in digital content on painless labor: A systematic evaluation of readability, reliability, quality, and accuracy

Ahmet Ridvan Dogan[a], *, Havva Kocayigit[b], Ayca Tas Tuna[b]

[a]Sakarya Training and Research Hospital, Department of Anesthesiology and Reanimation, Sakarya, Türkiye
[b]Sakarya University, Faculty of Medicine, Department of Anesthesiology and Reanimation, Sakarya, Türkiye
*Corresponding author: a.ridvandogan@gmail.com (Ahmet Ridvan Dogan)

MAIN POINTS

- Large language models, such as ChatGPT, Gemini, and DeepSeek, are increasingly used for patient education and health communication.
- The readability and reliability of AI-generated medical information vary widely across platforms.
- This study provides the first systematic comparison of these three models in the context of information on painless labor.
- It identifies distinct strengths and weaknesses: ChatGPT excels in readability, Gemini excels in quality and reliability, and DeepSeek exhibits variable performance.
- The findings emphasize the need for verifiable references and audience-specific optimization in AI-based health communication.

Cite this article as: Dogan AR, Kocayigit H, Tas Tuna A. A comparative analysis of large language models in digital content on painless labor: A systematic evaluation of readability, reliability, quality, and accuracy. *Ann Med Res*. 2026;33(6):267–276. doi: 10.5455/annalsmedres.2025.11.346.

ABSTRACT

Aim: Painless labor is an important concern in obstetric care, with both medical and psychosocial dimensions. With the increasing use of artificial intelligence (AI) in health communication, evaluating the quality of content generated by large language models (LLMs) is essential. This study aimed to compare AI-generated content on "painless labor" produced by three LLMs---ChatGPT (Chat Generative Pre-trained Transformer), Gemini, and DeepSeek---in terms of readability, reliability, content quality, and medical accuracy.

Materials and Methods: A total of 270 texts were generated using 30 frequently searched keywords on three dates in May 2025 via the free versions of the three LLMs. Readability was assessed with six validated indices. Reliability and quality were evaluated using the Modified DISCERN tool, the JAMA (Journal of the American Medical Association) Benchmark Criteria, the Ensuring Quality Information for Patients (EQIP)-36, and the Global Quality Score. The medical accuracy of 90 texts was assessed by an obstetric anesthesia expert.

Results: ChatGPT texts were the most readable and were written at lower grade levels, but scored lower in reliability and quality. Gemini ranked highest for both reliability and content quality, although it produced more complex language. DeepSeek showed variable performance. No significant differences were found in medical accuracy or content completeness. Gemini also demonstrated the most consistent performance across all time points.

Conclusion: LLMs vary substantially in how they present medical content. ChatGPT achieved higher readability scores, indicating simpler language structures, whereas Gemini demonstrated higher reliability and content quality metrics. However, the absence of source citations across all models raises concerns about the verifiability of content, highlighting the need for critical oversight in healthcare applications.

Keywords: Painless labor, Large language models, Readability, Health literacy, Patient education, Obstetric anesthesia

Received: Nov 10, 2025 **Accepted:** Jan 05, 2026 **Available Online:** Jun 25, 2026

INTRODUCTION

Painless labor aims to enhance the childbirth experience by reducing pain through pharmacological or non-pharmacological methods [1]. Expectations and anxiety during pregnancy influence delivery preferences, including cesarean rates [2].

In recent years, patients increasingly search online for labor analgesia options—such as epidural or spinal techniques—prior to consulting healthcare professionals [3,4]. However, online content varies significantly in readability and reliability [5].

Artificial intelligence (AI)-based large language models (LLMs), including ChatGPT (Chat Generative Pre-trained Transformer), Gemini, and DeepSeek, have transformed access to health information by providing conversational, simplified answers to complex queries [6]. Despite their growing use, the scientific accuracy, readability, and reliability of AI-generated content remain largely unexamined [7]. Especially for individuals with limited health literacy, this raises concerns about misinformation and content

quality [8].

Readability is a key parameter for digital health communication. Leading institutions recommend that patient materials be understandable to readers with basic literacy [9]. Individuals with low health literacy often struggle with technical texts, which may negatively affect decision-making [10]. In this context, painless labor serves as an ideal topic for evaluating LLMs in terms of readability, reliability, quality, and medical accuracy. Unlike many general medical topics, painless labor is a preference-sensitive and emotionally charged clinical context where preconsultation information may directly influence patients' attitudes and decision-making. Therefore, evaluating AI-generated content in this domain addresses not only informational quality but also potential clinical and psychosocial implications.

ChatGPT, Gemini, and DeepSeek differ in architecture, data sources, and citation capabilities [11–13]. These distinctions justify a comparative analysis of their responses to a sensitive and frequently searched medical topic. Beyond readability, the reliability of content (e.g., source usage, authorship, currency) and its quality (relevance, objectivity) are critical for evaluating health information [14]. This study aims to systematically evaluate texts generated by these LLMs on the topic of painless labor across four domains: readability, reliability, quality, and medical accuracy.

MATERIALS AND METHODS

Study design and ethical approval

This cross-sectional, comparative study was conducted in May 2025 using experimental content analysis. No human participants were involved, and only publicly accessible outputs from AI models were used; thus, ethical approval was not required.

Keyword selection and data collection

Browser history and cookies were cleared, and all queries were performed in incognito mode to avoid personalization bias. Google Trends was used to identify relevant keywords by searching for the terms "epidural delivery," "painless labor," and "painless delivery." After filtering to remove irrelevant results, 30 keywords were selected.

Each keyword was submitted to three AI models—ChatGPT (GPT-3.5, OpenAI), Gemini (Gemini 1.5, Google Bard), and DeepSeek (DeepSeek-V3)—on three dates: May 11, 18, and 25, 2025. All queries were conducted using new browser sessions in incognito mode. All queries for each predefined date were conducted in a single session within the same time window to minimize temporal variability. DeepSeek required account registration; therefore, a separate account was created. In total, 270 response texts were collected (3 models × 3 dates × 30 keywords).

Responses were saved in Microsoft Word 2016 after removing formatting elements, such as lists and bullet points. Full dataset was archived at: https://archive.org/details/all_answers_LLMs.

Readability analysis

Each response was uploaded to https://readabilityformulas.com, and analyzed using six readability indices: Automated Readability Index (ARI), Flesch Reading Ease Score (FRES), Gunning Fog Index (GFO), Flesch–Kincaid Grade Level (FKGL), Simple Measure of Gobbledygook (SMOG), and FORCAST Readability Formula (FORCAST) (Table 1). Word and sentence counts were also recorded. All texts were analyzed by the same researcher using a standardized method and were subsequently entered into the statistical dataset.

Reliability assessment

Two independent researchers evaluated the texts using two validated tools. The Modified DISCERN tool (0–5 points) assessed source credibility, balance, and clarity [15]. The JAMA (Journal of the American Medical Association) Benchmark Criteria (0–4 points) evaluated authorship, attribution, update date, and information transparency [16]. In the event of disagreement, a third researcher resolved any discrepancies.

Quality assessment

Text quality was assessed using the Ensuring Quality Information for Patients (EQIP)-36 scale and the Global Quality Score (GQS) [15]. EQIP-36 includes 36 binary items, and scores were categorized as high (76–100%), moderate (51–75%), low (26–50%), or poor (0–25%). GQS is a 1–5 scale reflecting overall usefulness to patients. Discrepancies were resolved through third-party consensus. Researchers used standardized forms and reviewed the responses in different orders to minimize bias.

Medical accuracy assessment

A random sample of 90 texts (30 per model) was selected using a Python-based algorithm. All identifying metadata were removed, and evaluations were conducted blindly by a specialist in obstetric anesthesia. Each text was scored 1–10 for scientific accuracy and 1–10 for content comprehensiveness, following prior methodologies [17]. Obvious medical errors and fabricated references were identified through qualitative assessment.

Statistical analysis

Analyses were performed using IBM SPSS v25.0. Descriptive statistics were reported as mean ± standard deviation, median (minimum–maximum), and frequency (%). The distribution of the data was assessed using the Shapiro–Wilk test. Kruskal–Wallis tests were used for group comparisons, followed by Bonferroni-corrected Mann–Whitney U tests when

https://doi.org/10.5455/annalsmedres.2025.11.346

appropriate. Chi-square tests were applied to categorical variables. The Friedman test was used to assess variation in performance across dates. A p-value < 0.05 was considered statistically significant, with adjustments for multiple comparisons where necessary.

RESULTS

A total of 30 English keywords frequently associated with "epidural delivery," "painless labor," and "painless delivery" were identified using Google Trends. Repetitive or semantically overlapping terms were excluded, and the most contextually relevant expressions were retained (Table 2).

Each keyword was queried using three AI models—ChatGPT, Gemini, and DeepSeek—on three separate dates (May 11, 18, and 25, 2025), resulting in a total of 270 response texts.

Comparison of readability indices

Significant differences were found for all readability indices, word counts, and sentence counts across the three models (all p<0.001) (Table 3). Gemini and DeepSeek had significantly higher ARI scores than ChatGPT (p<0.01). For FRES, ChatGPT scored highest, followed by DeepSeek and Gemini (p<0.05 to p<0.001). In GFI, Gemini had the highest scores; there was no difference between ChatGPT and DeepSeek.

FKGL scores were ranked as follows: Gemini > ChatGPT > DeepSeek (all p<0.001). Gemini also had the highest SMOG scores, while DeepSeek had the highest FORCAST scores (all p<0.001). Gemini produced significantly longer texts, in terms of both word and sentence counts, than the other two models. DeepSeek responses were also longer than ChatGPT's in terms of sentence count (all adjusted p<0.05).

In grade-level classifications, Gemini's content was consistently grouped into higher difficulty levels by ARI, GFI, FKGL, and SMOG (Figure 1). ChatGPT exhibited the lowest difficulty levels across most indices. FORCAST scores indicated that DeepSeek had higher complexity than both Gemini and ChatGPT (p<0.01). Detailed pairwise comparisons are provided in Supplementary Table 1.

Comparison of reliability and quality assessments

Significant differences between the three AI models were observed across all four evaluation metrics (all p<0.001). ChatGPT had significantly lower Modified-DISCERN scores than DeepSeek and Gemini, both of which showed similar reliability levels. In the JAMA Benchmark assessment, DeepSeek ranked highest, followed by Gemini; ChatGPT scored significantly lower than both. Gemini achieved the highest EQIP quality scores, significantly outperforming ChatGPT and DeepSeek. In GQS evaluations, both Gemini and DeepSeek scored significantly higher than ChatGPT, with no difference between them (Figure 2). Detailed pairwise comparisons are provided in Supplementary Table 2.

Time-based analyses

Temporal analyses using the Friedman test revealed some model-specific trends across the three data collection dates. For ChatGPT, both word and sentence counts significantly increased over time (p<0.01). ARI, FKGL, and SMOG scores rose between May 11 and May 18, then slightly decreased on May 25, indicating fluctuations in complexity. FRES scores dipped on May 18 and rebounded on May 25, while FORCAST scores gradually declined (p<0.01). No significant changes were found in reliability or quality scores across the three dates. In DeepSeek, only EQIP scores increased significantly over time (p<0.05), while other metrics remained stable. Gemini showed no significant changes in any metric, indicating consistent performance across all time points (Table 4).

Comparison of medical accuracy assessments

Median scores for both scientific accuracy and content completeness were similar across the three models, with no statistically significant differences (Table 5). All models demonstrated comparable medical accuracy.

DISCUSSION

This study systematically compared the content generated on the topic of "painless labor" by three popular large language models (ChatGPT, Gemini, and DeepSeek) across metrics of readability, reliability, quality, and medical accuracy. The findings demonstrated statistically significant differences among the models regarding readability, reliability, and quality. Beyond confirming previously reported performance patterns of large language models, the present study extends the existing literature by contextualizing these differences in painless labor, a preference-sensitive, emotionally charged clinical scenario. In addition, the time-based evaluation was designed as an exploratory assessment of output stability, highlighting consistency differences between models rather than serving as a primary outcome measure.

In terms of readability, ChatGPT produced texts with lower grade levels and greater ease of comprehension. In contrast, Gemini and DeepSeek generated longer texts with higher technical content, requiring more advanced reading levels. Regarding reliability, ChatGPT received the lowest scores, whereas DeepSeek and Gemini received higher scores. Regarding quality metrics, Gemini produced the most balanced and structured content for patient information, whereas DeepSeek showed competitive performance on certain quality criteria but occasionally yielded variable results in content consistency.

In time-based analyses, the Gemini model was distinguished by producing consistent output, whereas ChatGPT exhibited a more variable performance; however, the observed fluctuations were small in magnitude and are unlikely to be clinically or practically significant. In assessments of medical accuracy, no significant differences were observed among the three models; all produced similar scores for accuracy and content

https://doi.org/10.5455/annalsmedres.2025.11.346

Table 1. Readability indices used in the study, their formulas, and conceptual descriptions.

Index Name	Formula	Description
ARI (Automated Readability Index)	4.71 × (characters/word) + 0.5 × (words/sentence) -- 21.43	Based on average character count and sentence length; provides an approximate U.S. educational grade level.
FRES (Flesch Reading Ease Score)	206.835 -- 1.015 × (words/sentence) -- 84.6 × (syllables/word)	Higher scores indicate easier readability (range: 0--100).
GFO (Gunning Fog Index)	0.4 × [(words/sentence) + % complex words]	Estimates the number of years of education required to understand the text.
FKGL (Flesch-Kincaid Grade Level)	0.39 × (words/sentence) + 11.8 × (syllables/word) -- 15.59	Indicates the educational grade level needed to comprehend the text.
SMOG (Simple Measure of Gobbledygook)	1.0430 × √(number of polysyllabic words × (30/number of sentences)) + 3.1291	Recommended especially for health literacy; more reliable for texts exceeding 100 words.
FORCAST	20 -- (number of single-syllable words per 10 words)	Focuses solely on word level, not sentence structure; particularly suitable for documents other than prose text.

Complex Word: A word consisting of three or more syllables. Grade Level: The educational level required to comprehend the text. For example, "8th grade level" means the text can be understood by an 8th-grade student.

Table 2. List of the 30 most relevant English keywords associated with the terms "epidural delivery," "painless labor," and "painless delivery" according to Google Trends data (2004--May 2025). (Source: Google Trends, https://trends.google.com).

Keywords		
birth with epidural	epidural for normal delivery	painless delivery
birth without epidural	epidural labor	painless delivery cost
can contractions be painless	epidural painless delivery	painless delivery side effects
delivery pain	epidural pregnancy	painless labor
epidural anesthesia	epidural side effects	painless normal delivery
epidural birth	labor analgesia	painless pregnancy
epidural delivery	labor pain	what is epidural
epidural during labor	labor with epidural	what is epidural birth
epidural for birth	painless birth	what is epidural delivery
epidural for labor	painless contractions	what is painless delivery

Table 3. Comparison of readability indices, word count, and sentence count across artificial intelligence models.

Readability Index	ChatGPT (Median, (Min--Max))	DeepSeek (Median, (Min--Max))	Gemini (Median, (Min--Max))	p *	ChatGPT vs DeepSeek p †	ChatGPT vs Gemini p †	DeepSeek vs Gemini p †
ARI	12.71 (8.24--17.19)	13.33 (10.08--21.87)	13.34 (10.67--16.93)	**0.001**	**<0.001**	**0.002**	0.674
FRE	42.69 (27.85--60.67)	39.64 (2.40--54.84)	36.79 (19.16--52.02)	**<0.001**	<0.032	**<0.001**	**0.005**
GFI	9.79 (6.62--14.91)	10.08 (7.49--14.22)	10.49 (8.48--12.99)	**<0.001**	<0.166	**<0.001**	**0.001**
FKGL	11.70 (7.85--16.01)	10.23 (7.84--16.00)	12.52 (9.96--15.78)	**<0.001**	**<0.001**	**<0.001**	**<0.001**
SMOG	13.02 (8.69--16.18)	11.12 (9.35--15.05)	13.51 (11.57--16.26)	**<0.001**	**<0.001**	**<0.001**	**<0.001**
FORCAST	11.58 (10.30--14.11)	12.29 (10.77--14.13)	11.89 (10.73--12.93)	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Word Count	332.50 (82--775)	290.50 (134--424)	526.50 (159--1144)	**<0.001**	0.069	**<0.001**	**<0.001**
Sentence Count	18.00 (2--55)	28.00 (9--47)	30.00 (7--57)	**<0.001**	**<0.001**	**<0.001**	**0.004**

* Overall group comparisons were performed using the Kruskal–Wallis test. † For measures with significant differences, pairwise comparisons were conducted using the Bonferroni-corrected Mann–Whitney U test (significance level $p < 0.0167$). Abbreviations: ARI = Automated Readability Index; FRES = Flesch Reading Ease Score; GFI = Gunning Fog Index; FKGL = Flesch-Kincaid Grade Level; SMOG = Simple Measure of Gobbledygook; FORCAST = FORCAST Readability Formula.

https://doi.org/10.5455/annalsmedres.2025.11.346

Table 4. Comparison of temporal changes in readability, reliability, and quality metrics of artificial intelligence models (May 11–25, 2025).

Metric	ChatGPT (Median (min-max))			DeepSeek (Median (min-max))			Gemini (Median (min-max))		
	11.May.25	18.May.25	25.May.25	11.May.25	18.May.25	25.May.25	11.May.25	18.May.25	25.May.25
Word Count	210 (110–416)	315 (115–517)	385 (82–775)	278 (162–420)	284 (134–424)	267 (154–404)	520 (219–822)	528 (159–1144)	523 (164–796)
p *		**<0.01**			0.611			0.909	
Sentence Count	16 (2–28)	16 (6–29)	21 (4–55)	25 (12–37)	28 (9–43)	25 (12–47)	30 (12–48)	30 (9–57)	31 (7–50)
p *		**0.003**			0.57			0.995	
ARI Score	11.42 (8.40–14.56)	13.77 (10.23–17.19)	12.65 (8.24–15.28)	14.09 (10.08–19.77)	13.67 (10.29–18.07)	14.23 (11.53–21.87)	13.65 (10.67–16.93)	13.74 (11.46–16.18)	13.15 (10.95–16.80)
p *		**<0.01**			0.671			0.225	
FRES Score	42.83 (28.26–55.58)	39.50 (27.85–49.99)	43.61 (34.61–60.67)	38.74 (14.27–54.44)	39.55 (13.05–54.84)	38.17 (2.40–51.41)	35.28 (19.16–52.02)	35.56 (20.89–44.98)	37.86 (20.81–49.51)
p *		**0.024**			0.86			0.266	
Gunning Fog Score	9.85 (7.39–14.91)	9.92 (6.62–11.47)	9.62 (6.71–12.56)	10.14 (7.68–13.37)	10.12 (7.49–13.00)	9.99 (7.89–14.22)	10.61 (8.48–12.99)	10.65 (8.78–12.62)	10.47 (8.91–12.88)
p *		0.682			0.887			0.742	
FKGL Score	10.74 (8.38–13.87)	12.76 (10.21–16.01)	11.77 (7.85–14.10)	10.71 (7.84–14.60)	10.42 (8.11–15.14)	10.66 (8.36–16.00)	12.71 (9.96–15.78)	12.75 (11.16–15.32)	12.30 (10.04–15.73)
p *		**<0.01**			0.788			0.301	
SMOG Grade	11.81 (8.69–15.90)	13.71 (11.36–16.18)	13.10 (9.97–15.32)	11.51 (9.44–14.55)	11.27 (9.35–14.39)	11.25 (9.40–15.05)	13.64 (11.57–15.56)	13.74 (12.24–15.90)	13.45 (11.94–16.26)
p *		**<0.01**			0.656			0.49	
FORCAST Score	11.85 (10.54–14.11)	11.46 (10.30–12.24)	11.38 (10.51–12.14)	12.31 (11.26–14.13)	12.25 (10.77–13.17)	12.29 (11.45–13.87)	11.99 (10.73–12.93)	11.87 (11.08–12.69)	11.81 (10.99–12.57)
p *		**0.002**			0.888			0.301	
M-DISCERN	1.90 (1–3)	1.93 (1–3)	2.00 (1–3)	2.87 (2–4)	3.03 (2–4)	3.30 (2–4)	2.80 (2–4)	3.20 (2–4)	3.13 (2–4)
p *		0.898			0.126			0.12	
JAMA	0.50 (0–1)	0.47 (0–1)	0.37 (0–1)	1.37 (1–2)	1.63 (1–2)	1.43 (1–2)	1.00 (0–2)	1.13 (0–2)	0.93 (0–2)
p *		0.566			0.1			0.604	
EQIP (%)	53.7% (30.6–75.0%)	57.3% (33.3–77.8%)	54.4% (33.3–77.8%)	54.3% (33.3–69.4%)	58.6% (33.3–80.6%)	61.8% (38.9–77.8%)	63.4% (47.2–83.3%)	63.7% (47.2–86.1%)	63.4% (44.4–83.3%)
p *		0.479			**0.033**			0.99	
GQS	3.03 (2–4)	2.83 (2–4)	3.00 (2–4)	3.27 (2–4)	3.53 (2–4)	3.20 (2–4)	3.37 (2–4)	3.50 (2–4)	3.27 (2–4)
p *		0.583			0.125			0.378	

* Friedman test. Abbreviations: ARI: Automated Readability Index; FRES: Flesch Reading Ease Score; GFI: Gunning Fog Index; FKGL: Flesch-Kincaid Grade Level; SMOG: Simple Measure of Gobbledygook; FORCAST: FORCAST Readability Formula; M-DISCERN: Modified DISCERN; JAMA: Journal of the American Medical Association ;EQIP: Ensuring Quality Information for Patients; GQS: Global Quality Score.

Table 5. Comparison of medical accuracy and content completeness metrics among artificial ıntelligence models.

Metric	ChatGPT (Median, (Min–Max))	DeepSeek (Median, (Min-Max))	Gemini (Median, (Min–Max))	p *
Accuracy	8 (5-9)	8 (5-8)	8 (5-9)	0.765
Completeness	7 (5-8)	7 (5-7)	7 (5-8)	0.938

* Overall group comparisons were performed using the Kruskal–Wallis test.

completeness. Overall, the study highlighted notable differences among models in the technical sophistication, comprehensibility, reliability, and accuracy of their AI-generated content.

The ChatGPT model produced the simplest texts and at the lowest grade levels in terms of readability, a finding consistent with previous studies reporting that ChatGPT achieves the highest readability scores in patient information materials [18,19]. However, those studies also found ChatGPT inadequate in content comprehensiveness and guidance, noting that its simplified language reduced the depth of information. Similarly, in our study, although ChatGPT produced simpler

https://doi.org/10.5455/annalsmedres.2025.11.346

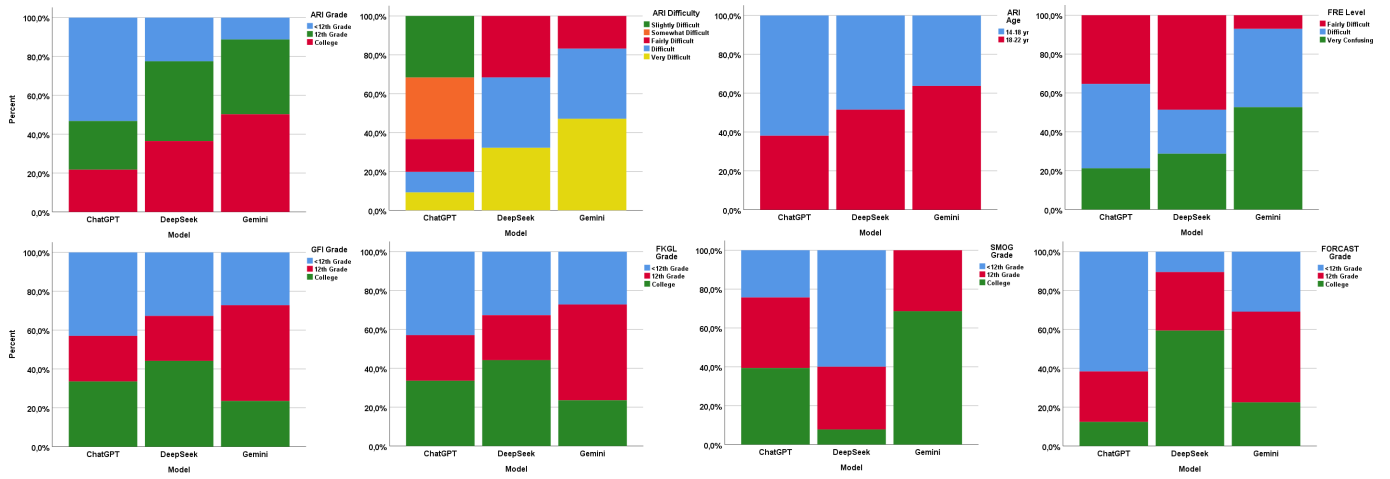

Figure 1. Grade-level classifications of texts generated by ChatGPT, DeepSeek, and Gemini across readability indices (ARI, ARI Difficulty, ARI Age, FRES, GFI, FKGL, SMOG, FORCAST). Bars represent the percentage distribution of responses in each grade-level category. Significant differences were observed across models for all indices except certain ARI subcategories (Chi-suqare tests, p<0.01). Detailed pairwise comparisons are provided in Supplementary Table 1.

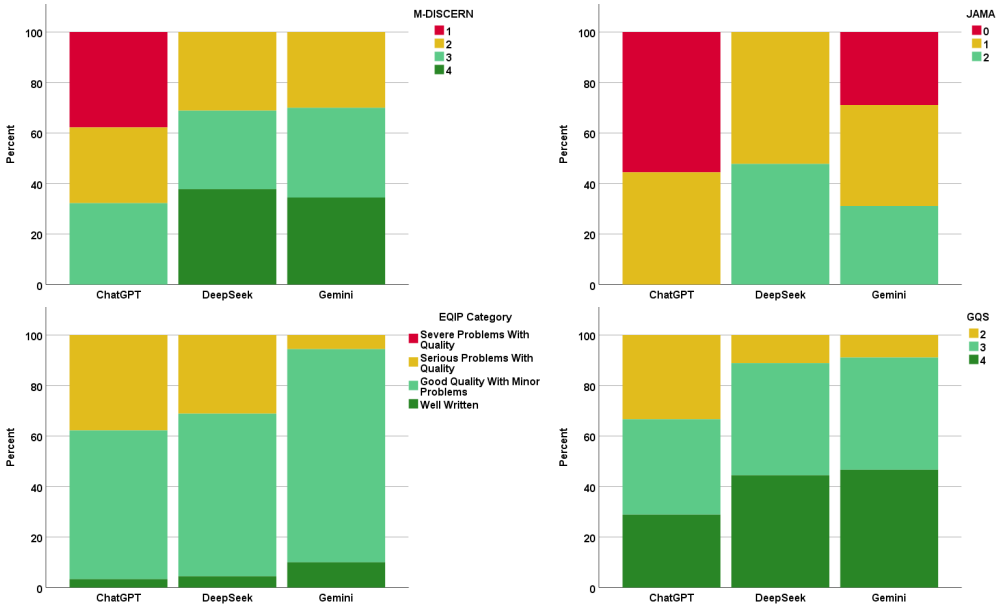

Figure 2. Reliability and quality assessments of AI-generated texts across models. Stacked bar charts show the percentage distribution of scores for M-DISCERN (top left), JAMA Benchmark Criteria (top right), EQIP-36 (bottom left), and Global Quality Score (GQS, bottom right) among ChatGPT, DeepSeek, and Gemini. Significant differences were observed across all four evaluation metrics (Chi-square tests, p<0.001). Detailed pairwise comparisons are provided in Supplementary Table 2.

texts, it scored significantly lower on reliability and quality metrics. This reflects the challenge of balancing simplification of information delivery with preservation of clinical value in the content. Notably, despite these relative differences, the absolute readability levels of outputs across all evaluated models predominantly corresponded to college-grade reading levels (FKGL and SMOG > 12), which remain well above the recommended $\leq 8^{th}$ grade level for patient education materials.

In our study, the Gemini model achieved the highest scores on reliability (JAMA, M-DISCERN) and quality (EQIP, GQS) metrics. This aligns with previous studies emphasizing Gemini's capacity to produce patient-friendly and balanced content However, those studies also reported that Gemini tends to produce higher Gunning Fog and SMOG scores, indicating that the texts are structurally more complex, which may limit comprehensibility. Similarly, in our study, Gemini generated content at the highest grade levels, particularly according to SMOG and FKGL indices, thereby reducing readability. This discrepancy may stem from Gemini's tendency to produce more comprehensive, explanatory content, which may be less accessible to individuals with limited health literacy.

Findings related to the DeepSeek model revealed a more het-

https://doi.org/10.5455/annalsmedres.2025.11.346

erogeneous profile, consistent with both the literature and the internal analyses of our study. Previous studies have reported that DeepSeek offers more technical content and provides more frequent reference links but has been criticized for higher hallucination rates and content inconsistencies [20,21]. In our study, DeepSeek demonstrated moderate performance in terms of readability but produced more variable results in quality scores compared to both ChatGPT and Gemini. Possible reasons include the model's open-source structure, incomplete optimization of its document-linking capabilities, and lack of training specifically focused on structured health-related content.

In this study, the medical accuracy and completeness scores of the three models were similar; however, the lack of source citations by any of the models limited the verifiability of the content. This finding aligns with previous reports indicating that although ChatGPT often demonstrates high content accuracy, it lacks transparency in referencing and occasionally includes misleading information [22]. Similarly, earlier studies have reported that while GPT-4 can maintain accuracy and comprehensiveness even in simplified responses, models like Bard tend to experience content loss during simplification [19]. In our study, the Gemini and DeepSeek models produced accuracy scores comparable to ChatGPT, but, like ChatGPT, did not provide sources to support clinically critical information. The absence of references poses a significant risk to patient safety, as texts may appear accurate but remain unverifiable. The findings underscore that large language models still require careful oversight for generating medical content and are not suitable as standalone tools for clinical decision support. It should be noted that instruments such as the Modified DISCERN and JAMA Benchmark Criteria were originally developed to evaluate static, authored health websites rather than conversational AI-generated outputs. Therefore, lower reliability scores may partly reflect limitations in transparency and attribution inherent to current general-purpose LLMs, rather than definitive deficits in the factual correctness of the information provided.

Previous studies focusing on obstetric anesthesia have demonstrated that large language models have significant potential for patient education, although this potential varies substantially with respect to content quality, source transparency, and readability. In a study on epidural analgesia during labor, ChatGPT was reported to perform well in terms of accuracy (97.7%) and content quality; however, the texts were generally written above the 11th-grade reading level [23]. This finding aligns with our study, in which ChatGPT, despite achieving high accuracy scores, offered only a relative advantage in readability. Another study found that both ChatGPT and Gemini provided similar accuracy rates (87–85%) when responding to frequently asked questions about childbirth; however, the Gemini model was considered superior in terms of using simpler language and delivering more actionable content [5]. In contrast, our study found that Gemini produced longer and more complex texts while achieving the highest quality scores. This difference may be attributable to the nature of the questions used in our study, which were neither directly instructive nor action-oriented, and to Gemini's tendency to offer more comprehensive explanations. In a comparative study involving American Society of Anesthesiologists (ASA) materials, Gemini outperformed ChatGPT in terms of accuracy and comprehensiveness; however, both models were found to produce more challenging texts in terms of readability compared to ASA content [17]. Similarly, in our study, the texts generated by Gemini and DeepSeek were classified at significantly higher grade levels than those recommended for health literacy. Another analysis focusing on pregnancy-related questions showed that the Gemini model achieved a higher rate of acceptable answers (83%) compared to ChatGPT (58%) and also demonstrated significantly fewer errors in source citation [24]. However, in our study, none of the evaluated models provided references, resulting in equivalent limitations regarding verifiability. In a comparative analysis focusing on obstetric pathologies, ChatGPT was found to perform significantly better than Gemini in terms of understandability and actionability [25]. Conversely, in our study, the Gemini model achieved the highest scores on the GQS and EQIP metrics, while ChatGPT exhibited lower reliability scores. This discrepancy may be due to the specific patient scenarios and prompts used in different studies, which can directly influence the text structure and content strategies of the models. Overall, consistent with findings from other studies in the obstetric field, our study highlights that, while the technical accuracy of these models is generally high, significant differences exist among models in user-centered aspects such as presentation style, readability, and content organization.

These findings from the literature largely align with the main results of our study. However, previous research has often focused on broader pregnancy-related topics rather than on specific themes such as labor analgesia. The unique aspect of our study lies in analyzing the content-generation performance of AI models on the topic of painless labor—a subject with both medical and socio-cultural implications—and in systematically revealing differences in performance through a large dataset that also includes temporal consistency assessments.

The model-based differences observed in this study are related not only to the content output but also to the structural and functional parameters governing how this output is produced. The tendency of ChatGPT to generate texts that are more readable but with lower reliability and content quality likely stems from how the model balances its prioritization of simplification against information density. Gemini's propensity to produce more technical, longer, and more instructive content may be explained by the diversity of its training data and differences in its source pool [26]. Athough the DeepSeek model demonstrated performance similar to Gemini in terms of readability in our study, it produced more variable and at

https://doi.org/10.5455/annalsmedres.2025.11.346

times superficial responses regarding content coherence and clarity of guidance. This may be attributable to the model's open-source nature, which allows continuous updates, but which has not yet undergone sufficient structured evaluation in the health domain [27]. Additionally, by structuring the evaluation around model–time–text interactions, rather than isolated outputs, this study adopted a multilayered analytical approach that allowed for a more nuanced understanding of model behavior within clinical communication contexts.

Moreover, the differences among models are related to the prompt-processing strategies and response-generation architectures they employ. For example, the tendency of ChatGPT's free versions to produce shorter and more summarized content, due to token limits, may lead to information gaps and superficial narratives [28]. Conversely, models like Gemini, which generate longer responses, may increase information coverage but adversely affect readability scores. Additionally, previous studies have shown that some models are more prone to "hallucination" behavior, incorporating non-existent sources or unverified information [11]. Although these phenomena were not quantitatively assessed in the present study, they represent well-documented concerns in the literature that may influence the perceived reliability of AI-generated health information. Another factor explaining these differences lies in the scope and currency of the datasets used during model training. Models such as ChatGPT-3.5, trained only on data up to 2021, may present certain contemporary medical practices incompletely or inaccurately, whereas newer LLMs like Gemini, with access to broader and more updated data pools, may offer enhanced content coverage [26]. However, if this advantage is not effectively integrated into the model architecture, it can result in texts that lack readability and user-friendliness.

The use of AI-generated texts in sensitive areas such as labor analgesia depends not only on technical accuracy but also on comprehensibility. Our study found that content lacking high readability remains common, potentially hindering information access for individuals with low health literacy [10]. Information that is insufficiently understood can negatively influence patients' decisions, particularly in clinical areas such as labor, where anxiety levels are high. While some models produce technically accurate yet complex content, potentially complicating patient education, overly simplified, superficial responses may create a false sense of confidence among users. Both situations underscore that, beyond accuracy, the manner in which information is presented is equally critical. Previous studies have also highlighted that while online peer-to-peer interactions can enhance access to information and support, they may simultaneously carry risks when content promotes harmful behaviors, underscoring the importance of careful moderation and presentation of health information [29].

Limitations

This study provides a systematic and multidimensional evaluation of large language models on the clinically and socially significant topic of labor analgesia. However, certain limitations should be acknowledged. First, all analyses were based on the free versions of the models; therefore, the results may not reflect the performance of premium versions, such as GPT-4. Second, textual outputs in English were evaluated, limiting generalizability to other languages and to multimedia capabilities. Third, the accuracy assessment was performed by a single expert on a limited subset of responses, which may introduce subjectivity and limit statistical generalizability. Although the texts were randomly selected and assessed in a blinded manner, the absence of multiple independent evaluators and of a standardized accuracy checklist limits findings on medical accuracy. Accordingly, the accuracy results should be interpreted as exploratory and hypothesis-generating rather than definitive. Additionally, the study employed a relative model-to-model comparison design and did not benchmark against established gold-standard patient education materials, which limits the interpretation of their absolute adequacy for clinical use. Furthermore, the use of single-turn, keyword-based prompts does not reflect the iterative and conversational nature of real-world LLM use, potentially underrepresenting the performance of models optimized for multi-turn dialogue. Lastly, evaluations did not include end-user perspectives; how patients perceive and interpret the content remains unknown.

CONCLUSION

This study demonstrates that large language models differ substantially in the informational content they generate about labor analgesia. While ChatGPT offers the highest readability and may therefore be more suitable for patient education, Gemini provides superior reliability and content quality, making it more appropriate for professional or information-dense contexts. DeepSeek demonstrates variable performance, indicating the need for further refinement.

Importantly, none of the models provided source citations, a critical limitation that undermines verifiability and clinical trust. Optimizing LLM outputs for target audiences and establishing mechanisms for transparent referencing will be essential for their safe integration into healthcare communication. Future research should incorporate real patient perspectives, premium model versions, and multilingual analyses to better assess the practical implications of AI-generated medical information.

Acknowledgements: The authors would like to thank Ali İrfan Doğan, an undergraduate student in Management Information Systems at Esenyurt University, for his valuable technical support during the data collection and archiving phases of this study.

https://doi.org/10.5455/annalsmedres.2025.11.346

Ethics Committee Approval: This study did not involve human participants, identifiable personal data, or animals, and therefore did not require approval by an institutional ethics committee.

Informed Consent: This study does not involve human participants, identifiable personal data, or animals.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors declare that they have no competing interests related to the content of this article.

Author Contributions: Conception: A.R.D; Design: A.R.D, H.K; Supervision: H.K, A.T.T; Materials: A.R.D; Data Collection or Processing: A.R.D, H.K, A.T.T; Analysis or Interpretation: A.R.D, H.K, A.T.T; Literature Review: A.R.D, A.T.T; Writing: A.R.D; Critical Review: H.K, A.T.T.

Financial Disclosure: No external funding was received for this study.

Artificial Intelligence Disclosure: The authors declare that an artificial intelligence–based tool (ChatGPT, OpenAI, version 3.5) 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. Nori W, Kassim MAK, Helmi ZR, et al. Non-Pharmacological Pain Management in Labor: A Systematic Review. *J Clin Med.* 2023;12(23):7203. doi: 10.3390/jcm12237203.
2. Fuglenes D, Aas E, Botten G, Øian P, Kristiansen IS. Why do some pregnant women prefer cesarean? The influence of parity, delivery experiences, and fear. *Am J Obstet Gynecol.* 2011;205(1):45.e1-45.e9. doi: 10.1016/j.ajog.2011.03.043.
3. Gjestvang C, Haakstad LAH. Navigating Pregnancy: Information Sources and Lifestyle Behavior Choices—A Narrative Review. *J Pregnancy.* 2024;2024(1):4040825. doi: 10.1155/2024/4040825.
4. Murphy J, Vaughn J, Gelber K, Geller A, Zakowski M. Readability, content, quality and accuracy assessment of internet-based patient education materials relating to labor analgesia. *Int J Obstet Anesth.* 2019;39:82–7. doi: 10.1016/j.ijoa.2019.01.003.
5. Lee D, Brown M, Hammond J, Zakowski M. Readability, quality and accuracy of generative artificial intelligence chatbots for commonly asked questions about labor epidurals: a comparison of ChatGPT and Bard. *Int J Obstet Anesth.* 2025;61:104317. doi: 10.1016/j.ijoa.2024.104317.
6. Ayers JW, Zhu Z, Poliak A, et al. Evaluating Artificial Intelligence Responses to Public Health Questions. *JAMA Netw Open.* 2023;6(6):e2317517. doi: 10.1001/jamanetworkopen.2023.17517.
7. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare (Basel).* 2023;11(6):887. doi: 10.3390/healthcare11060887.
8. Murugesu L, Damman OC, Timmermans DRM, et al. Health literate-sensitive shared decision-making in maternity care: needs for support among maternity care professionals in the Netherlands. *BMC Pregnancy Childbirth.* 2023;23(1):594. doi: 10.1186/s12884-023-05915-9.
9. Kripalani S, Weiss BD. Teaching about health literacy and clear communication. *J Gen Intern Med.* 2006;21(8):888–90. doi: 10.1111/j.1525-1497.2006.00543.x.
10. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low Health Literacy and Health Outcomes: An Updated Systematic Review. *Ann Intern Med.* 2011;155(2):97–107. doi: 10.7326/0003-4819-155-2-201107190-00005.
11. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health.* 2023;2(2):e0000198. doi: 10.1371/journal.pdig.0000198.
12. Islam R, Ahmed I. Gemini-the most powerful LLM: Myth or Truth. 2024 5th Inf Commun Technol Conf ICTC. Nanjing, China, 2024, pp. 303-308. doi: 10.1109/ICTC61510.2024.10602253.
13. Peng Y, Malin BA, Rousseau JF, et al. From GPT to DeepSeek: Significant gaps remain in realizing AI in healthcare. *J Biomed Inform.* 2025;163:104791. doi: 10.1016/j.jbi.2025.104791.
14. Li Y, Ouyang H, Lin G, Peng Y, Yao J, Chen Y. Evaluation of the measurement properties of online health information quality assessment tools: A systematic review. *Int J Nurs Sci.* 2025;12(2):130–6. doi: 10.1016/j.ijnss.2025.02.015.
15. Ozdemir ZM, Yavuz SA, Surmelioglu DG. YouTube as an information source in deep margin elevation: Reliability, accuracy and quality analysis. *PLoS One.* 2025;20(2):e0318568. doi: 10.1371/journal.pone.0318568.
16. Silberg WM, Lundberg GD, Musacchio RA. Assessing, Controlling, and Assuring the Quality of Medical Information on the Internet: Caveant Lector et Viewor—Let the Reader and Viewer Beware. *JAMA.* 1997;277(15):1244–5. doi: 10.1001/jama.1997.03540390074039.
17. Gondode PG, Singh R, Mehta S, Singh S, Kumar S, Nayak SS. Artificial intelligence chatbots versus traditional medical resources for patient education on "Labor Epidurals": an evaluation of accuracy, emotional tone, and readability. *Int J Obstet Anesth.* 2025;61:104302. doi: 10.1016/j.ijoa.2024.104302.
18. Demir M, Bedir N, Boğa D, et al. Comparison of Artificial Intelligence Responses About Penile Prosthesis implantation in Terms of Quality and Readability: A Methodological and Validity Study on ChatGPT, Gemini, and DeepSeek. *J Reconstr Urol.* 2025;15(1):21–6. doi: 10.5336/urology.2025-109438.
19. Srinivasan N, Samaan JS, Rajeev ND, Kanu MU, Yeo YH, Samakar K. Large language models and bariatric surgery patient education: a comparative readability analysis of GPT-3.5, GPT-4, Bard, and online institutional resources. *Surg Endosc.* 2024;38(5):2522–32. doi: 10.1007/s00464-024-10720-2.
20. Aydin S, Karabacak M, Vlachos V, Margetis K. Large language models in patient education: a scoping review of applications in medicine. *Front Med.* 2024;11:1477898. doi: 10.3389/fmed.2024.1477898.
21. Zhou J, Cheng Y, He S, Chen Y, Chen H. Large Language Models for Transforming Healthcare: A Perspective on DeepSeek-R1. *MedComm – Future Med.* 2025;4(2):e70021. doi: 10.1002/mef2.70021.
22. Rahimli Ocakoglu S, Coskun B. The Emerging Role of AI in Patient Education: A Comparative Analysis of the Accuracy of Large Language Models for Pelvic Organ Prolapse. *Med Princ Pract.* 2024;33(4):330–7. doi: 10.1159/000538538.
23. Tan CW, Chan JCY, Chan JJI, Nagarajan S, Sng BL. Information about labor epidural analgesia: an updated evaluation on the readability, accuracy, and quality of ChatGPT responses incorporating patient preferences and complex clinical scenarios. *Int J Obstet Anesth.* 2025;63:104688. doi: 10.1016/j.ijoa.2025.104688.
24. Khromchenko K, Shaikh S, Singh M, Vurture G, Rana RA, Baum JD. ChatGPT-3.5 Versus Google Bard: Which Large Language Model Responds Best to Commonly Asked Pregnancy Questions? *Cureus.* 2024;16(7):e65543. doi: 10.7759/cureus.65543.
25. Khemlani A, Singavarapu J, Vasa R, et al. Comparing the Accuracy of Large Language Models(LLM)in Trending Obstetrical Topics. *Open Access J Gynecol Obstet.* 2025;7(1):18–22. doi: 10.22259/2638-5244.0701003.
26. Rane N, Choudhary S, Rane J. Gemini versus ChatGPT: applications, performance, architecture, capabilities, and implementation. *J Appl Artif Intell.* 2024;5(1):69–93. doi: 10.48185/jaai.v5i1.1052.

27. Temsah A, Alhasan K, Altamimi I, et al. DeepSeek in Healthcare: Revealing Opportunities and Steering Challenges of a New Open-Source Artificial Intelligence Frontier. *Cureus.* 2025;17(2):e79221. doi: 10.7759/cureus.79221.
28. Roberts RH, Ali SR, Hutchings HA, Dobbs TD, Whitaker IS. Comparative study of ChatGPT and human evaluators on the assessment of medical literature according to recognised reporting standards. *BMJ Health Care Inform.* 2023;30(1):e100830. doi: 10.1136/bmjhci-2023-100830.
29. Ferraro G, Loo Gee B, Ji S, Salvador-Carulla, L. Lightme: analysing language in internet support groups for mental health. *Health Inf Sci Syst.* 2020;8(34):1-10. doi: 10.1007/s13755-020-00115-7.

https://doi.org/10.5455/annalsmedres.2025.11.346

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Association between serum chloride levels and 28-day mortality in geriatric septic ICU patients: A retrospective cohort study

Irem Durmus[a,*], Kubra Taskin[a], Merve Bulun Yediyildiz[a], Reyhan Fidan[b]

[a]University of Health Sciences, Kartal Dr. Lütfi Kırdar City Hospital, Department of Anesthesiology and Reanimation, Istanbul, Türkiye
[b]Tuzla State Hospital, Clinic of Anesthesiology and Reanimation, Istanbul, Türkiye
*Corresponding author: irem.durmus89@gmail.com (Irem Durmus)

MAIN POINTS

- Hyperchloremia in elderly septic ICU patients was found to be an independent predictor of 28-day all-cause mortality.
- The frequency of hyperchloremia was significantly higher in non-survivors than in survivors.
- In multivariate regression analysis, hyperchloremia and the APACHE II score were independently associated with mortality risk.
- Routine chloride monitoring during sepsis management may help identify high-risk patients earlier and guide individualized fluid therapy.
- Optimizing chloride balance could improve prognosis and reduce mortality in geriatric intensive care patients with sepsis.

Cite this article as: Durmus I, Taskin K, Bulun Yediyildiz M, Fidan R. Association between serum chloride levels and 28-day mortality in geriatric septic ICU patients: A retrospective cohort study. *Ann Med Res*. 2026;33(6):277–283. doi: 10.5455/annalsmedres.2025.11.344.

ABSTRACT

Aim: This study aimed to examine whether serum chloride concentrations obtained within the first 24 hours after intensive care unit (ICU) admission are associated with 28-day all-cause mortality in elderly patients with sepsis, which was defined according to Sepsis-3 criteria (suspected infection plus ≥2-point increase in SOFA score) and to explore their potential prognostic value.

Materials and Methods: This retrospective cohort study was conducted at a single center and involved 221 patients aged 65 years and older who were hospitalized in the ICU with sepsis between December 1, 2023 and May 31, 2024. Patient categorisation was based on serum chloride levels measured within the first 24 hours. Patients were categorised as hypochloremic (<95 mEq/L), normochloremic (95–110 mEq/L), or hyperchloremic (>110 mEq/L). Information on demographics, laboratory results, and Acute Physiology and Chronic Health Evaluation II (APACHE II) scores was collected. Data analysis involved the Mann-Whitney U test, the Chi-square test, and logistic regression.

Results: A total of 221 patients were evaluated, with a mean age of 80.87 ± 8.04 years. The 28-day all-cause mortality rate was found to be 67.4%. The prevalence of hyperchloremia (62.4%) was significantly higher than that of normochloremia (25.8%) and hypochloremia (11.8%). Multivariate logistic regression analysis showed that hyperchloremia was an independent predictor of 28-day mortality (OR=4.95; 95% CI, 1.51–16.21; p=0.008). A rise in the APACHE II score was found to be a highly significant predictor of mortality (p<0.001).

Conclusion: An independent association between elevated chloride levels and 28-day mortality was observed in elderly septic patients requiring intensive care. Routine monitoring of serum chloride levels and careful management of chloride load may contribute to an improved prognosis in this vulnerable patient population.

Keywords: Sepsis, Chlorides, Mortality, Critical care, Intensive care units, Aged

Received: Nov 25, 2025 **Accepted:** Jan 05, 2026 **Available Online:** Jun 25, 2026

INTRODUCTION

Sepsis represents a condition in which an infection triggers a dysregulated and maladaptive host immune response, leading to widespread inflammation [1]. This aberrant response can disrupt microcirculatory function, impair tissue perfusion and ultimately progress to multi-organ dysfunction. Clinically, a substantial proportion of affected individuals require management in the intensive care unit, and sepsis continues to be recognised as one of the primary contributors to mortality among intensive care unit (ICU) patients [2]. In this study, sepsis was defined according to the Sepsis-3 criteria, requiring suspected infection and a ≥2-point increase in the SOFA score. Geriatric patients were defined as individuals aged ≥65 years [3].

Serum chloride represents the predominant extracellular anion and is essential for multiple physiological processes, including the regulation of acid–base homeostasis, osmotic equilibrium, neuromuscular signalling and various enzyme functions [4,5]. In addition, chloride channels have been implicated in intracellular inflammatory pathways, which are known to contribute substantially to the underlying mech-

anisms of sepsis [6]. Notwithstanding chloride's vital functions, it is frequently not evaluated with the same rigour as sodium or potassium in ICU practice, resulting in its clinical importance being somewhat overlooked.

Recent evidence suggests that disturbances in serum chloride concentrations can contribute to unfavourable clinical outcomes. In septic patients, reduced chloride levels may lead to metabolic alkalosis and may be associated with increased mortality risk. Conversely, hyperchloremia has been shown to induce hyperchloremic metabolic acidosis and renal vasoconstriction, consequently impairing organ function and contributing to mortality [7, 8].

We hypothesized that serum chloride levels measured during the first 24 hours of ICU admission would be independently associated with 28-day mortality in geriatric septic patients.

This study aimed to examine how serum chloride levels, obtained during the first 24 hours of ICU admission, relate to 28-day mortality in patients with sepsis. A secondary aim was to explore whether different chloride categories (hypochloremia, normochloremia, hyperchloremia) might be useful as prognostic indicators in this population.

MATERIALS AND METHODS

Study design and ethical approval

Ethical approval for the study was obtained from the Kartal Koşuyolu High Specialization Training and Research Hospital Clinical Research Ethics Committee (Date: 20/05/2025, Approval No: 2025/08/1128). The research was conducted in accordance with the Declaration of Helsinki. This retrospective cohort investigation was performed at a single centre, the Intensive Care Unit of Tuzla State Hospital, and covered the period from 1 December 2023 to 31 May 2024.

Study population

A total of 221 individuals aged 65 years or older who were admitted to the intensive care unit with a diagnosis of sepsis were included in the study population. Patient information was retrospectively extracted from the electronic medical records. Demographic variables included age and sex. Laboratory parameters included white blood cell (WBC) count, haemoglobin level, platelet count (PLT), red cell distribution width (RDW), albumin, calcium, sodium, potassium, creatinine, lactate, C-reactive protein (CRP), and procalcitonin. In addition, the Acute Physiology and Chronic Health Evaluation II (APACHE II) score and serum chloride values obtained during the first 24 hours after ICU admission were recorded.

Inclusion and exclusion criteria

Patients were eligible for inclusion if they were 65 years of age or older, had a documented diagnosis of sepsis, and had serum chloride levels measured during the initial 24 hours of their ICU stay. Individuals without chloride measurements within this timeframe, those discharged or deceased within the first 24 hours, and those presenting with acute kidney injury prior to ICU admission were excluded from the analysis. To ensure a homogeneous cohort of elderly patients admitted for medical reasons, postoperative and trauma-related admissions were excluded.

Data collection and variables

Serum chloride levels were measured within the first 24 hours after ICU admission and recorded for all patients. The classification of patients was based on the following values: hypochloremia(<95 mEq/L),normochloremia(95–110 mEq/L), and hyperchloremia (>110 mEq/L). Each patient was categorised according to the earliest measured chloride value within the first 24 hours following admission to the ICU.

At the conclusion of the ICU stay, patients were evaluated according to two clinical outcomes: survivors and non-survivors. The principal outcome of interest was overall mortality within 28 days of ICU admission. In addition, comorbidities such as diabetes mellitus, hypertension, coronary artery disease, heart failure, chronic obstructive pulmonary disease, cerebrovascular disorders, and chronic kidney disease were documented and evaluated to determine their relationship with mortality.

Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics 22 (IBM Corp., Armonk, NY, USA). The Kolmogorov–Smirnov test indicated that the numerical variables were not normally distributed. Accordingly, comparisons of continuous variables between groups were performed with the Mann–Whitney U test. The Kruskal–Wallis test was used for comparisons of parameters among the chloride groups, with Dunn's test applied for post hoc analyses. Categorical variables were analysed using the Chi-square test, Fisher's exact test, the Fisher–Freeman–Halton exact test, or Yates' continuity correction, when appropriate. Variables that were significant in univariate analyses were further assessed using a multivariable logistic regression model. The multivariable logistic regression model was adjusted for hyperchloremia, albumin, calcium, creatinine, lactate, C-reactive protein, procalcitonin, and APACHE II score. A p-value below 0.05 was accepted as the threshold for statistical significance. Additionally, a post-hoc power analysis for the observed association between hyperchloremia and 28-day mortality ($\alpha = 0.05$, two-tailed, confidence level 0.95) demonstrated an achieved power ($1-\beta$) of 0.99.

RESULTS

A total of 221 elderly patients diagnosed with sepsis and admitted to the ICU were included in the study, with ages ranging from 66 to 97 years. Among them, 117 (52.9%) were women and 104 (47.1%) were men. The overall mean age

https://doi.org/10.5455/annalsmedres.2025.11.344

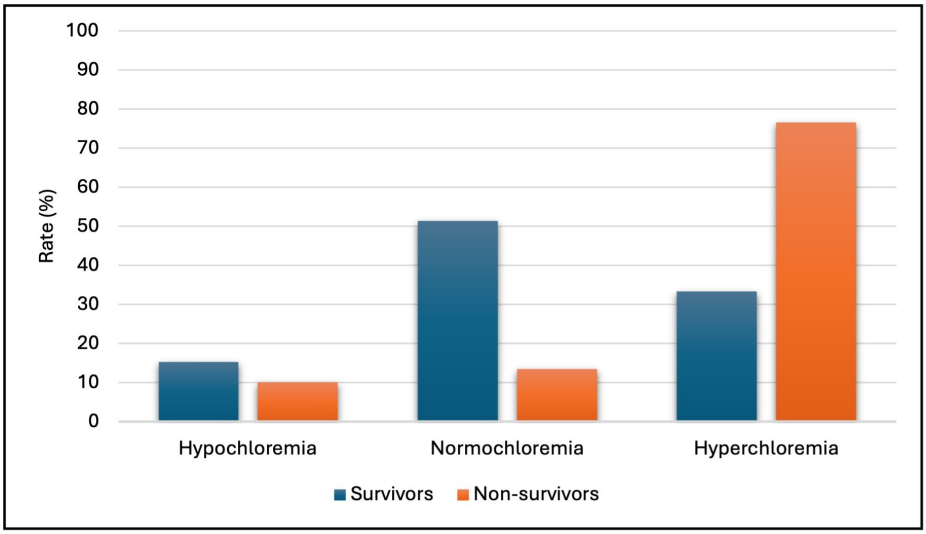

Figure 1. Distribution of survival and non-survival rates according to serum chloride levels.

Table 1. Distribution of serum chloride levels.

Chloride Level	n	%
Hypochloremia	26	11.8
Normochloremia	57	25.8
Hyperchloremia	138	62.4
Total	221	100

n = number.

of the study population was 80.87 ± 8.04 years. In total, 149 patients (67.4%) were non-survivors, while 72 (32.6%) were survivors. As shown in Table 1, hypochloremia, normochloremia, and hyperchloremia were identified in 11.8%, 25.8%, and 62.4% of patients, respectively.

No significant differences were detected between the surviving and non-surviving groups with respect to age or sex (p>0.05).

Non-survivors exhibited significantly lower albumin (p=0.001) and calcium levels (p=0.005), whereas creatinine (p=0.011), lactate (p=0.002), CRP (p=0.033), procalcitonin (p=0.009), and APACHE II scores (p=0.001) were significantly higher compared with survivors.

Serum chloride concentrations were strongly and significantly associated with mortality (p=0.001).

Chronic obstructive pulmonary disease was observed more frequently among survivors (35.4%) than non-survivors (18.4%), a difference that reached statistical significance (p=0.039).

Other comorbidities showed no significant relationship with mortality (p>0.05) (Table 2).

Hyperchloremia was more common in patients who died (76.5%) than in those who survived (33.3%) (Figure 1).

When patients were stratified according to serum chloride levels, no significant difference in age was observed among the chloride groups (p>0.05). Albumin (p=0.004) and calcium levels (p=0.006) differed significantly among the groups; both parameters were lower in the hyperchloremia group than in the hypochloremia and normochloremia groups (p<0.05). Sodium levels also varied significantly among the chloride groups (p=0.001) and were highest in patients with hyperchloremia. Similarly, APACHE II scores differed significantly between groups (p=0.001), with higher scores observed in the hyperchloremia group. No statistically significant differences were detected among the chloride groups in terms of WBC, hemoglobin, platelet count, RDW, potassium, creatinine, lactate, CRP, or procalcitonin levels (p>0.05) (Table 4).

The prevalence of comorbidities differed significantly between chloride groups (p=0.048), with a higher prevalence in the hypochloremia group than the hyperchloremia group. Chronic kidney disease was not observed in the normochloremia group, but it was significantly more frequent in patients with hypochloremia and hyperchloremia (p=0.042). Mortality rates differed significantly among the chloride groups (p=0.001), with the highest mortality observed in the hyperchloremia group; no significant difference in mortality was found between the hypochloremia and normochloremia groups (p>0.05).

In the backward stepwise logistic regression model constructed using variables that were significant in the univariate analyses (hyperchloremia, albumin, calcium, creatinine, lactate, CRP, procalcitonin, and APACHE II) the overall model reached statistical significance (p=0.001; p<0.05) and explained 83.2% of the variance (Nagelkerke R^2). The explanatory accuracy of the model was 92.8%. Hyperchloremia and APACHE II score remained independent predictors of mortality (p<0.05), with hyperchloremia increasing the odds of mortality by approximately 4.952-fold and each unit rise in APACHE II score being associated with a 1.335-fold higher mortality risk. Although lactate was not statistically significant (p>0.05), it remained in the model and was associated with a 1.329-fold increase in mortality risk (Table 3).

https://doi.org/10.5455/annalsmedres.2025.11.344

Table 2. Comparisons according to mortality.

	Survivors (n=72) Median (IQR)	Non-survivors(n=149) Median (IQR)	Total (n=221) Median (IQR)	p
Age (years)	81.5 (73-88)	81 (75-87)	81 (74-88)	0.712[1]
Sex n (%)				0.543[2]
Female	36 (50%)	81 (54.4%)	117 (52.9%)	
Male	36 (50%)	68 (45.6%)	104 (47.1%)	
WBC (×10^9/L)	11.87 (9.04-16.05)	13.2 (9.72-18.44)	12.9 (9.55-18.07)	0.109[1]
HGB (g/dL)	10.6 (8.53-11.98)	9.7 (8.3-11)	9.9 (8.4-11.55)	0.066[1]
PLT (×10^9/L)	240 (148.75-297)	211 (128.5-300)	217 (135-299)	0.232[1]
RDW (%)	51.15 (46.9-55.15)	52.3 (48.35-59.25)	52 (47.75-57.7)	0.058[1]
Albumin (g/L)	29 (24.25-31)	25 (21-29.5)	27 (22-30)	0.001*[1]
Calcium (mmol/L)	8.19 (7.81-8.51)	7.9 (7.3-8.4)	8 (7.48-8.42)	0.005*[1]
Sodium (mEq/L)	140.5 (137-145)	141 (136.5-147)	141 (137-146)	0.839[1]
Potassium (mEq/L)	3.88 (3.54-4.47)	3.96 (3.52-4.6)	3.92 (3.53-4.54)	0.476[1]
Creatinine (mg/dL)	1.01 (0.82-1.32)	1.3 (0.82-2.07)	1.19 (0.82-1.81)	0.011*[1]
Lactate (mmol/L)	1.95 (1.3-2.88)	2.5 (1.6-4.2)	2.3 (1.45-3.6)	0,002*[1]
CRP (mg/L)	131.25 (65.33-182)	152.8 (89.1-219.5)	140.3 (80.9-215.01)	0.033*[1]
Procalcitonin (ng/mL)	1.12 (0.28-4.76)	2.71 (0.71-10.15)	1.82 (0.57-9.2)	0.009*[1]
APACHE II	26 (23-32.75)	52 (44.5-56.5)	45 (32-55)	0.001*[1]
Chloride level n (%)				0.001*[2]
Hypochloremia	11 (15.3%)	15 (10.1%)	26 (11.8%)	
Normochloremia	37 (51.4%)	20 (13.4%)	57 (25.8%)	
Hyperchloremia	24 (33.3%)	114 (76.5%)	138 (62.4%)	
Comorbidities n (%)	48 (66.7%)	98 (65.8%)	146 (66.1%)	0.895[1]
Diabetes mellitus n (%)	13 (27.1%)	31 (31.6%)	44 (30.1%)	0.711[3]
Hypertension n (%)	25 (52.1%)	55 (56.1%)	80 (54.8%)	0.645[2]
Coronary artery disease n (%)	8 (16.7%)	10 (10.2%)	18 (12.3%)	0.397[3]
Conjestive heart failure n (%)	8 (16.7%)	27 (27.6%)	35 (24.0%)	0.215[3]
Chronic obstrructive pulmonary disease n (%)	17 (35.4%)	18 (18.4%)	35 (24%)	0.039*[3]
Cerebrovascular disease n (%)	9 (18.8%)	29 (29.6%)	38 (26.0%)	30.229[3]
Chronic kidney disease n (%)	3 (6.3%)	9 (9.2%)	12 (8.2%)	0.751[4]

[1]Mann Whitney U Test, [2]Chi-square test, [3]Continuity (yates) correction, [4]Fisher's Exact test *p<0.05. Continuous variables are expressed as median (IQR); categorical variables as n(%). WBC: white blood cells HGB: hemoglobin, PLT: platelet RDW:red cell distribution width CRP: C-reactive protein APACHE II: Acute Physiologic and Chronic Health Evaluation II.

Table 3. Logistic regression analysis of factors associated with mortality.

	Variables in the Equation	OR	95% CI		p
			Lower	Upper	
Step 6a	Hyperchloremia	4.952	1.512	16.219	0.008*
	Lactate	1.329	0.939	1.88	0.109
	APACHE II	1.335	1.224	1.456	0.001*

OR = Odds Ratio, CI = Confidence Interval *p<0.05. Variable(s)entered on step 1: Hyperchloramia, Albumin, Calcium, Creatinine, Lactate, CRP, Procalcitonin, APACHE II. APACHE II = Acute Physiology and Chronic Health Evaluation II CRP: C-reactive protein.

DISCUSSION

This study analyzed the association between serum chloride levels and 28-day mortality in elderly intensive care patients with sepsis. This study demonstrated that hyperchloremia was prevalent in the study population and significantly associated with mortality. Multivariate analysis demonstrated that hyperchloremia was an independent risk factor for 28-day mortality. This observation highlights the potential of hyperchloremia as an indicator of adverse outcomes in critically ill septic individuals. The notably high 28-day mortality rate observed in this study may be explained by the advanced age of the cohort and the markedly elevated APACHE II scores at ICU admission. Unlike previous studies that primarily included mixed adult ICU populations, our study focuses exclusively on geriatric (≥65 years) septic ICU patients and evaluates serum chloride concentrations obtained within the first 24 hours after ICU admission, thereby providing incremental evidence on the prognostic significance of chloride disturbances in this high-risk group. This higher mortality rate differs from that reported in several previous studies and may be partially explained by differences in patient characteristics. Our study population consisted exclusively of elderly patients admitted for medical conditions; postoperative and trauma-related cases were excluded and therefore were not represented in the cohort.

In 2011, a prospective cohort study was the first to demon-

https://doi.org/10.5455/annalsmedres.2025.11.344

Table 4. Comparison of study parameters by chloride levels.

	Hypochloremia (n=26) Median (IQR)	Normochloramia (n=57) Median (IQR)	Hyperchloremia (n=138) Median (IQR)	p
Age (years)	81 (70-86.5)	82 (73-88.5)	80.5 (75-88)	0.517[1]
Sex n (%)				0.042*[2]
Female	15 (57.7%)	22 (38.6%)	80 (58%)	
Male	11 (2.3%4)	35 (61.4%)	58 (42%)	
WBC(×10^9/L)	12,55 (9.4-15.8)	12.4 (8.8-15.4)	13.15 (9.8-18.8)	0.287[1]
HGB (g/dL)	9.15 (8.4-10.9)	10.4 (8.9-11.7)	9.85 (8.3-11.5)	0.292[1]
PLT(×10^9/L)	269 (155.5-325.8)	221 (136-298.5)	207.5 (133.8-299)	0.4021
RDW(%)	52.3 (48.4-58.9)	50.5 (46.3-54.6)	52.3 (48.4-59.2)	0.105[1]
Albumin (g/L)	29 (23.8-32.5)	28 (24-31)	25 (21-30)	0.004*[1]
Calcium (mmol/L)	8.225 (7.8-8.7)	8.12 (7.8-8.6)	7.9 (7.3-8.3)	0.006*[1]
Sodium (mEq/L)	136 (128.8-140.5)	139 (136-143)	143.5 (138-150)	0.001*[1]
Potassium (mEq/L)	4.015 (3.5-4.7)	3.91 (3.5-4.5)	3.905 (3.5-4.5)	0.782[1]
Creatinine (mg/dL)	1.235 (0.9-1.8)	1.11 (0.9-1.5)	1.25 (0.8-2.1)	0.473[1]
Lactate (mmol/L)	2.35 (1.3-3.7)	2 (1.3-3.3)	2.4 (1.6-3.8)	0.385[1]
CRP (mg/L)	149.5 (115.2-197)	119 (52.1-208.7)	146.35 (87.2-217)	0.189[1]
Procalcitonin (ng/mL)	2.355 (0.6-9.4)	1.09 (0.3-5)	2.3 (0.7-9.6)	0.108[1]
APACHE II	42 (25-52)	33 (24-44)	48 (39.8-56)	0.001*[1]
Comorbidities n (%)	22 (84.6%)	40 (70.2%)	84 (60.9%)	0.048*[2]
Diabetes mellitus n (%)	7 (31.8%)	11 (27.5%)	26 (31%)	0.910[2]
Hypertension n (%)	12 (54.5%)	24 (60%)	44 (52.4%)	0.728[2]
Coronary artery disease n (%)	0 (0%)	7 (17.5%)	11 (13.1%)	0.112[3]
Conjestive heart failure n (%)	8 (36.4%)	7 (17.5%)	20 (23.8%)	0.250[2]
Chronic obstrructive pulmonary disease n (%)	7 (31.8%)	16 (40%)	12 (14.3%)	0.005*[2]
Cerebrovascular disease n (%)	3 (13.6%)	8 (20%)	27 (32.1%)	0.126[2]
Chronic kidney disease n (%)	3 (13.6%)	0 (0%)	9 (10.7%)	0.042*[3]
Mortality n (%)	15 (57.7%)	20 (35.1%)	114 (82.6%)	0.001*[2]

[1]Kruskal Wallis Test, [2]Chi-square test, [3]Fisher Freeman Halton Exact test *p<0.05. Continuous variables are expressed as median (IQR); categorical variables as n(%).

strate an association between hyperchloremia and poorer outcomes in critically ill patients.The study reported a significant relationship between increased serum chloride levels and mortality among ICU patients [9]. Previous evidence also indicates that increased serum chloride levels predict mortality in both early and late phases. Supporting those observations, the present results show a clear relationship between high chloride levels and 28-day mortality among geriatric septic patients [10].

Aggressive fluid therapy in sepsis management frequently involves chloride-rich crystalloid solutions. However, beyond achieving hemodynamic stabilization, this approach may lead to metabolic disturbances associated with increased chloride load. Recent observational data suggest that differences in the content of resuscitation fluids may play a determining role in patient outcomes.The administration of resuscitation fluids containing low chloride or balanced crystalloids has been documented to result in a reduction in mortality [11,12]. Overall, the evidence indicates that fluid composition is an important factor shaping prognosis in critical illness, while limiting chloride excess may positively affect patient outcomes. Although we were unable to analyze specific fluid administration patterns, the observed association suggests that minimizing chloride load may benefit geriatric septic patients, aligning aligns with emerging evidence in fluid resuscitation research.

Multiple studies suggest that patients with low chloride levels experience higher mortality during sepsis. Evidence from one report further indicates that a moderate increase in chloride within 24 hours may be associated with better outcomes in those initially presenting with hypochloremia. In addition, other studies have described an L-shaped relationship between serum chloride levels and mortality, suggesting that particularly low chloride levels may be linked to worse outcomes [4]. A further study was conducted to compare the mortality rates of patients with septic shock according to their chloride status (hypochloremia, normochloremia, and hyperchloremia). Prior work has shown hypochloremia to be an independent determinant of elevated in hospital mortality [13]. In contrast, other investigators have linked early hyperchloremia with poorer outcomes, independent of metabolic or renal variables [14]. Our current findings similarly demonstrate a significant association between elevated chloride levels and mortality among older patients with sepsis.This finding suggests that both low and high chloride levels can disrupt physiological balance; however, in elderly septic patients, elevated chloride levels appear to represent a more prominent prognostic risk. The observed variations across studies may be attributed to heterogeneity in patient populations and variability in fluid therapy practices.

The relationship between hyperchloremia and mortality can be explained by several pathophysiological mechanisms. It has been demonstrated that an elevation in chloride concentra-

https://doi.org/10.5455/annalsmedres.2025.11.344

tion induces vasoconstriction at the macula densa, thereby reducing renal blood flow. Additionally, elevated chloride levels have been observed to induce vasoconstriction in renal arterioles, thereby exerting a deleterious effect on renal perfusion [15]. These mechanisms have been demonstrated to contribute to renal hypoperfusion and acidosis. Hyperchloremic metabolic acidosis resulting from chloride overload has been demonstrated to decrease myocardial contractility and impair vasopressor responsiveness [16]. As demonstrated by experimental sepsis models, elevated chloride levels have been shown to trigger proinflammatory responses [17]. Collectively, these mechanisms provide plausible physiological explanations for the negative impact of hyperchloremia on mortality.

Multivariate analysis in our study further suggested that the detrimental effect of hyperchloremia on mortality might be more pronounced in patients with higher APACHE II scores. Concomitant hyperchloremia and elevated APACHE II scores are hypothesised to result in deleterious clinical outcomes through mechanisms including an enhanced inflammatory response, renal hypoperfusion, and metabolic acidosis. This finding suggests that the prognostic impact of chloride imbalance may also be associated with disease severity.

Multiple studies involving heterogeneous ICU populations have reached similar conclusions. Hyperchloremia consistently emerged as a predictor of higher mortality in surgical and medical intensive care units and among patients with cerebral bleeding [18-20]. Our data echo these observations, adding further support to the growing recognition of serum chloride as a prognostic indicator in critical illness.

Recent research on artificial-intelligence-based clinical decision support systems has also highlighted the potential of serum chloride as a prognostic parameter. In one study, the incorporation of baseline serum chloride levels into predictive models resulted in enhanced prognostic accuracy [21]. This finding underscores that chloride is not merely a biochemical parameterbut also a clinically meaningful variable that should be considered in decision-making processes.

A notable strength of this study is its focus on geriatric patients with sepsis. Given that electrolyte imbalances may have distinct clinical consequences in elderly patients, this analysis constitutes an original contribution to the literature. Moreover, incorporating multivariate analysis and the APACHE II score into the model enhanced the reliability of evaluating independent effects on mortality.

Limitations

There are a few notable limitations to this research. Because the analysis was conducted retrospectively, some degree of observational bias or the effects of variables not captured in the dataset may persist. The absence of detailed data regarding the type, duration, and amount of fluid therapy makes it challenging to ascertain whether hyperchloremia was a cause or a consequence. This limitation impairs the ability to determine the extent to which hyperchloremia reflects underlying disease severity rather than iatrogenic factors. However, incorporation of clinical variables such as the APACHE II score, lactate, and creatinine into the model was intended to minimise this effect.

Secondly, because this investigation was conducted at a single institution, the external validity of the results may be limited. The limited sample size and relatively short data collection period may reduce the generalisability and statistical reliability of the findings; therefore, the results should be interpreted with appropriate caution. Moreover, disease severity was evaluated solely through the APACHE II scoring system; the omission of additional indices such as SOFA or SAPS II might have limited a more multidimensional assessment of disease severity.

Third, serum chloride concentrations were recorded only during the initial 24 hours after ICU admission, which prevented the analysis of temporal variations and their potential prognostic implications.

Additionally, the wide confidence interval around the adjusted odds ratio for hyperchloremia indicates imprecision in the effect estimate and therefore warrants cautious interpretation.

Lastly, because of the retrospective design, therapeutic decisions regarding fluid administration and electrolyte correction were not entirely uniform across patients.This may represent a significant confounding factor influencing chloride variability.

The absence of pathogen-specific data, detailed information regarding antibiotic modification, and subgroup comparisons across chloride categories constitutes a methodological limitation and should be considered when interpreting the study's findings.

CONCLUSION

Our findings indicate that elevated serum chloride levels correlate strongly with 28-day mortality among elderly septic patients in the ICU. Multivariate analysis revealed hyperchloremia as an independent predictor of death, with its adverse impact becoming more evident as disease severity increased. These findings suggest that routine monitoring of serum chloride levels and careful management of chloride load during fluid therapy may play an important role in improving prognosis among elderly patients with sepsis.

Ethics Committee Approval: Ethical approval for this study was obtained from the Kartal Koşuyolu High Specialization Training and Research Hospital Clinical Research Ethics Committee (Date: 20/05/2025, Approval No: 2025/08/1128).

Informed Consent: Not applicable due to the retrospective design.

https://doi.org/10.5455/annalsmedres.2025.11.344

Peer-review: Externally peer-reviewed.

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

Author Contributions: Conceptualization: İD; Data curation: İD, KT; Formal analysis: İD; Investigation: İD, MBY; Methodology: İD; Project administration: RF; Resources: İD, MBY; Software: KT; Supervision: MBY; Validation: İD, RF; Visualization: İD; Roles/Writing - original draft: İD, KT; and Writing - review & editing: İD, MBY.

Financial Disclosure: The authors received no financial support for the research, authorship, or publication of this article.

Artificial Intelligence Disclosure: The authors declare that an artificial intelligence–based tool (ChatGPT, OpenAI, GPT-4.0) 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 the development of scientific conclusions. All AI-assisted content was critically reviewed and approved by the authors to ensure its accuracy, originality, and scientific integrity. The authors take full responsibility for the content of the manuscript and confirm that its use complies with the journal's ethical guidelines.

REFERENCES

1. Fleischmann C, Scherag A, Adhikari NKJ, et al. Assessment of global incidence and mortality of hospital-treated sepsis—current estimates and limitations. *Am J Respir Crit Care Med.* 2016;193(3):259-272. doi: 10.1164/rccm.201504-0781OC.
2. Font MD, Thyagarajan B, Khanna AK. Sepsis and septic shock: basics of diagnosis, pathophysiology, and clinical decision making. *Med Clin North Am.* 2020;104(4):573-585. doi: 10.1016/j.mcna.2020.02.011.
3. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA.* 2016;315(8):801-810. doi: 10.1001/jama.2016.0287.
4. Zhao Z, Huang F, Shen J, Lv W, Huang Y, Pan J, Shen T. L-shape association between serum chloride and 365-day mortality in critically ill patients with sepsis based on the MIMIC-IV database. Sci Rep. 2025;15(1):20391. doi: 10.1038/s41598-025-08630-x.
5. Marunaka Y. Physiological roles of chloride ions in bodily and cellular functions. *J Physiol Sci.* 2023;73(1):31. doi: 10.1186/s12576-023-00889-x.
6. Lin Z, Deng Z, Liu J, et al. Chloride channel and inflammation-mediated pathogenesis of osteoarthritis. *J Inflamm Res.* 2022;15:953-964. doi: 10.2147/JIR.S350432.
7. Oh HJ, Kim SJ, Kim YC, et al. An increased chloride level in hypochloremia is associated with decreased mortality in patients with severe sepsis or septic shock. *Sci Rep.* 2017;7(1):15883. doi: 10.1038/s41598-017-16238-z.
8. Noritomi DT, Soriano FG, Kellum JA, et al. Metabolic acidosis in patients with severe sepsis and septic shock: a longitudinal quantitative study. *Crit Care Med.* 2009;37(10):2733-2739. doi: 10.1097/CCM.0b013e3181a59165.
9. Boniatti MM, Cardoso PR, Castilho RK, Vieira SR. Is hyperchloremia associated with mortality in critically ill patients? A prospective cohort study. *J Crit Care.* 2011;26(2):175-179. doi: 10.1016/j.jcrc.2010.04.013.
10. Huang K, Hu Y, Wu Y, Ji Z, Wang S, Lin Z, et al. Hyperchloremia is associated with worse outcomes in critically ill stroke patients. *Front Neurol.* 2018;9:485. doi: 10.3389/fneur.2018.00485.
11. Shaw AD, Raghunathan K, Peyerl FW, et al. Association between intravenous chloride load during resuscitation and in-hospital mortality among patients with SIRS. *Intensive Care Med.* 2014;40(12):1897-1905. doi: 10.1007/s00134-014-3505-3.
12. Raghunathan K, Shaw A, Nathanson B, et al. Association between choice of IV crystalloid and in-hospital mortality in adults with sepsis. *Crit Care Med.* 2014;42(7):1585-1591. doi: 10.1097/CCM.0000000000000305.
13. Ruan X, Gao Y, Lai X, Wang B, Wu J, Yu X. Trimatch comparison of the prognosis of hypochloremia, normochloremia, and hyperchloremia in patients with septic shock. *J Formos Med Assoc.* 2025;124(5):426-431. doi: 10.1016/j.jfma.2024.05.012.
14. Neyra JA, Canepa-Escaro F, Li X, et al; Acute Kidney Injury in Critical Illness Study Group. Association of hyperchloremia with hospital mortality in critically ill septic patients. *Crit Care Med.* 2015;43(9):1938-1944. doi: 10.1097/CCM.0000000000001161.
15. Chowdhury AH, Cox EF, Francis ST, Lobo DN. A randomized, controlled, double-blind crossover study on the effects of 2-L infusions of 0.9% saline and plasma-lyte® 148 on renal blood flow velocity and renal cortical tissue perfusion in healthy volunteers. *Ann Surg.* 2012;256(1):18-24. doi: 10.1097/sla.0b013e318256be72.
16. Levy B, Collin S, Sennoun N, et al. Vascular hyporesponsiveness to vasopressors in septic shock: from bench to bedside. *Intensive Care Med.* 2010;36(12):2019-2029. doi: 10.1007/s00134-010-2045-8.
17. Kellum JA, Song M, Almasri E. Hyperchloremic acidosis increases circulating inflammatory molecules in experimental sepsis. *Chest.* 2006;130(4):962-967. doi: 10.1378/chest.130.4.962.
18. Song K, Yang T, Gao W. Association of hyperchloremia with all-cause mortality in surgical intensive care unit patients: a retrospective cohort study. *BMC Anesthesiol.* 2022;22(1):14. doi: 10.1186/s12871-021-01558-5.
19. Zhou D, Li T, Zhao D, et al. Increase in chloride from baseline is independently associated with mortality in intracerebral hemorrhage patients admitted to intensive care unit: a retrospective study. *J Intensive Med.* 2022;2(4):274-281. doi: 10.1016/j.jointm.2022.04.002.
20. Yılmaz H, Güney BÇ. Evaluation of the effect of hyperchloremia on the prognosis and mortality of medical intensive care patients: a single-center study. *Turk J Med Sci.* 2022;52(5):1689-1696. doi: 10.55730/1300-0144.5510.
21. Wang J, Li Q, Xie C, et al. Predicting in-hospital mortality in intensive care unit patients using Causal SurvivalNet with serum chloride and other causal factors: a cross-country study. *J Med Internet Res.* 2025;27:e70118. doi: 10.2196/70118.

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

Evaluation of swallowing and voice functions in children with different degrees of adenoid hypertrophy and after adenoidectomy

Suheyla Kandemir[a,*], Tugay Basar[a], Nuray Bayar Muluk[a], Ela Comert[a], Burak Mustafa Tas[a], Ziya Sencan[a]

[a]Kırıkkale University, Faculty of Medicine, Department of Otorhinolaryngology, Kırıkkale, Türkiye
*Corresponding author: dr.suheyla_bostan@hotmail.com (Suheyla Kandemir)

MAIN POINTS

- Children with severe adenoid hypertrophy (>50% nasopharyngeal obstruction) exhibited significantly higher PEDI-EAT-10 and CVHI-10 scores, indicating greater self-reported severity of swallowing and voice-related symptoms.
- Self-reported dysphonia (CVHI-10 ≥9) and swallowing difficulties (PEDI-EAT-10 >3) were most prevalent in patients with severe adenoid hypertrophy, whereas such symptoms were rare or absent in controls.
- Adenoidectomy led to significant postoperative improvements in self-reported swallowing and voice symptoms, as assessed by questionnaire scores.

Cite this article as: Kandemir S, Basar T, Bayar Muluk N, Comert E, Tas BM, Sencan Z. Evaluation of swallowing and voice functions in children with different degrees of adenoid hypertrophy and after adenoidectomy. *Ann Med Res*. 2026;33(6):284–291. doi: 10.5455/annalsmedres.2025.11.356.

ABSTRACT

Aim: To evaluate self-reported swallowing difficulties and voice-related symptoms in children with varying degrees of adenoid hypertrophy and in those who had undergone adenoidectomy.

Materials and Methods: Children aged 5–12 years diagnosed with adenoid hypertrophy were included in this prospective cohort study. Patients were categorized into three groups based on the degree of nasopharyngeal obstruction: no detectable adenoid hypertrophy (control), adenoid hypertrophy occupying less than ≤50% of the nasopharyngeal space, and adenoid hypertrophy occupying more than >50% of the nasopharyngeal space. Swallowing and voice functions were assessed using validated self-report screening questionnaires: the Children's Voice Handicap Index-10 (CVHI-10) and the Pediatric Eating Assessment Tool-10 (PEDI-EAT-10). For patients who underwent adenoidectomy, both questionnaires were re-administered one month postoperatively and compared with preoperative results.

Results: A total of 113 children (mean age 7.41 ± 2.00 years) were included. Group comparisons showed that PEDI-EAT-10 and CVHI-10 scores were significantly higher in Group 3 than in Group 1 ($p<0.001$). Adenoidectomy significantly reduced both scores postoperatively ($p<0.001$). Self-reported dysphonia was observed in 15% of patients, and self-reported swallowing difficulties in 25.7% of patients, with the highest prevalence in Group 3.

Conclusion: Severe adenoid hypertrophy was associated with greater patient-reported impairments in swallowing and voice function compared with controls. Adenoidectomy led to significant postoperative improvements in swallowing and voice symptoms, as assessed by screening tools.

Keywords: Adenoid hypertrophy, Swallowing, Voice, Self-reported symptoms, Pediatric otolaryngology

Received: Nov 17, 2025 **Accepted:** Jan 12, 2026 **Available Online:** Jun 25, 2026

INTRODUCTION

The adenoid is lymphoid tissue located on the roof and posterior wall of the nasopharynx [1]. Adenoid hypertrophy leads to prominent symptoms such as nasal obstruction, snoring, chronic oral breathing, and sleep-disordered breathing in children [2]. The tonsils and adenoids, situated within the vocal tract, can narrow the airway, hinder sound production, and alter voice quality [3]. Enlargement of these structures may result in nasal obstruction, subsequent oral breathing, and dehydration of the laryngeal mucosa. Indeed, it has been demonstrated that even 15 minutes of oral breathing in healthy adults increases phonation threshold and vocal effort [4]. Mucosal dryness elevates vocal fold friction, leading to vocal edema and fatigue. Furthermore, children with adenoid hypertrophy frequently present with speech disorders; hyponasality commonly occurs because nasal resonance is reduced by the adenoids [5].

Adenoidectomy, performed alone or with tonsillectomy, is a common procedure in pediatric otorhinolaryngology [6]. However, postoperative complications such as pain, bleeding,

edema, nasopharyngeal stenosis, hypernasality, and velopharyngeal insufficiency may occur [7]. Velopharyngeal dysfunction arises from inadequate closure during speech and swallowing, resulting in hypernasality and nasal air escape or nasal regurgitation of food [8]. Adenoid hypertrophy may play a modulatory role in velopharyngeal closure, acting as a passive cushion that facilitates velar contact in some children; therefore, adenoid hypertrophy alone is not considered a direct cause of velopharyngeal insufficiency. Instead, velopharyngeal function appears to depend on the balance between upper airway anatomy and neuromuscular coordination, and latent velopharyngeal insufficiency may become evident following adenoidectomy. Moreover, effective velopharyngeal closure is relevant not only for speech but also for swallowing.

Incomplete nasopharyngeal closure may result in nasal air escape or, less commonly, nasal regurgitation of food [8]. Adequate velopharyngeal function is essential for normal speech development, and surgical interventions involving the oropharynx and nasopharynx, such as adenoidectomy, may alter the pharyngeal musculature and contribute to velopharyngeal insufficiency [9].

There is a close relationship among airway muscles, airway anatomy, and the coordination of respiration and swallowing. Chronic oral breathing may lead to muscle prolapse and positional changes in orofacial structures [10]. In mouth-breathing children, swallowing has been shown to involve altered tongue activity patterns, with delayed and less coordinated tongue–palate contact, reflecting impaired neuromuscular control during the oral phase of swallowing [10]. The primary function of the upper airway is to facilitate the passage of air and food from the oral cavity to the oropharynx; however, obstructive factors may disrupt this process, resulting in dysphagia during the pharyngeal phase [11]. Swallowing consists of the oral preparatory, oral transit, pharyngeal, and esophageal phases. During this process, food is chewed, mixed with saliva, and transported to the pharynx by the tongue, while the soft palate closes the choanae to prevent nasal regurgitation and the pharyngeal muscles transfer the bolus to the lower pharynx [12,13]. Effective swallowing requires precise coordination between breathing and deglutition, and studies have demonstrated that alterations in upper airway patency and nasal airflow can disrupt breathing–swallowing coordination, leading to reduced swallowing efficiency [14]. Dysphagia may arise from various factors that interfere with bolus transit or sensitivity at any stage of this sequence [15]. Radiological, endoscopic, and manometric methods used for definitive diagnosis are often invasive and costly.

The aim of this study is to identify pediatric patients with self-reported swallowing difficulties and voice-related symptoms among those with varying degrees of adenoid hypertrophy and after adenoidectomy, using two non-invasive screening questionnaires: the Children's Voice Handicap Index-10 (CVHI-10) and the Pediatric Eating Assessment Tool-10 (PEDI-EAT-10). Although swallowing and voice functions have been investigated in patients undergoing tonsillectomy, studies focusing exclusively on adenoid hypertrophy remain limited. Moreover, the evaluation of patient-reported swallowing and voice symptoms following adenoidectomy has received relatively little attention in the literature. Therefore, this study aims to provide additional data on the association between severity of adenoid hypertrophy and self-reported swallowing and voice symptoms in the pediatric population.

MATERIALS AND METHODS

This study was conducted on pediatric patients aged 5–12 years who presented to the Otorhinolaryngology (ENT) Outpatient Clinic of the University Faculty of Medicine Hospital and were diagnosed with adenoid hypertrophy. The age range of 5–12 years was selected because children under 5 might have difficulty understanding and completing the questionnaires, whereas children over 12 might exhibit voice changes related to developmental factors that could affect the results. Adenoidectomy was indicated for sleep-disordered breathing.

Exclusion criteria included craniofacial malformations (including cleft palate), prior adenotonsillar surgery, neuromuscular disorders, systemic diseases, candidacy for tonsillectomy, nasal septal deviation or hypertrophic obstructive turbinates, tonsillar hypertrophy (grades 3 and 4), or palatal anomalies.

The study protocol was approved by the Kırıkkale University Non-Interventional Research Ethics Committee (Date: 14.11.2024, Decision No: 2024.11.14). Written informed consent was obtained from the parents of all participants, and the principles of the Declaration of Helsinki were followed.

Assessment of adenoid size

All patients underwent flexible nasopharyngoscopy (Karl Storz, Tuttlingen, Germany), and the degree of adenoid obstruction was assessed based on the proportion of the nasopharyngeal airway occupied by adenoid hypertrophy. The "endoscopic adenoid percentage" was defined as the estimated percentage of choanal obstruction caused by the adenoid hypertrophy, visually determined during endoscopic examination. Each examination was independently evaluated by two experienced otorhinolaryngologists who were blinded to the questionnaire results. The adenoid size was categorized as causing $\leq$50% or >50% nasopharyngeal obstruction by consensus of the two observers. In cases of initial disagreement, the recorded endoscopic images were reviewed jointly, and a final decision was reached by consensus. This approach was used to minimize inter-observer variability and improve measurement reliability.

Study groups

All patients underwent a general ENT examination and flexible nasopharyngoscopy (Karl Storz, Tuttlingen, Germany). Based on the size of the adenoid hypertrophy in the nasopharynx, patients were categorized into three groups:

https://doi.org/10.5455/annalsmedres.2025.11.356

- Group 1 (Control group): The control group consisted of patients who were referred for evaluation of nasal obstruction or snoring but were found to have no detectable adenoid hypertrophy on nasopharyngoscopic examination (Figure 1a).
- Group 2: Children with adenoid hypertrophy occupying ≤50% of the nasopharynx (Figure 1b).
- Group 3: Children with adenoid hypertrophy occupying >50% of the nasopharynx (Figure 1c) Most patients in this group were candidates for surgical treatment. Those who underwent adenoidectomy were re-evaluated in the first postoperative month.

Demographic data (age and sex) for all patients were recorded. The participants were administered the Children's Voice Handicap Index-10 (CVHI-10) and the Pediatric Eating Assessment Tool-10 (PEDI-EAT-10) questionnaires. For Group 3 patients undergoing adenoidectomy, the same Questionnaires were re-administered at one month postoperatively, and the preoperative and postoperative results were compared. Validated parent- or self-reported questionnaires were used as non-invasive tools to assess symptom severity and functional impact related to voice and swallowing in the pediatric population. Both CVHI-10 and PEDI-EAT-10 are screening instruments and do not provide a definitive diagnosis of dysphonia or dysphagia.

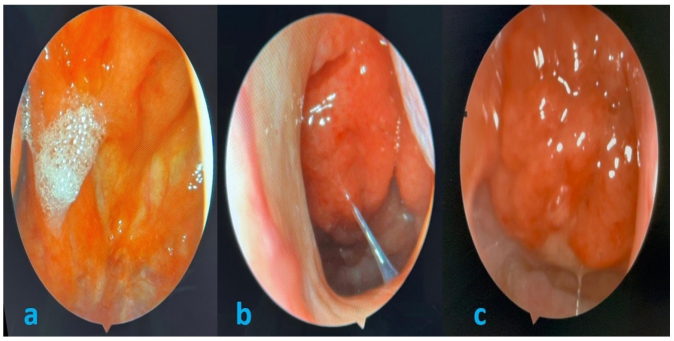

Figure 1. Nasopharyngoscopic images: (a) absence of adenoid hypertrophy; (b) adenoidhypertrophy occupying ≤50% of the nasopharyngeal space; (c) adenoid hypertrophy occupying >50% of the nasopharyngeal space.

Children's Voice Handicap Index-10 (CVHI-10)

The CVHI-10 is a self-assessment tool for pediatric dysphonia, completed directly by children. It was developed by Ricci-Maccarini et al. [16]. The scale consists of 10 items and assesses how dysphonia affects the child's daily life and their perception of their voice disorder. In younger children, particularly those aged 5–7 years, the questionnaire was completed with parental assistance to ensure comprehension and response accuracy. Responses are rated using a Likert scale ranging from 0 to 3. The Turkish version of the CVHI-10 has been validated as a reliable and sensitive instrument for evaluating children's perceptions of voice disorders [17]. The internal consistency of the Turkish-CVHI-10 was 0.87. The Cronbach's alpha coefficient is considered good. The cut-off point of the Turkish version of the CVHI-10 is 9, whereas in the original CVHI-10, it was 4. A CVHI-10 score of 9 or higher is considered indicative of dysphonia [17].

Pediatric Eating Assessment Tool-10 (PEDI-EAT-10)

The Eating Assessment Tool-10 (EAT-10) is a widely used, self-administered, symptom-specific outcome measure that has been validated for clinical use [18]. Beyond assessing the risk of dysphagia from the patients' perspective, it also serves as a valuable tool for identifying dysphagia symptoms and evaluating their prevalence in associated clinical conditions [19]. In children aged 5–7 years, the PEDI-EAT-10 was completed with parental assistance, whereas older children completed the questionnaire independently under supervision.

The PEDI-EAT-10 version consists of ten questions and is a valid and reliable symptom-specific instrument for children, with its validity demonstrated in Turkish pediatric populations [20,21]. An EAT-10 score greater than 3 was considered indicative of aspiration risk [18].

Surgical technique (Adenoidectomy)

All surgical procedures were performed under general anesthesia with oral endotracheal intubation. The child was positioned supine in the Rose position, and adenoidectomy was performed using the conventional curettage technique with

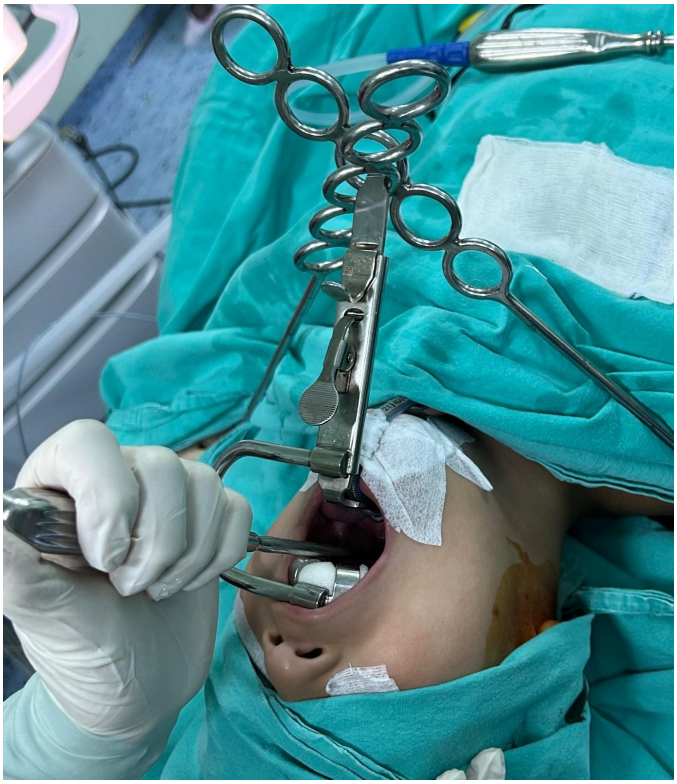

Figure 2. Conventional curettage technique performed using a Boyle--Davis mouth gag.

https://doi.org/10.5455/annalsmedres.2025.11.356

Table 1. Patient demographics and clinical characteristics.

Parameter	Group 1 (No Adenoid) (n=36)	Group 2 (≤50% Adenoid) (n=36)	Group 3 (>50% Adenoid) (n=41)	p
Age	7.67±1.82	7.39±2.01	7.20±2.15	0.591
Gender (Female/Male)	14/22	22/14	8/33	**0.001**
Endoscopic Adenoid Size, %	0	42.2±9.29	73.54±10.62	**<0.001**

Bold values indicate statistical significance.

Table 2. PEDI-EAT-10 and CVHI-10 scores.

	Group 1 (No Adenoid) (n=36)	Group 2 (≤50% Adenoid) (n=36)	Group 3 (>50% Adenoid) (n=41)	p
PEDI-EAT-10	0.64±1.07	3.22±4.24	6.20±8.61	**<0.001**
CVHI-10	0.94±1.95	3.81±4.45	6.00±7.66	**<0.001**

All scores are given as mean±SD. Bold values indicate statistical significance.

Table 3. Pre- and Postoperative PEDI-EAT-10 and CVHI-10 scores in adenoidectomy patients.

	Group 3 (>50% Adenoid) (n=30 Undergoing adenoidectomy)
PEDI-EAT-10	
Preoperative	2.00 (0.00-10.25)
Postoperative	0.00 (0.00-1.25)
p	**<0.001**
CVHI-10	
Preoperative	2.50 (0.00-8.25)
Postoperative	0.00 (0.00-2.00)
p	**<0.001**

All scores are given as median (25^{th}-75^{th} percentile). P-values were obtained using the Wilcoxon signed-rank test. Bold values indicate statistical significance.

the aid of a Boyle-Davis mouth gag (Figure 2). A nasopharyngeal pack was placed for five minutes. After removal of the pack, hemostasis was achieved, and the procedure was completed. No serious intraoperative or postoperative complications occurred.

Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics, version 21.0 (Armonk, NY: IBM Corp.). The distributions of the variables were assessed using both analytical (Kolmogorov–Smirnov test) and visual methods (histograms and probability plots). Categorical variables were compared using the chi-square test or Fisher's exact test, as appropriate. Continuous variables were analyzed by one-way analysis of variance (ANOVA) or the Kruskal–Wallis test, depending on their distributional characteristics. For within-group comparisons of preoperative and postoperative measurements in patients undergoing adenoidectomy, the Wilcoxon signed-rank test was used because the paired data were not normally distributed. Post-hoc pairwise comparisons were adjusted using the Bonferroni correction. Results are presented as counts and percentages for categorical variables, mean ± standard deviation for normally distributed continuous variables, and median with interquartile range (25^{th}–75^{th} percentiles) for non-normally distributed continuous variables. A post hoc power analysis was performed using the observed effect size (Cohen's d = 0.88) obtained from the comparison between Group 1 and Group 3 on both the PEDI-EAT-10 and CVHI-10 scores. At a two-tailed significance level of 0.05 (95% confidence), the achieved statistical power exceeded 0.95, indicating that the sample size was sufficient to detect clinically meaningful differences between groups.

RESULTS

A total of 113 pediatric patients with a mean age of 7.41 ± 2.00 years were included in the study. Of these, 44 were female (38.9%) and 69 were male (61.1%). The overall mean Children's Voice Handicap Index-10 (CVHI-10) score was 3.69 ± 5.72, and the overall mean Pediatric Eating Assessment Tool-10 (PEDI-EAT-10) score was 3.48 ± 6.15. Demographic and clinical characteristics of the patients are presented in Table 1. There were no significant differences in age among the groups (p>0.05). However, sex distribution differed significantly among groups (p = 0.001), with a higher proportion of females in Group 2 and of males in Groups 1 and 3. Endoscopic adenoid percentage also differed significantly among groups (Table 1). Group 2 had a mean adenoid percentage of 42.2 ± 9.29%, whereas Group 3 had a mean adenoid percentage of 73.54 ± 10.62%.

Group PEDI-EAT-10 and CVHI-10 scores are shown in Table 2. A significant difference was observed in PEDI-EAT-10 scores among the groups (p<0.001). Post hoc analysis revealed that this difference was attributable to higher scores in Group 3 than in Group 1 (p<0.001). The magnitude of this difference was large (Cohen's d = 0.88, 95% CI: 0.41–1.35). No significant differences were found between Groups 1 and 2 (p = 0.178) or between Groups 2 and 3 (p = 0.077). CVHI-10

https://doi.org/10.5455/annalsmedres.2025.11.356

Table 4. Distribution of self-reported dysphonia and swallowing difficulties across the study groups.

Parameter	Group 1 (No Adenoid) (n=36)	Group 2 (≤50% Adenoid) (n=36)	Group 3 (>50% Adenoid) (n=41)	Total
Normal Voice/ Self-reported dysphonia symptoms (CVHI-10 ≥9)	36/0	30/6	30/11	96/17
Percentage (%)	100%/0%	83.3%/16.7%	73.2%-26.8%	85.0%-15.0%
Normal Swallowing/ Self-reported swallowing difficulties (PEDI-EAT-10 >3)	35/1	26/10	23/18	84/29
Percentage (%)	97.2%-2.8%	72.2%-27.8%	56.1%-43.9%	74.3%-25.7%

All values are presented as number (n) and percentage (%).

scores also differed significantly among the three groups. Post hoc analysis showed that the difference was driven by higher scores in Group 3 compared with Group 1 ($p<0.001$), with a large effect size (Cohen's d = 0.88, 95% CI: 0.41–1.35). No significant differences were observed between Groups 1 and 2 ($p = 0.078$) or between Groups 2 and 3 ($p = 0.230$).

In Group 3, 30 patients underwent adenoidectomy and 11 did not. Pre- and postoperative PEDI-EAT-10 and CVHI-10 scores for the surgical subgroup are presented in Table 3. Both PEDI-EAT-10 and CVHI-10 scores improved significantly after adenoidectomy ($p<0.001$), indicating that the procedure reduced the symptom burden.

The risk of dysphonia (CVHI-10 ≥9) was observed in 15% of patients (Table 4). The highest prevalence was noted in Group 3 (26.8%), whereas no patients in Group 1 exhibited this risk. In Group 3, four of the 11 nonsurgical patients and seven of the 30 surgical patients exhibited risk of dysphonia preoperatively. Among the surgical subgroup, 23 patients had normal voice, while 7 exhibited risk of dysphonia before surgery. At the first postoperative month, only one patient continued to exhibit risk of dysphonia.

A Risk of dysphagia (PEDI-EAT-10 >3) was identified in 25.7% of patients (Table 4). The highest prevalence was noted in Group 3 (43.9%), while Group 1 had the lowest prevalence (2.8%). In Group 3, five of the 11 nonsurgical patients exhibited a risk of dysphagia. Among the 30 patients who underwent adenoidectomy, 13 exhibited risk of dysphagia preoperatively, but only two continued to exhibit risk of dysphagia postoperatively.

DISCUSSION

Hypertrophic adenoids may facilitate velopharyngeal closure by narrowing the space between the soft palate and the posterior pharyngeal wall. Their removal during adenoidectomy suddenly enlarges this space, which can cause transient hypernasality, whereas hypertrophy itself may lead to hyponasality [5, 22-24]. Because the vocal tract acts as an acoustic resonator, any structural alterations of the oral or pharyngeal cavities can modify resonance characteristics and affect perceived voice quality [3,25].

In this study, larger adenoid size was associated with higher Children's Voice Handicap Index-10 (CVHI-10) scores and increased self-reported dysphonia symptoms. After adenoidectomy, CVHI-10 scores improved significantly, indicating a positive effect of surgery on voice-related quality of life. These findings suggest that mechanical obstruction and altered resonance due to adenoid hypertrophy predominantly influence patients subjective perception of voice, as reflected by patient-reported outcome measures. Although CVHI-10 scores were higher in Group 3 than in Group 2, the lack of a statistically significant difference between these two groups may suggest that once adenoid hypertrophy exceeds a certain threshold of nasopharyngeal obstruction, further increases in size do not proportionally worsen perceived voice-related symptoms. Previous research has yielded comparable findings. Salami et al. reported a significant improvement in Voice Handicap Index (VHI) scores after tonsillectomy and/or adenoidectomy [26]. Naraghi et al. observed postoperative improvements in Pediatric Voice Handicap Index (PVHI) and Pediatric Voice-Related Quality of Life (PVRQOL) scores among 86 children who underwent adenotonsillectomy [27]. Similarly, Karakurt et al. found improved PVHI scores among children with tonsillar hypertrophy after adenotonsillectomy, although objective acoustic parameters remained unchanged; this parallels our observation that symptom-based voice outcomes improve following relief of upper airway obstruction [28].

Abdel-Aziz et al. showed that children with severe adenoid hypertrophy frequently exhibited preoperative hyponasality and transient postoperative hypernasality, which typically resolved within three months, reflecting temporary resonance adaptation after surgery [29]. Pinto et al. also reported improved nasal resonance following adenoidectomy or adenotonsillectomy, without significant changes in acoustic parameters [3]. Together with the present results, these findings suggest that adenoidectomy may improve perceived voice quality, even in the absence of objective acoustic assessment in the present study.

Liu et al. found no significant differences in acoustic parameters after tonsillectomy with or without adenoidectomy. However, aerodynamic analyses demonstrated increased maximum phonation time and reduced subglottic pressure, suggesting improved phonatory efficiency postoperatively [30]. These physiological improvements may explain the postoperative reduction in CVHI-10 scores observed in our cohort.

https://doi.org/10.5455/annalsmedres.2025.11.356

Brkic et al. evaluated 1,258 pediatric patients undergoing adenotonsillectomy, tonsillectomy, or adenoidectomy and observed transient postoperative increases in fundamental frequency and decreases in jitter and shimmer values, indicating voice-parameter fluctuations [31], which is consistent with the early postoperative improvement seen in our surgically treated patients.

Lundeborg et al. compared tonsillectomy plus adenoidectomy with tonsillotomy plus adenoidectomy in 67 children and reported preoperative hyponasality and altered acoustic parameters. Postoperatively, perceptual voice quality normalized, although minor acoustic differences persisted; these did not differ significantly between surgical methods [5]. This aligns with our findings that subjective voice outcomes improved following adenoidectomy. Inja et al. observed improvements only in subjective measures [32]. Furthermore, a meta-analysis by Wang et al., including 23 studies and 2,154 children, reported mild improvements in shimmer and the soft phonation index at one month postoperatively, but no significant changes at three months, suggesting that postoperative voice alterations are usually transient [33], indicating that the primary benefit of adenoidectomy may be symptom relief rather than sustained acoustic modification. In contrast, Lee et al. demonstrated significant postoperative reductions in both perceptual and acoustic voice parameters, including PVHI, jitter, shimmer, and noise-to-harmonic ratio, suggesting more sustained voice improvement following surgery [25]. These discrepancies may be attributable to heterogeneity among studies, including differences in patient populations, surgical procedures, follow-up duration, and outcome measures.

In the present study, self-reported swallowing difficulties (PEDI-EAT-10>3) were detected in 25.7% of all participants, with the highest prevalence observed in the group with severe adenoid hypertrophy (43.9%). This finding suggests that enlarged adenoids may contribute to swallowing difficulties in children by mechanically interfering with normal bolus passage and oropharyngeal coordination. Group 3, with the highest mean adenoid percentage (73.54 ± 10.62), also exhibited significantly higher PEDI-EAT-10 scores than Group 1 ($p<0.001$), indicating a dose–response relationship between adenoid size and symptom severity. The absence of a significant difference between Groups 2 and 3 may reflect compensatory swallowing mechanisms or symptom adaptation once moderate obstruction is present, limiting further symptom escalation despite increasing adenoid size. Furthermore, in Group 3 patients in patients who underwent surgery, postoperative evaluations revealed significant decreases in PEDI-EAT-10 scores ($p<0.001$), while the prevalence of self-reported swallowing difficulties decreased from 43.3% to 6.6%, indicating that adenoidectomy effectively alleviated swallowing-related symptoms. These results align with previous evidence demonstrating that adenotonsillar hypertrophy can impair oral motor and swallowing function.

Lundeborg et al., evaluated children using the Nordic Orofacial Test–Screening (NOT-S) and reported marked improvement in preoperative oral motor difficulties, such as chewing and swallowing, six months after adenotonsillar surgery [34]. This is consistent with our findings that relief of upper airway obstruction is associated with improvement in swallowing-related symptoms. Similarly, de Oliveira Branco et al., described persistent alterations in swallowing dynamics, including reduced labial seal and vallecular residue, even years after adenotonsillectomy, suggesting that subtle functional adaptations may persist [35], suggesting that although symptom-based improvement is common, subtle functional adaptations may persist beyond clinical recovery. Walsh et al. also demonstrated significant postoperative improvement in feeding behavior, assessed by the Pediatric Eating Assessment Tool, particularly in meal fatigue and avoidance of chewable foods, thereby confirming the beneficial impact of adenotonsillectomy on feeding efficiency [36]. This finding parallels the significant reduction in PEDI-EAT-10 scores observed in our surgical cohort.

Collectively, these findings indicate that adenoid hypertrophy is associated with measurable impairment in swallowing, which can be substantially reduced following adenoidectomy. However, the persistence of subtle or long-term orofacial motor adaptations, as reported by de Oliveira Branco et al., underscores the importance of postoperative follow-up, particularly in children with severe preoperative obstruction [35].

A major strength of this study lies in its prospective design and the inclusion of a relatively large pediatric cohort with clearly defined diagnostic criteria. The use of validated questionnaires (PEDI-EAT-10 and CVHI-10) allowed for the standardized assessment of swallowing and voice functions. Moreover, preoperative and postoperative comparisons provided valuable insights into functional recovery following adenoidectomy.

However, the absence of a phoniatric unit and a speech therapist limited the ability to perform objective acoustic and swallowing analyses. Although instrumental diagnostic methods such as videofluoroscopic swallowing study (VFSS) and fiberoptic endoscopic evaluation of swallowing (FEES) are considered the gold standard for detecting swallowing disorders, they were not included because they are invasive, costly, and often impractical in pediatric populations. In addition, the unequal gender distribution among the study groups and the relatively small number of patients undergoing adenoidectomy may limit the generalizability of the findings.

CONCLUSION

This study demonstrates that adenoid hypertrophy negatively affects self-reported swallowing difficulties and voice-related symptoms in children, particularly when nasopharyngeal obstruction exceeds 50%. Functional impairment was significant only in those with severe obstruction, compared with

https://doi.org/10.5455/annalsmedres.2025.11.356

controls, as assessed using validated self-report screening questionnaires rather than by direct instrumental evaluation. Adenoidectomy led to marked improvements in patient-reported outcomes for swallowing and voice, emphasizing its role in alleviating symptom burden. Because these findings are based on screening tools (PEDI-EAT-10 and CVHI-10), they reflect perceived functional impact rather than objectively measured swallowing or voice function. Incorporating subjective tools such as the PEDI-EAT-10 and CVHI-10 as initial screening instruments may aid clinical evaluation; however, objective instrumental assessments (e.g., VFSS or FEES) should be considered in future studies to confirm and further characterize functional outcomes.

Ethics Committee Approval: The study protocol was approved by the Kırıkkale University Non-Interventional Research Ethics Committee (Date: 14.11.2024, Decision No: 2024.11.14).

Informed Consent: Written informed consent was obtained from the parents of all participants.

Peer-review: Externally peer-reviewed.

Availability of Data and Materials: The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest: No potential conflict of interest was reported by the authors.

Author Contributions: Conception: SK, TB, NBM, EC, BMT, ZS; Design: SK, TB, NBM, EC, BMT, ZS; Supervision: SK, TB, NBM, EC, BMT, ZS; Materials: SK, TB, EC, BMT, ZS; Data Collection and/or Processing: SK, TB, EC, BMT, ZS; Analysis and/or Interpretation: SK, NBM, EC; Literature Review: SK, NBM; Writing: SK, TB, NBM; Critical Review: SK, TB, NBM, EC, BMT, ZS.

Financial Disclosure: The authors received no financial support for the research, authorship, and/or publication of this article.

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 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. Abo Elmagd EA, Khalifa MS, Abeskharoon BK, El Tahan AA. Comparative study between conventional adenoidectomy and adenoidectomy using micro-debrider. *Egypt J Otolaryngol.* 2021;37(56). doi: 10.1186/s43163-021-00102-z.
2. Niedzielski A, Chmielik LP, Mielnik-Niedzielska G, Kasprzyk A, Bogusławska J. Adenoid hypertrophy in children: a narrative review of pathogenesis and clinical relevance. *BMJ Paediatr Open.* 2023;7(1):e001710. doi: 10.1136/bmjpo-2022-001710.
3. Pinto I, Firmino-Machado J, Castro E, Alves S, Helena D, Condé A. The effects on the acoustic parameters and auditory-perceptive characteristics of voice in children submitted to adenoidectomy with or without tonsillectomy. *Int J Pediatr Otorhinolaryngol.* 2019;125:51-5. doi: 10.1016/j.ijporl.2019.06.021.
4. Sivasankar M, Fisher KV. Oral breathing increases Pth and vocal effort by superficial drying of vocal fold mucosa. *J Voice.* 2002;16(2):172-81. doi: 10.1016/s0892-1997(02)00087-5.
5. Lundeborg I, Hultcrantz E, Ericsson E, McAllister A. Acoustic and perceptual aspects of vocal function in children with adenotonsillar hypertrophy–effects of surgery. *J Voice.* 2012;26(4):480-7. doi: 10.1016/j.jvoice.2010.11.003.
6. Kim JW, Kim HJ, Lee WH, et al. Comparative Study for Efficacy and Safety of Adenoidectomy according to the Surgical Method: A Prospective Multicenter Study. *PLoS One*. 2015;10(8):e0135304. doi: 10.1371/journal.pone.0135304.
7. Ramos SD, Mukerji S, Pine HS. Tonsillectomy and adenoidectomy. *Pediatr Clin North Am.* 2013;60(4):793-807. doi: 10.1016/j.pcl.2013.04.015.
8. Kummer AW, Marshall JL, Wilson MM. Non-cleft causes of velopharyngeal dysfunction: implications for treatment. *Int J Pediatr Otorhinolaryngol.* 2015;79(3):286-95. doi: 10.1016/j.ijporl.2014.12.036.
9. Wallner LJ, Hill BJ, Waldrop W, Monroe C. Voice changes following adenotonsillectomy. A study of velar function by cinefluorography and video tape. *Laryngoscope.* 1968;78(8):1410-8. doi: 10.1288/00005537-196808000-00011.
10. Knösel M, Klein S, Bleckmann A, Engelke W. Coordination of tongue activity during swallowing in mouth-breathing children. *Dysphagia.* 2012;27(3):401-7. doi: 10.1007/s00455-011-9383-8.
11. Jobin V, Champagne V, Beauregard J, Charbonneau I, McFarland DH, Kimoff RJ. Swallowing function and upper airway sensation in obstructive sleep apnea. *J Appl Physiol (1985).* 2007;102(4):1587-94. doi: 10.1152/japplphysiol.00439.2006.
12. Dodds WJ, Stewart ET, Logemann JA. Physiology and radiology of the normal oral and pharyngeal phases of swallowing. *AJR Am J Roentgenol.* 1990;154(5):953-63. doi: 10.2214/ajr.154.5.2108569.
13. Matsuo K, Palmer JB. Anatomy and physiology of feeding and swallowing: normal and abnormal. *Phys Med Rehabil Clin N Am.* 2008;19(4):691-707. doi: 10.1016/j.pmr.2008.06.001.
14. Ayuse T, Hisamatsu N, Yamaguchi T, et al. Efficacy of nasal high flow therapy on the coordination between breathing and swallowing of saliva during daytime nap in chronic obstructive pulmonary disease patients: A single center, randomized crossover controlled study. *Medicine (Baltimore).* 2020;99(34):e21778. doi: 10.1097/MD.0000000000021778.
15. Clavé P, Shaker R. Dysphagia: current reality and scope of the problem. *Nat Rev Gastroenterol Hepatol.* 2015;12(5):259-70. doi: 10.1038/nrgastro.2015.49.
16. Ricci-Maccarini A, De Maio V, Murry T, Schindler A. Development and validation of the children's voice handicap index-10 (CVHI-10). *J Voice.* 2013;27(2):258.e23-258.e28. doi: 10.1016/j.jvoice.2012.10.006.
17. Tadihan Özkan E, Tüzüner A, Çiyiltepe M, Balo E, Özçelik Korkmaz M, Çağli A. Reliability and validity of the Turkish children's voice handicap index-10 (TR-CVHI-10). *Int J Pediatr Otorhinolaryngol.* 2017;96:131-4. doi: 10.1016/j.ijporl.2017.03.015.
18. Belafsky PC, Mouadeb DA, Rees CJ, et al. Validity and reliability of the Eating Assessment Tool (EAT-10). *Ann Otol Rhinol Laryngol.* 2008;117(12):919-24. doi: 10.1177/000348940811701210.
19. Zhang PP, Yuan Y, Lu DZ, et al. Diagnostic Accuracy of the Eating Assessment Tool-10 (EAT-10) in Screening Dysphagia: A Systematic Review and Meta-Analysis. *Dysphagia.* 2023;38(1):145-58. doi: 10.1007/s00455-022-10486-6.
20. Demir N, Serel Arslan S, İnal Ö, Karaduman AA. Reliability and Validity of the Turkish Eating Assessment Tool (T-EAT-10). *Dysphagia.* 2016;31(5):644-9. doi: 10.1007/s00455-016-9723-9.

https://doi.org/10.5455/annalsmedres.2025.11.356

21. Serel Arslan S, Demir N, Karaduman AA, Belafsky PC. The Pediatric Version of the Eating Assessment Tool: a caregiver administered dyphagia-specific outcome instrument for children. *Disabil Rehabil.* 2018;40(17):2088-92. doi: 10.1080/09638288.2017.1323235.
22. Abdel-Aziz M, Dewidar H, El-Hoshy H, Aziz AA. Treatment of persistent post-adenoidectomy velopharyngeal insufficiency by sphincter pharyngoplasty. *Int J Pediatr Otorhinolaryngol.* 2009;73(10):1329-33. doi: 10.1016/j.ijporl.2009.05.026.
23. Konstantinidou S, Hall A, Pinkstone M, Cochrane L. Paediatric velopharyngeal insufficiency following adenotonsillar surgery. *Int J Pediatr Otorhinolaryngol.* 2021;149:110847. doi: 10.1016/j.ijporl.2021.110847.
24. Abdel-Aziz M, Nassar A, Nashed R, Elsherbeeny M, Sabry O. The benefits of endoscopic look after curettage adenoidectomy. *Egypt J Otolaryngol.* 2020;36(26). doi: 10.1186/s43163-020-00027-z.
25. Lee JS, Keum BR, Kim JE, Park IS, Kim H. Prevalence of Dysphonia in Children with Adenotonsillar Problems and the Impact of Surgery on Voice. *Laryngoscope.* 2021;131(10):2369-75. doi: 10.1002/lary.29507.
26. Salami A, Jankowska B, Dellepiane M, Crippa B, Mora R. The impact of tonsillectomy with or without adenoidectomy on speech and voice. *Int J Pediatr Otorhinolaryngol.* 2008;72(9):1377-84. doi: 10.1016/j.ijporl.2008.05.019.
27. Naraghi M, Adil S, Bastaninejad S, Dabiran S. Evaluation of pediatric voice handicap index and pediatric voice related quality of life before and after adenotonsillectomy in pediatric population. *Int J Pediatr Otorhinolaryngol.* 2015;79(3):388-91. doi: 10.1016/j.ijporl.2014.12.034.
28. Karakurt SE, Cetin MA, Yamur AR, Ikinciogullari A, Ensari S, Dere HH. Is Adenotonsillar Size a Significant Factor on Voice in Children Undergoing Adenotonsillectomy? *J Coll Physicians Surg Pak.* 2019;29(6):524-7. doi: 10.29271/jcpsp.2019.06.524.
29. Abdel-Aziz M, Hady AFA, Sheikhany AR, Yousef AI, Sabry OA, Farag HM. Effect of adenoid size on the post-adenoidectomy hypernasality in children with a normal palate. *Eur Arch Otorhinolaryngol.* 2023;280(10):4555-60. doi: 10.1007/s00405-023-08049-y.
30. Liu X, Zheng Y, Tian P, Yang J, Zou H. The impact of tonsillectomy with or without adenoidectomy on voice: acoustic and aerodynamic assessments. *J Voice.* 2015;29(3):346-8. doi: 10.1016/j.jvoice.2014.08.014.
31. Brkic FF, Liu DT, Campion NJ, et al. Changes in Acoustic Aspects of Vocal Function in Children After Adenotonsillectomy. *J Voice.* 2022;36(3):438.e19-438.e24. doi: 10.1016/j.jvoice.2020.06.022.
32. Inja RR, Paul RR, Varghese L, Santosh S, Sebastian T, Mathews SS. Voice Change Following Adenotonsillectomy in Pediatric Population: Myth or Reality?-A Pilot Study. *Indian J Otolaryngol Head Neck Surg.* 2019;71(Suppl 1):574-9. doi: 10.1007/s12070-018-1412-6.
33. Wang ST, Kang KT, Chang CF, Lin MT, Hsu WC. Voice Change After Adenotonsillectomy in Children: A Systematic Review and Meta-Analysis. *Laryngoscope.* 2024;134(6):2538-50. doi: 10.1002/lary.31140.
34. Lundeborg I, McAllister A, Graf J, Ericsson E, Hultcrantz E. Oral motor dysfunction in children with adenotonsillar hypertrophy–effects of surgery. *Logoped Phoniatr Vocol.* 2009;34(3):111-6. doi: 10.1080/14015430903066937.
35. de Oliveira Branco AA, de Castro Corrêa C, de Souza Neves D, Huehara T, Theresa Weber SA. Swallowing patterns after adenotonsillectomy in children. *Pediatr Investig.* 2019;3(3):153-8. doi: 10.1002/ped4.12142.
36. Walsh ME, Retzler R, Huang J, Daglish A, Tweedie D, Pepper C. A prospective cohort study exploring the impact of tonsillectomy on feeding difficulties in children. *Clin Otolaryngol.* 2024;49(3):314-9. doi: 10.1111/coa.14148.

https://doi.org/10.5455/annalsmedres.2025.11.356

Ann Med Res

Current issue list available at Ann Med Res

Annals of Medical Research

journal page: annalsmedres.org

HBV immunity gaps in older adults undergoing cataract surgery: Opportunistic screening and linkage to care

Ezgi Karatas[a], *, Seher Koksaldi[a], Mustafa Kayabasi[b], Yavuz Oruc[a]

[a]Ağrı İbrahim Çeçen University, Faculty of Medicine, Department of Ophthalmology, Ağrı, Türkiye
[b]Muş State Hospital, Clinic of Ophthalmology, Muş, Türkiye
*Corresponding author: e.karatas.2015@gmail.com (Ezgi Karatas)

MAIN POINTS

- Nearly half (47.0%) of cataract surgery candidates had non-protective anti-HBs (<10 mIU/mL), and two-thirds (67.7%) had levels ≤100 mIU/mL, indicating substantial immunity gaps.
- HBsAg positivity was 4.4%, highlighting silent carriers identified at preoperative screening with infection-control implications.
- Female sex was independently associated with non-protective titers (aOR 1.49, 95% CI 1.04–2.13).
- Pragmatic contribution: routine preoperative serology in ophthalmology functions as an opportunistic public-health checkpoint to trigger booster vaccination and linkage-to-care.

Cite this article as: Karatas E, Koksaldi S, Kayabasi M, Oruc Y. HBV immunity gaps in older adults undergoing cataract surgery: Opportunistic screening and linkage to care. *Ann Med Res*. 2026;33(6):292–299. doi: 10.5455/annalsmedres.2025.10.328.

ABSTRACT

Aim: To assess whether routine preoperative testing in cataract surgery candidates can serve as a practical public health checkpoint for detecting waning hepatitis B immunity and initiating booster-based protection pathways.

Materials and Methods: In this descriptive cross-sectional study, preoperative anti-HBs, HBsAg, anti-HCV, and anti-HIV serologies were evaluated in 430 patients (January 2023 -- December 2024) who underwent cataract surgery. Anti-HBs levels were classified as non-protective (<10 mIU/mL), low-protective (10–<100 mIU/mL), or highly protective (≥100 mIU/mL). Chi-square tests and logistic regression were used to analyze immunity levels and associated risk factors.

Results: Among 430 patients (mean age 67.4 ± 9.6 years; 47.4% women), anti-HBs titers were <10 mIU/mL in 47.0% (202/430), 10–<100 mIU/mL in 20.7% (89/430), and ≥100 mIU/mL in 32.3% (139/430). Women had higher odds of non-protective titers (<10 mIU/mL) than men (OR 1.52, 95% CI 1.04–2.23; P=0.031), which persisted after adjustment for age and comorbidity (aOR 1.49, 95% CI 1.04–2.13; P=0.028). HBsAg positivity was 4.4% (19/430); anti-HCV was 2.1% (9/430) and HIV was 0%. Across major comorbidity groups, anti-HBs distributions were broadly similar.

Conclusion: A substantial proportion of patients undergoing cataract surgery lack adequate HBV immunity, emphasizing the need for routine preoperative serological screening and booster vaccination. This pioneering study from Türkiye provides valuable insights into infection control strategies for ophthalmic surgery.

Keywords: Cataract surgery, Hepatitis B virus, Hepatitis B virus antibodies, Infection control, Vaccination

Received: Nov 03, 2025 **Accepted:** Jan 29, 2026 **Available Online:** Jun 25, 2026

INTRODUCTION

Cataract surgery is minimally invasive and involves minimal blood loss. Although the risk of blood-borne infection transmission is infrequent, universal precautions and preoperative screening protocols are essential to ensure the safety of both patients and staff. The incidence of needlestick (sharps) injuries in ophthalmic practice has been reported to be 0.06–0.08 per 1,000 procedures; although low, this risk cannot be dismissed given the potential severity of HBV and other blood-borne infections [1]. Asymptomatic seropositive individuals, who are often unaware of their infection status, present unique challenges in perioperative settings, necessitating stringent infection control protocols [2].

HBV remains a major global public health challenge, with an estimated 296 million people living with chronic HBV infection in 2019, resulting in approximately 820,000 deaths annually due to liver cirrhosis and hepatocellular carcinoma [3]. The World Health

Organization (WHO) classifies regions based on hepatitis B surface antigen (HBsAg) prevalence (<2%) [4]. In Türkiye, HBsAg prevalence has decreased from 12.3% in 2000 to 4–6% in recent years, largely due to the national HBV vaccination program initiated in 1998 [5]. However, regional dispari-

ties persist, with a higher prevalence (up to 10%) in Eastern Türkiye than in metropolitan areas [6]. Globally, approximately one-third of the population is exposed to HBV, underscoring the need for robust preventive strategies [7].

Vaccination has significantly reduced HBV incidence [8]. However, long-term immunity wanes in some individuals, particularly older adults who may not have been vaccinated during childhood [9]. Cataract surgery, one of the most common procedures globally, involves older adults who may have waning immunity due to incomplete vaccination coverage in earlier decades. Studies indicated that anti-HBs levels decline over time, with seroprotection rates ranging from 30% to 80% in adults, depending on age, vaccination history, and comorbidities [10].

In surgical settings, where the risk of HBV transmission is heightened owing to exposure to blood and bodily fluids, preoperative serological screening is increasingly recognized as essential [11]. While HBV screening is routine in high-risk surgeries, its application in

elective procedures, such as cataract surgery, remains underexplored, particularly in

intermediate-prevalence settings such as Türkiye.

This study aimed to characterize the preoperative HBV serological status of cataract surgery candidates by quantifying anti-HBs levels to identify individuals with inadequate immunity, guide targeted booster vaccination, and strengthen infection control practices in ophthalmic settings, thereby supporting the global HBV elimination goals for 2030 [12]. We defined seroprotection as anti-HBs ≥10 mIU/mL and, given blunted immune responses in certain high-risk settings (e.g., haemodialysis, profound immunosuppression), considered ≥100 mIU/mL a conservative benchmark cited in existing guidance [13]. Accordingly, we reported anti-HBs in three strata (<10, 10–<100, ≥100 mIU/mL) and evaluated factors associated with non-protection. Our objective was to characterize serological protection levels rather than the aetiology of immunity; accordingly, differentiation between vaccine- and infection-induced anti-HBs responses lay outside the study's predefined scope.

MATERIALS AND METHODS

This study was approved by the Ağrı İbrahim Çeçen University Scientific Research Ethics Committee (date: 26.12.2024, decision no: 500) and adhered to the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrolment. This cross-sectional study (January 2023–December 2024) was designed to quantify the degree and robustness of protective immunity against HBV—using anti-HBs titer strata (<10, 10–<100, ≥100 mIU/mL) in patients who underwent cataract surgery, based on blood samples obtained preoperatively. We also estimated the preoperative seroprevalence of HBV (HBsAg), HCV (anti-HCV), and HIV (anti-HIV). Inclusion required a clinical diagnosis of cataract and written informed consent; patients with incomplete serologic data were excluded. Before surgery, all participants underwent a standard ophthalmic examination. As part of the routine preoperative work-up, venous blood was collected before surgery for serologic testing (quantitative anti-HBs [mIU/mL], qualitative HBsAg, anti-HCV antibody, and HIV antibody) using chemiluminescent immunoassays according to manufacturers' instructions.

Demographic characteristics, medical history, and comorbidities were recorded. The results were categorized as follows:

- Non-protective: anti-HBs <10 mIU/mL
- Low protective: anti-HBs 10–<100 mIU/mL
- Highly protective: anti-HBs ≥100 mIU/mL [13].

Sample size calculation

Based on previous reports indicating that the prevalence of non-protective anti-HBs levels in older adults ranges between 40% and 50%, a prevalence estimate of p = 0.47 was assumed. With a 95% confidence level (Z = 1.96) and a margin of error of ±5% (d = 0.05), the minimum required sample size was calculated as 383 patients [14,15].

Statistical analysis

Analyses were performed using IBM SPSS Statistics, version 26 (IBM Corp., Armonk, NY). Descriptive statistics summarized demographics and serologic measures. Normality of continuous variables was assessed using the Shapiro–Wilk test and visual inspection of histograms and Q–Q plots. Normally distributed variables were presented as means ± standard deviations (SD) and compared using Student's t tests, whereas non-normally distributed variables were summarized as median (interquartile range, IQR) and compared using the Mann–Whitney U test. Comparisons across anti-HBs categories (<10, 10-<100, and ≥100 mIU/mL) were performed using one-way ANOVA for normally distributed variables or the Kruskal–Wallis test for non-normally distributed variables. For variables showing overall significance, post hoc pairwise comparisons were conducted with Bonferroni adjustment. Categorical variables were expressed as counts and percentages and compared using χ^2 tests or Fisher's exact test, as appropriate. Multivariable logistic regression analysis was performed to identify independent factors associated with non-protective anti-HBs status (<10 mIU/mL), and results were reported as odds ratios (ORs) with 95% confidence intervals (CIs). Two-sided P values < 0.05 were considered statistically significant.

RESULTS

Among 430 patients (67.4 ± 9.6 years; 204/430 [47.4%] women), anti-HBs titers were <10 mIU/mL in 202 (47.0%), 10–<100 mIU/mL in 89 (20.7%), and ≥100 mIU/mL in 139 (32.3%). Women were more likely than men to have non-protective titers <10 mIU/mL (107/204 [52.5%] vs 95/226

https://doi.org/10.5455/annalsmedres.2025.10.328

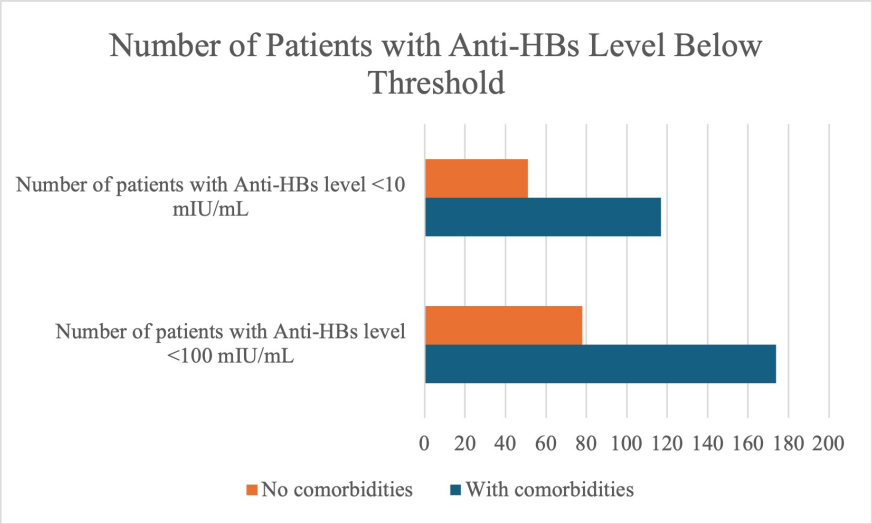

Figure 1. Represents the number of cataract surgery patients with anti-HBs levels below the immunity thresholds (<10 mIU/mL and <100 mIU/mL), categorized by the presence or absence of comorbidities. [Grouped bar chart showing hepatitis B surface antibody (anti-HBs) categories (<10, 10-<100, ≥100 mIU/mL) by comorbidity status (none vs ≥1 chronic disease). Distributions are broadly similar between groups; no significant global difference is observed. The largest category overall is anti-HBs <10 mIU/mL (202/430; 47.0%), followed by 10-<100 mIU/mL (89/430; 20.7%) and ≥100 mIU/mL (139/430; 32.3%). This figure emphasizes the sizeable proportion with non-protective titers, supporting preoperative screening and linkage to booster vaccination].

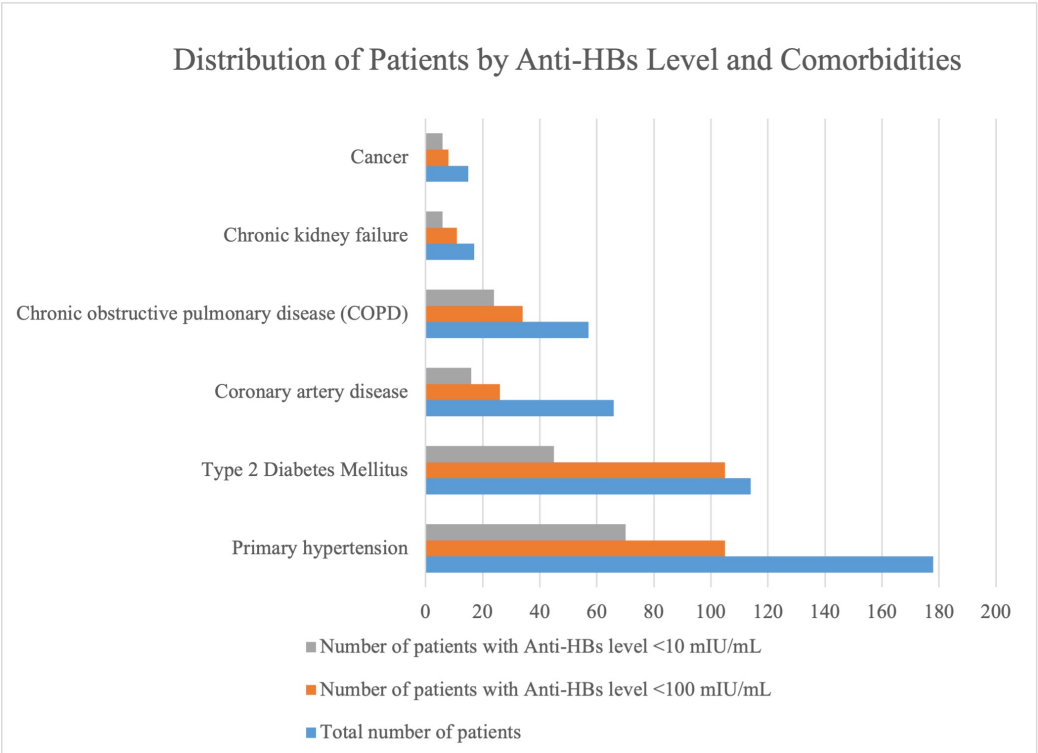

Figure 2. Depicts patient distribution according to Anti-HBs levels and comorbid conditions. [Bar chart comparing hepatitis B surface antibody (anti-HBs) categories (<10, 10-<100, ≥100mIU/mL) across multiple chronic conditions: primary hypertension, type 2 diabetes mellitus, coronary arterydisease, chronic obstructive pulmonary disease, chronic kidney disease, and a no-comorbidity referencegroup. Each condition displays three bars reflecting the distribution of anti-HBs categories. Overall patternsare broadly similar across conditions; in most groups, <10 mIU/mL (non-protective) constitutes the largestshare, while ≥100 mIU/mL indicates a smaller, long-term immunity subgroup. A modest tilt toward ≥100mIU/mL is observable in the coronary artery disease group, whereas hypertension and diabetes approximatecohort averages. The figure underscores that non-protective titers are prevalent across comorbidity profiles, reinforcing the rationale for preoperative screening and linkage to booster vaccination].

[42.0%]; χ^2 test, P=0.031). In age-adjusted models (age and any comorbidity), female sex remained associated with anti-HBs <10 mIU/mL (aOR 1.49, 95% CI 1.04–2.13; P=0.028). The proportion with anti-HBs ≥100 mIU/mL was 32.3% (139/430), while 67.7% (291/430) had titers <100 mIU/mL (Table 1). HBsAg was positive in 19/430 (4.4%) patients;

https://doi.org/10.5455/annalsmedres.2025.10.328

Table 1. Demographic characteristics and anti-HBs* antibody concentrations of the patients.

Parameters	Female (n, %)	Male (n, %)	P-value
Number (n, percentage)	204 (47.4%)	226 (52.6%)	0.2
Age (mean ± SD)	68.21 ± 10.2	66.65 ± 8.8	0.09
Anti-HBs level (<10 mIU/mL)	107 (52.4%)	95 (44%)	0.03
Anti-HBs level (10-<100 mIU/mL)	31 (15.2%)	58 (25.6%)	0.00
Anti-HBs level (≥100 mIU/mL)	66 (47.4%)	73 (52.5%)	0.9

* Antibodies against Hepatitis B surface antigen.

Table 2. Distribution of patients according to systemic chronic diseases and gender.

Diseases	Female n (%)	Male n (%)	P-value
Primary hypertension	116 (56.9%)	62 (27.4%)	<0.001
Type 2 diabetes mellitus	72 (35.3%)	42 (18.6%)	<0.001
Coronary artery disease	31 (15.2%)	35 (15.5%)	1.0
Chronic obstructive pulmonary disease	23 (11.2%)	34 (15.0%)	0.25
Cancer	3 (1.5%)	12 (5.3%)	0.03
Chronic Hepatitis B	5 (2.5%)	7 (3.1%)	0.68
Chronic Hepatitis C	0 (0.0%)	2 (0.9%)	0.5
Other*	68 (33.0%)	54 (23.0%)	0.2

* Other diseases include aortic aneurysm, hydatid cyst, nephrectomy, ankylosing spondylitis, systemic lupus erythematosus, cirrhosis, polycythemia vera, psychosis, nephrotic syndrome, rheumatoid arthritis, bronchitis, hepatosteatosis, epilepsy, cerebrovascular occlusion, heart failure, chronic renal failure, and dermatomyositis.

12/19 (63.2%) met criteria for chronic hepatitis B, and the remainder represented previously undiagnosed current infection identified at preoperative screening. Results were unchanged after excluding HBsAg-positive cases (n=19) (all P>0.05).

Based on chronic disease status, 111 patients had no comorbidities, whereas 319 had ≥1 comorbidity. The most frequent conditions were hypertension (179/430, 41.6%), type 2 diabetes mellitus (116/430, 27.0%), coronary artery disease (66/430, 15.3%), and chronic obstructive pulmonary disease (59/430, 13.7%) (Table 2). Table 3 summarizes anti-HBs strata across comorbidity groups and specific diagnoses (HT, T2DM, CAD, COPD, CKD); Figure 1 and Figure 2 visualizes these patterns.

In multivariable logistic regression analysis (Table 4), female sex remained independently associated with non-protective anti-HBs status (≤10 mIU/mL) (adjusted OR [aOR] 1.74, 95% CI 1.14–2.63; P=0.010). Increasing age was associated with lower odds of non-protective titers (aOR 0.96 per 1-year increase, 95% CI 0.94–0.99; P=0.001). COPD was associated with increased odds of anti-HBs ≤10 mIU/mL (aOR 1.80, 95% CI 1.01–3.22; P=0.048), whereas diabetes mellitus, hypertension, and coronary artery disease were not significant predictors (all P>0.05).

DISCUSSION

This study evaluated the preoperative serological status of patients scheduled for cataract surgery by assessing markers of HBV, HCV, and HIV infection. Among 430 participants, 47.0% (202/430) had anti-HBs <10 mIU/mL, indicating insufficient protection against HBV. Additionally, 20.7% (89/430) had anti-HBs 10–<100 mIU/mL, suggesting waning long-term immunity. The proportion with long-term immunity (≥100 mIU/mL) was 32.3% (139/430), whereas 67.7% (291/430) had titers <100 mIU/mL, indicating a substantial gap in durable protection. These results highlight the substantial proportion of patients at risk of HBV infection in the surgical setting, particularly among older adults. In addition, HBsAg positivity was 4.4% (19/430). The coexistence of limited durable protection, a sizable pool of HBV carriers, and a non-negligible subset of previously undiagnosed cases identified during preoperative screening underscores the need for targeted booster vaccination, systematic linkage to care, and strict adherence to universal precautions in ophthalmic theatres. Our primary aim was to quantify the degree and robustness of HBV protective immunity using anti-HBs titer strata (<10, 10–<100, ≥100 mIU/mL); a secondary aim was to estimate the preoperative seroprevalence of HBV (HBsAg), HCV, and HIV. Differentiating vaccine-induced from infection-induced immunity was not a prespecified objective.

Following vaccination or natural infection, anti-HBs antibody levels gradually decline and may fall below protective thresholds within around 5 to 20 years [16–19]. The rate and degree of antibody decline are influenced by factors like vaccination age, the original peak antibody response, and certain health problems (e.g., immunosuppression, comorbidities) [20–22]. While primary HBV vaccination provides enduring protection, the decline of humoral immunity requires continuous monitoring, especially in elderly individuals who may not have received the universal childhood immunization implemented in Türkiye in 1998 [23].

Although the incidence of breakthrough infections among vaccinated individuals remains low, the observed decrease in anti-HBs levels raises concerns regarding the durability of vaccine-induced protection [24]. T-lymphocyte-mediated immune memory often persists longer than humoral responses; however, a progressive decline in this memory has been documented, with anamnestic response rates decreasing from 85.3% at 10 years to 73.6% at 15 years post-vaccination [24,25]. This attrition could lead to HBsAg seroconversion and increased vulnerability to HBV infection.

In accordance with previous booster-response studies, those with very low pre-booster titers (<2 mIU/mL) are more likely to demonstrate only marginal post-booster responses compared to those with detectable although sub-protective titers (2.00–9.99 mIU/mL) [25–27]. Earlier research found that 65% of adults with pre-booster <10 mIU/mL had undetectable titers (<2 mIU/mL). After the booster, low responders with <2 mIU/mL were overrepresented among those who stayed <10 mIU/mL (≈11% vs 2.3% for the 2.00–9.99 mIU/mL subgroup), and they also more often fell into the 10–100 mIU/mL band (≈26.6% vs 10.2%) instead of exceed-

https://doi.org/10.5455/annalsmedres.2025.10.328

Table 3. Distribution of Anti-HBs Titer Strata (<10, 10-<100, ≥100 mIU/mL) by Overall Comorbidity Statusand Specific Chronic Diseases, with P-values.

Variable	Group	<10 mIU/mL	10–<100 mIU/mL	≥100 mIU/mL	Total	P value (test)
Any comorbidity	No	54 (48.6%)	25 (22.5%)	32 (28.8%)	111	0.639 (χ^2)
	Yes	148 (46.4%)	64 (20.1%)	107 (33.5%)	319	
Hypertension	No	118 (47.0%)	49 (19.5%)	84 (33.5%)	251	0.724 (χ^2)
	Yes	84 (46.9%)	40 (22.3%)	55 (30.7%)	179	
Type 2 diabetes mellitus	No	148 (47.1%)	68 (21.7%)	98 (31.2%)	314	0.616 (χ^2)
	Yes	54 (46.6%)	21 (18.1%)	41 (35.3%)	116	
Coronary artery disease	No	179 (49.2%)	76 (20.9%)	109 (29.9%)	364	0.036 (χ^2)
	Yes	23 (34.8%)	13 (19.7%)	30 (45.5%)	66	
COPD	No	170 (45.8%)	75 (20.2%)	126 (34.0%)	371	0.191 (χ^2)
	Yes	32 (54.2%)	14 (23.7%)	13 (22.0%)	59	
Chronic kidney disease	No	197 (47.2%)	86 (20.6%)	134 (32.1%)	417	0.819 (FE)
	Yes	5 (38.5%)	3 (23.1%)	5 (38.5%)	13	

Anti-HBs: hepatitis B surface antibody; HT: hypertension; T2DM: type 2 diabetes mellitus; CAD: coronary artery disease; COPD (KOAH): chronic obstructive pulmonary disease; CKD: chronic kidney disease; Data are presented as n (%). P values were calculated using Pearson's chi-square test (χ^2) unless otherwise indicated. FE indicates Fisher's exact test, used when expected cell counts were <5.

Table 4. Multivariable logistic regression analysis for predictors of non-protective anti-HBs (≤10 mIU/mL).

Predictor	β (SE)	Wald χ^2	Adjusted OR	95% CI	P value
Female sex (vs male)	0.551 (0.213)	6.70	1.74	1.14–2.63	0.010
Age (per 1-year increase)	-0.036 (0.011)	10.71	0.96	0.94–0.99	0.001
Diabetes mellitus	-0.202 (0.235)	0.74	0.82	0.52–1.30	0.390
Hypertension	0.040 (0.224)	0.03	1.04	0.67–1.62	0.858
Coronary artery disease	-0.499 (0.290)	2.96	0.61	0.34–1.07	0.085
COPD	0.587 (0.297)	3.91	1.80	1.01–3.22	0.048
Constant	2.656 (0.766)	12.01	---	---	<0.001

OR: odds ratio; CI: confidence interval; SE: standard error; AUC: area under the receiver operating characteristic curve. Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test. Constant indicates the intercept term of the logistic regression model.

ing 100 mIU/mL [26]. This pattern is consistent with waned immune memory rather than a strong anamnestic response. Interestingly, individuals who stayed below 10 mIU/mL after one booster shot were able to respond well following a full secondary vaccination course. This suggests that they were not true primary "non-responders" but had lost their immunological memory about 18–20 years after their first immunization. Our study did not give boosters or check anti-HBc levels, but our preoperative distribution, which showed that 19.5% of participants fell in the 10–100 mIU/mL range, likely includes a large group of people with partial or waning immunity who could benefit from targeted booster vaccination and, if possible, post-booster titer checks to confirm durability. From a practical point of view, these data support a pragmatic approach in ophthalmic settings: (i) offer boosters to patients with anti-HBs 10–100 mIU/mL (and strongly to those <10 mIU/mL), (ii) consider aiming for ≥100 mIU/mL in higher-risk contexts (e.g., haemodialysis, profound immunosuppression), and (iii) reserve etiologic attribution (vaccine- vs infection-derived immunity) for settings where anti-HBc is available. This approach is in line with our main goal, which is to map actionable immunity gradients before surgery.

Data from immunosuppressed groups shows that having an anti-HBs level of at least 100 mIU/mL (100mIU/mL) protects against clinically significant HBV reactivation. On the other hand, patients with low (≤100 mIU/mL) or no anti-HBs levels have higher rates of reactivation and more frequent HBsAg seroreversion during treatments like anti-TNF [28–30]. These findings underscore the therapeutic significance of baseline quantitative anti-HBs prior to the commencement of anti-TNF (and other potent immunosuppressants) and advocate for vaccination/booster dose to elevate titers to ≥100 mIU/mL before and/or during treatment. Our study did not assess treatment-related reactivation; nevertheless, the preoperative identification of individuals with 10–<100 mIU/mL or <10 mIU/mL offers a chance to identify vaccination requirements and facilitate care coordination with treating specialists during planned or continuing immunosuppression.

In this perspective, it is significant that, despite a substantial reduction in acute HBV incidence in Italy, 212 cases were recorded in 2018 (incidence 0.8 per 100,000), including 16 cases among vaccinated patients, five of whom had seemingly completed a full and properly timed vaccination series [26]. This persistent burden, especially in highly vaccinated environments, highlights the necessity of quantitative anti-HBs evaluation, targeted boosters for diminishing immunity, and the facilitation of care for carriers detected opportunistically

https://doi.org/10.5455/annalsmedres.2025.10.328

in surgical procedures.

Recent guidelines for haemodialysis patients suggest keeping anti-HBs levels at least 100 mIU/mL because they tend to go down over time [28–30]. As a result, experts recommend checking anti-HBs levels once a year and giving a booster shot if they drop below 100 mIU/mL. A comparable high-titer approach has been utilized following liver transplantation to avert HBV-related disease. In line with this methodology, Khiangte et al. found no HBV reactivation in individuals who sustained anti-HBs levels exceeding 100 mIU/mL, despite the lack of antiviral treatment [31,32]. These studies underscore the clinical significance of identifying preoperative patients between the 10–<100 mIU/mL or <10 mIU/mL categories and prioritizing targeted boosters to provide more sustained protection, particularly among high-risk populations such as those with chronic kidney disease or haemodialysis and immunosuppressed persons. In accordance with this, our group of participants exhibited significant immune deficiencies within comorbidity subgroups: 39.5% of patients with diabetes mellitus and 64.7% of those with chronic renal disease demonstrated reduced or non-existent seroprotection (anti-HBs <100 mIU/mL, including <10 mIU/mL). These results highlight the necessity of regular preoperative screening for hepatitis B indicators in cataract surgical protocols, particularly for elderly or immunocompromised individuals, and advocate for targeted booster immunization to improve seroprotection in these groups.

Meta-analytic evidence indicates diminished seroprotection following HB vaccination in older adults, individuals with a body mass index ≥25, smokers, and those with comorbidities [33]. Other pooled data show that in people who are immunocompetent and have been properly vaccinated, protection usually lasts for about 20 years, with just a few short-lived HBsAg seroconversions [34]. In light of this, our preoperative screening of patients who underwent cataract surgery showed that a lot of older people had weak immunity: 39.1% had anti-HBs levels below 10 mIU/mL, 19.5% had levels between 10 and 100 mIU/mL, and 39.5% of patients with diabetes and 64.7% of patients with chronic kidney disease had levels below 100 mIU/mL. These findings advocate for a risk-stratified, opportunistic strategy within cataract pathways: uphold standard vaccination practices universally, while identifying and administering targeted boosters to individuals with anti-HBs <10 mIU/mL or 10–<100 mIU/mL—especially older adults and patients with diabetes or chronic kidney disease—to augment seroprotection identified during routine preoperative assessments.

Although our cohort showed a higher proportion of women with anti-HBs <10 mIU/mL, multiple studies and reviews report that males often mount weaker HBV vaccine responses than females, with sex-based differences observed across regions and settings [33,35–37]. When combined, our female-dominant non-protection may indicate cohort-specific factors (e.g., demographic characteristics, comorbidity burden, time since vaccination, undocumented vaccination histories) or residual confounding if body mass index (BMI), metabolic status, or the interval from vaccination to testing varied by sex. From a methodological standpoint, these observations endorse the presentation of sex-stratified results in conjunction with adjusted models that incorporate BMI and the interval since vaccination. Operationally, irrespective of sex-based directionality in a specific context, the identification of individuals with anti-HBs <10 mIU/mL or 10–<100 mIU/mL continues to be therapeutically relevant for targeted booster methods identified via standard preoperative screening.

Furthermore, the detection of HBsAg positivity in 4.4% of patients (n:19; females: 8 [42.1%], males: 11 [57.9%])— which is comparable to the reported national prevalence [38–40] - underscores the silent burden of HBV infection in surgical populations. Routine serological screening facilitates early diagnosis and targeted infection control interventions, thereby reducing intraoperative transmission risk. Although anti-HCV positivity was detected in only 0.5% of patients and HIV seroprevalence was minimal, universal screening remains essential to ensure the safety of patients and staff.

Comprehensive infection control protocols were implemented for seropositive patients undergoing surgery. These included preoperative infectious disease consultations, enhanced barrier precautions (double gloves, impermeable gowns, and eye protection), use of dedicated surgical instruments, and meticulous administration of local anaesthesia to prevent needlestick injuries. The surgical team received targeted training on universal precautions and post-exposure management protocols, which significantly mitigated the risk of blood-borne pathogen transmission. Routine booster doses are not advised for the general populace in low-endemicity areas; nevertheless, our context indicates moderate endemicity (HBsAg ≈4%), where exposure risk is significant, especially among older persons and individuals with comorbidities. In light of our preoperative findings of substantial immunity gaps (anti-HBs <10 or 10–<100 mIU/mL), a targeted, opportunistic booster strategy integrated into preoperative workflows is warranted: evaluate seroprotection, administer boosters as necessary, and facilitate care linkage for HBsAg-positive individuals. The implementation of WHO recommendations has altered HBV endemicity in several Asian nations; nonetheless, the underlying risk remains variable, being elevated in certain contexts and moderate in others. Guidance established in low-endemicity regions (Europe/North America) typically emphasizes the completion of the primary series and does not support routine boosters for the general population—except for immunocompromised groups (e.g., chronic kidney disease/haemodialysis, advanced immunosuppression), for whom periodic monitoring and booster dosing are recommended.

Translating these principles to moderate-endemicity contexts like ours, the higher lifetime probability of exposure argues

https://doi.org/10.5455/annalsmedres.2025.10.328

against blanket boosting yet for a risk-stratified, opportunistic approach: use the preoperative visit to identify low/non-protected patients (anti-HBs <10 or 10–<100 mIU/mL), offer targeted boosters, and ensure linkage to care for those who are HBsAg-positive.

Limitations

This study is limited by its cross-sectional, single-centre design, which restricts causal inference and limits the generalizability of the findings. Longitudinal multicentre studies are warranted to assess the durability of vaccine-induced immunity and effectiveness of booster doses in maintaining long-term protection.

CONCLUSION

This study revealed that a considerable proportion of patients undergoing cataract surgery, predominantly older adults, exhibited inadequate protective anti-HBs levels. These findings highlight the urgent need for targeted booster immunization strategies and early vaccination initiatives to ensure sustained protection against HBV infection. Enhancing serological screening and vaccination efforts within surgical settings aligns with the World Health Organization's goal of achieving 90% global hepatitis B vaccination coverage by 2030, representing a critical step towards HBV elimination. Ophthalmologists and surgical teams must maintain heightened vigilance regarding the patients' serological status to ensure optimal infection control and patient safety.

Abbreviations

HBV: Hepatitis B Virus

HBsAg: Hepatitis B Surface Antigen

Anti-HBs: Antibodies Against Hepatitis B Surface Antigen

HCV: Hepatitis C Virus

HIV: Human Immunodeficiency Virus

WHO: World Health Organization

OR: Odds Ratio

CI: Confidence Interval

SD: Standard Deviation

SPSS: Statistical Package for the Social Sciences

Ethics Committee Approval: This study was approved by the Ağrı İbrahim Çeçen University Scientific Research Ethics Committee (date: 26.12.2024, decision no: 500).

Informed Consent: Written informed consent was obtained from all participants.

Peer-review: Externally peer-reviewed.

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

Author Contributions: Conception: E.K, S.K, M.K, Y.O; Design: E.K, S.K, M.K, Y.O; Supervision: S.K, M.K, Y.O; Materials: S.K; Data Collection and/or Processing: E.K, M.K; Analysis and/or Interpretation: E.K, M.K; Literature Review: E.K, Y.O; Writing: E.K, S.K, Y.O; Critical Review: E.K, S.K, M.K, Y.O.

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

Artificial Intelligence Disclosure: The authors declare that ChatGPT (OpenAI, GPT-4 series) was used during the preparation of this manuscript for language editing, grammatical correction, and improving clarity and readability of the text. The artificial intelligence tool did not contribute to the study design, data collection, statistical analysis, data interpretation, or the generation of scientific conclusions. All content produced with the assistance of ChatGPT was carefully reviewed, verified, and approved by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript. The AI tool was not listed as an author and did not assume responsibility for any part of the work.

REFERENCES

1. Kuhar DT, Henderson DK, Struble KA, et al. Updated US Public Health Service Guidelines for the Management of Occupational Exposures to Human Immunodeficiency Virus and Recommendations for Postexposure Prophylaxis. *Infect Control Hosp Epidemiol.* 2013;34(9):875–92. doi: 10.1086/672271.
2. Nayak B, Bhatt B. Commentary: Viral serology testing before cataract surgery: Where do we stand? *Indian J Ophthalmol.* 2022;70(2):405. doi: 10.4103/ijo.IJO_2484_21.
3. WHO. Global Health Sector Strategies on HIV, Viral Hepatitis and the sexually transmitted infections. World Health Organization. 2022–30.
4. WHO. Epidemiological update: increasing mortality calls for action 03. Global Hepatitis Report, 2017. 2017:7–20.
5. Tozun N, Ozdogan O, Cakaloglu Y, Idilman R, et al. Seroprevalence of hepatitis B and C virus infections and risk factors in Turkey: A fieldwork TURHEP study. *Clin Microbiol Infect.* 2015;21(11):1020–6. doi: 10.1016/j.cmi.2015.06.028.
6. Mehmet D, Meliksah E, Serif Y, Gunay S, Tuncer Ö, Zeynep S. Prevalence of Hepatitis B Infection in the Southeastern Region of Turkey: Comparison of Risk Factors for HBV Infection in Rural and Urban Areas. *Jpn J Infect Dis.* 2005;58(1):15–9. doi: 10.7883/yoken.JJID.2005.15.
7. Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *Lancet.* 2015;386(10003):1546–55. doi: 10.1016/S0140-6736(15)61412-X.
8. Hepatitis B vaccines: WHO position paper – July 2017. https://www.who.int/publications/i/item/WER9227 (accessed April 28, 2025).
9. Leuridan E, Van Damme P. Hepatitis B and the Need for a Booster Dose. *Clin Infect Dis.* 2011;53(1):68–75. doi: 10.1093/cid/cir270.
10. Bruce MG, Bruden D, Hurlburt D, et al. Antibody Levels and Protection After Hepatitis B Vaccine: Results of a 30-Year Follow-up Study and Response to a Booster Dose. *J Infect Dis.* 2016;214(1):16–22. doi: 10.1093/infdis/jiv748.
11. Redd JT, Baumbach J, Kohn W, Nainan O, Khristova M, Williams I. Patient-to-patient transmission of hepatitis B virus associated with oral surgery. *J Infect Dis.* 2007;195(9):1311–4. doi: 10.1086/513435.

https://doi.org/10.5455/annalsmedres.2025.10.328

12. Global health sector strategy on viral hepatitis 2016-2021. Towards ending viral hepatitis. https://www.who.int/publications/i/item/WHO-HIV-2016.06 (accessed April 28, 2025).
13. Seekircher L, Mühlbacher A, Tschiderer L, et al. Anti-HBs Seroprevalence in Blood Donors from Tyrol, Austria. *Vaccines (Basel).* 2024;12(10):1156. doi: 10.3390/vaccines12101156.
14. Sahana H V., Sarala N, Prasad SR. Decrease in Anti-HBs Antibodies over Time in Medical Students and Healthcare Workers after Hepatitis B Vaccination. *Biomed Res Int.* 2017;2017:1327492. doi: 10.1155/2017/1327492.
15. Elrashidy H, Elbahrawy A, El-Didamony G, et al. Antibody levels against hepatitis B virus after hepatitis B vaccination in Egyptian diabetic children and adolescents. *Hum Vaccin Immunother.* 2013;9(9):2002–6. doi: 10.4161/hv.25426.
16. Coppeta L, Pompei A, Balbi O, et al. Persistence of Immunity for Hepatitis B Virus among Heathcare Workers and Italian Medical Students 20 Years after Vaccination. *Int J Environ Res Public Health.* 2019;1616(9):1515. doi: 10.3390/ijerph16091515.
17. Dumaidi K, Al-Jawabreh A. Persistence of Anti-HBs Among Palestinian Medical Students After 18-22 Years of Vaccination: A Cross-Sectional Study. *Hepat Mon*. 2015;15(11):e29325. doi: 10.5812/hepatmon.29325.
18. Verso MG, Costantino C, Marrella A, Immordino P, Vitale F, Amodio E. Kinetics of Anti-Hepatitis B Surface Antigen Titers in Nurse Students after a Two-Year Follow-Up. *Vaccines (Basel).* 2020;88(3):467. doi: 10.3390/vaccines8030467.
19. Phattraprayoon N, Kakheaw J, Soonklang K, et al. Duration of Hepatitis B Vaccine-Induced Protection among Medical Students and Healthcare Workers following Primary Vaccination in Infancy and Rate of Immunity Decline. *Vaccines (Basel).* 2022;10(2):267. doi: 10.3390/vaccines10020267.
20. Bagheri-Jamebozorgi M, Keshavarz J, Nemati M, et al. The persistence of anti-HBs antibody and anamnestic response 20 years after primary vaccination with recombinant hepatitis B vaccine at infancy. *Hum Vaccin Immunother.* 2014;10(12):3731-6. doi: 10.4161/hv.34393.
21. Karabulut N, Çatak Z. Frequency of Hepatitis B Virus, Hepatitis C Virus and HIV Infections in Cannabis and Opioid Addicts. *Viral Hepat J.* 2017;23:26–9. doi: 10.4274/vhd.74046.
22. Zheng Z, Li G, Liao F, et al. Seroconversion of hepatitis B surface antigen among those with previously successful immune response in Southern China. *Hum Vaccin Immunother.* 2021;17(3):845-851. doi: 10.1080/21645515.2020.1801076.
23. Jan CF, Huang KC, Chien YC, et al. Determination of immune memory to hepatitis B vaccination through early booster response in college students. *Hepatology.* 2010;51(5):1547-54. doi: 10.1002/hep.23543.
24. Papadopoli R, De Sarro C, Torti C, Pileggi C, Pavia M. Is There Any Opportunity to Provide an HBV Vaccine Booster Dose before Anti-Hbs Titer Vanishes? *Vaccines (Basel).* 2020;8(2):227. doi: 10.3390/vaccines8020227.
25. Lu CY, Ni YH, Chiang BL, et al. Humoral and Cellular Immune Responses to a Hepatitis B Vaccine Booster 15–18 Years after Neonatal Immunization. *J Infect Dis.* 2008;197(10):1419–26. doi: 10.1086/587695.
26. Garthwaite E, Reddy V, Douthwaite S, Lines S, Tyerman K, Eccles J. Clinical practice guideline management of blood borne viruses within the haemodialysis unit. *BMC Nephrol.* 2019;20(1):388. doi: 10.1186/s12882-019-1529-1.
27. Paul S, Dickstein A, Saxena A, et al. Role of surface antibody in hepatitis B reactivation in patients with resolved infection and hematologic malignancy: A meta-analysis. *Hepatology.* 2017;66(2):379–88. doi: 10.1002/hep.29082.
28. Poola S, Kratzer MK, Sewell K, Tillmann HL. Size Matters! Anti-HBs Titer and HBV Reactivation During Anti-TNF Therapy. *Dig Dis Sci.* 2023;68(12):4511–20. doi: 10.1007/s10620-023-08141-7.
29. Khiangte B, Kothakota SR, Sasidharan M, et al. Hepatitis B Reactivation in Liver Transplant Recipients With Hepatitis B Virus Core Antibody Positive Grafts: a Retrospective Study. *J Clin Exp Hepatol.* 2020;10(6):548–54. doi: 10.1016/j.jceh.2020.05.001.
30. Nasir M, Wu GY. Prevention of HBV Recurrence after Liver Transplant: A Review. *J Clin Transl Hepatol.* 2020;8(2):150–60. doi: 10.14218/JCTH.2020.00003.
31. Yang S, Tian G, Cui Y, et al. Factors influencing immunologic response to hepatitis B vaccine in adults. *Sci Rep.* 2016;6:27251. doi: 10.1038/srep27251.
32. Poorolajal J, Mahmoodi M, Majdzadeh R, Nasseri-Moghaddam S, Haghdoost AA, Fotouhi A. Long-term protection provided by hepatitis B vaccine and need for booster dose: A meta-analysis. *Vaccine.* 2010;28(3):623–31. doi: 10.1016/j.vaccine.2009.10.068.
33. Hudu SA, Jimoh AO, Shinkafi SH, et al. Hepatitis B Vaccine Nonresponse and Associated Risk Factors: Insights From a Cohort Study. *J Trop Med.* 2025;2025:3879562. doi: 10.1155/jotm/3879562.
34. Trevisan A, Giuliani A, Scapellato ML, et al. Sex Disparity in Response to Hepatitis B Vaccine Related to the Age of Vaccination. *Int J Environ Res Public Health.* 2020;17(1):327. doi: 10.3390/ijerph17010327.
35. Ruggieri A, Anticoli S, D'ambrosio A, Giordani L, Viora M. Monographic section The influence of sex and gender on immunity, infection and vaccination. *Ann Ist Super Sanità.* 2016;52(2):198–204. doi: 10.4415/ANN_16_02_11.
36. Naeem SS, Siddiqui EU, Kazi AN, Khan S, Abdullah FE, Adhi I. Prevalence of Hepatitis 'B' and Hepatitis 'C' among preoperative cataract patients in Karachi. *BMC Res Notes.* 2012;5:492. doi: 10.1186/1756-0500-5-492.
37. Khan S, Rafique A, Khizer MA, Zafar O. Frequency of Hepatitis B and C in Patients Undergoing Cataract Surgery in a Tertiary Care Eye Hospital. *Pak Armed Forces Med J.* 2021;71(1):234–7. doi: 10.51253/pafmj.v71i1.4834.
38. Mishra D, Singh H, Gogate P, et al. Prevalence of incidental and total human immunodeficiency virus, hepatitis B and hepatitis C seropositivity among patients posted for cataract surgery at a tertiary care center in India. *Indian J Ophthalmol.* 2022;70(2):400-404. doi: 10.4103/ijo.IJO_1970_21.
39. John TJ, Cooksley G, Chan VF, et al. Hepatitis B vaccine boosters: Is there a clinical need in high endemicity populations? *J Gastroenterol Hepatol.* 2005;20(1):5–10. doi: 10.1111/j.1440-1746.2004.03398.x.
40. Al-Busafi SA, Alwassief A. Global Perspectives on the Hepatitis B Vaccination: Challenges, Achievements, and the Road to Elimination by 2030. *Vaccines (Basel).* 2024;12(3):288. doi: 10.3390/vaccines12030288.

https://doi.org/10.5455/annalsmedres.2025.10.328